Viable Models of Energy Metabolism

Contemplating Uncertainty in Measured Data, Parameter Estimates and Predictions

Ph.D. Thesis

Johannes Hettling

Vrije Universiteit Amsterdam, 2014

Cover: design by Brian McKenna

ISBN 978-94-6259-207-0

Printed by Ipskamp Drukkers, Amsterdam

VRIJE UNIVERSITEIT

Viable Models of Energy Metabolism

Contemplating Uncertainty in Measured Data, Parameter Estimates and Predictions

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duijn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Exacte Wetenschappen
op donderdag 12 juni 2014 om 9.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Johannes Hettling

geboren te Berlijn, Duitsland

promotor:	prof.dr. J. Heringa
copromotor:	dr. J.H.G.M. van Beek

Contents

Chapter 1

Introduction

1.1 Systems Biology

Over the last decade, the term Systems Biology has become pervasive throughout many biological disciplines [Kohl et al., 2010]. Biological research has advanced enormously from a boost in new automated technologies introduced around the turn of the millennium. New high throughput methods allowed for collecting large amounts of biological data, instigating research fields such as genomics, proteomics or metabolomics. At that time, rapid technical progress fuelled expectations that many complex biological processes underlying disease could soon be elucidated.

However, it is apparent that the integration of information about different components of the biological system is inevitable in order to understand how biological functions emerge from the interactions of single components. Rather than examining the characteristics of isolated parts of a cell or organism, structure and dynamics of cellular and organismal function must be investigated to understand biology at a system level [Kitano, 2002a]. Hence, the study of the interplay between different processes which enables biological function in the system as a whole is called Systems Biology. The aim to integrate quantitative knowledge and the emphasis on how biological systems behave and develop over time are major aspects of Systems Biology.

1.2 Computer models in Systems Biology

The information available to understand biological systems is typically incomplete. Many components are not accessible to direct measurement particularly under *in vivo* conditions.

By encoding the available information into computational models, the processes in the biological system can be described to a certain extent. Missing information can then be inferred by analysing the models. A sound understanding of complex biological systems thus requires the integration of experimental and computational research [Kitano, 2002b].

In principle, a computational model in Systems Biology is the representation of a biological process in mathematical terms. There are manifold ways to encode biological information into models and the choice of formalism depends on the biological question at hand and on the availability of measured data on the system. Often, computational models consist of a set of equations describing the quantitative properties of the system. The behaviour of the modelled system over time can be predicted by solving the model equations. Models are usually calibrated and validated on experimental data. Having a calibrated model, possible interventions to the system can be tested *in silico* to prove or reject a certain hypothesis about the modelled system or to predict the effect of certain manipulations (e.g. mutations) to the system. Experimental validation of model prediction yields additional data which is then, in turn, used to improve the computational model. The cycle of iterative model improvement

and experimental validation constitutes a hypothesis driven approach to research in biology [Kitano, 2002b].

1.2.1 Modelling metabolism

Metabolism is the total of biochemical processes within the cells of a living organism. The routes connecting intake and excretion of metabolites serving important functions for the organism are called metabolic pathways. The biochemical conversions between different metabolites are catalysed by enzymes. Individual enzyme reactions in metabolic pathways are in many cases well studied and information about these reactions is available in text books and in public repositories, for instance the metabolic reconstruction of human metabolism [Thiele et al., 2013].

Despite the relatively well-known qualitative aspects of metabolic pathways which is available nowadays, to study the physiology of an organism, quantitative knowledge on the whole system level is desirable. The combination of experimental and computational approaches is essential to gain such knowledge on living cells and organisms.

For the most part in this thesis, we investigate the physiological properties of the processes of energy production by mitochondria in muscle cells. As many biological processes, energy production can be described on different levels of complexity. In some cases, a model representation of a metabolic network describing the metabolites and the conversion rates between the metabolites is enough to study a specific problem. For other applications, such a representation might be too coarse and more detailed kinetic descriptions have to be taken into account, for instance the binding and dissociation constants of a metabolite to a particular enzyme. In this thesis, the following modelling techniques were used to study energy metabolism:

- In the *Bottom-up modelling* approach, detailed models of metabolism are build upon the experimental study of isolated components of the system. Quantitative knowledge on single system components obtained in *in vitro* experiments, such as enzyme turnover rates or molecular binding constants, is integrated on the system level. Mathematical equations are used to describe the kinetic properties of the reactions in the system in detail.

- *Constraint-based modelling* is usually used to estimate the velocity of metabolic reactions (fluxes) in a given system of metabolic reactions subject to a specific objective. The objective is generally the maximisation or minimisation of a certain flux within the system, for instance fluxes of reactions that produce adenosine triphosphate (ATP), a Co-enzyme that carries most energy produced and transported in cells. In constraint based modelling, the metabolic system is characterised by a number of constraints, for instance the constraint of mass conservation for a system in steady-state.

- *Carbon-13 Metabolic Flux Analysis* (^{13}C MFA) is a method to quantify fluxes through metabolic reactions. It is based on the incorporation of stable isotopes

(here ^{13}C) into substrates of metabolism. The isotope distribution in intermediate metabolites is then measured. Analysing the data requires sophisticated computational models and optimisation methods.

1.2.2 Uncertainty of parameters in Systems Biology models

Most computational models in Systems Biology contain parameters which describe physical, chemical or biological characteristics and features within the modelled system. A parameter could for example be the binding constant of a metabolite to an enzyme, the rate of nutrition uptake by a cell or the flux through a certain biochemical reaction. Since model dynamics depend on parameters, models can be calibrated to match experimental data by adjusting their parameter values. This procedure, often called parameter fitting, can be automated using computational routines. In several cases, biological parameters are well determined and experimentally validated whereas quantitative information about other parameters is completely missing.

A famous quote in theoretical physics states 'Give me four parameters, I can fit an elephant. Give me five and I can wag it's tail' [Brown and Sethna, 2003]. The quote manifests a major challenge in modelling (biological) systems: Often, if the parameters in a model are varied, the model is able to describe a variety of different system responses. If parameter values are not known with sufficient precision, model simulations can yield a variety of different outcomes when changing parameter values. Thus, the computational model effectively becomes futile for hypothesis-testing due to its low predictive power. Most models in Systems Biology contain parameters that are unknown or uncertain, because either experimental validation is infeasible or parameters have been determined under different conditions outside the scope of the model (e.g. a reaction flux has been measured under a different temperature in a different species).

Practically all models in Systems Biology are subject to uncertainty in the parameter values which may appear discouraging, since one would prefer a model to be as descriptive and predictive as possible. However, it is possible to estimate the effect of uncertainties in parameters on the uncertainty of the model's predictions such that the predictions can be more realistically interpolated, making the model more valuable. The evaluation of uncertainty of a model response with respect to changes in parameter values has been coined as parametric sensitivity analysis [Perumal and Gunawan, 2011].

In [Brown et al., 2004] and [Gutenkunst et al., 2007b] a method to analyse parameter sensitivity in Systems Biology models is introduced in which the multidimensional parameter space is extensively explored using a random walk. The approach is based on the idea that more than one valid parameter combination should be used to extract predictions from a model. To this end, whole sets of different parameter combinations that can describe the experimental data (parameter ensembles) are sampled in a random walk and used to generate model predictions. The contribution of single

parameters or combinations of parameters that cause uncertainty in model predictions can thereby be identified. By simulating an entire parameter ensemble, it is possible to obtain the uncertainty of model predictions with respect to uncertainties in the parameters directly.

In this thesis, the theoretical framework of Brown and Gutenkunst is frequently used in order to explore parameter sensitivities in models of energy metabolism. Emphasis is placed on analysing the models' sensitivity spectra with the aim to improve the work-flow of model analysis and data fitting to yield model predictions that are statistically more sound. To improve the predictive capacity of the models, quantitative *a priori* information on parameter values is explicitly taken into account.

1.3 Viable models of energy metabolism

Energy metabolism comprises production, transport and consumption of energy carrying compounds and metabolites. We investigated various aspects of energy metabolism, for example the energy production by mitochondria and the subsequent transport to energy consuming sites and the citric acid cycle and glycolysis which are both fundamental metabolic pathways in living organisms. Throughout this thesis, mathematical models that describe these processes have been used.

We believe that taking into account error in measured data, parameter values and model predictions is not only helpful practice to bound the uncertainty in extracting relevant information about biological systems, but moreover should be an elemental part of any analysis with mathematical models. We therefore describe approaches that consider various sources of uncertainty in order to design robust models that are feasible, useful and thus *viable* for the describing biological reality.

A general introduction on how metabolic systems can be modelled quantitatively is provided in chapter 2. Relevant techniques of modelling are described in detail. Further, in this chapter the creatine kinase system which is the subject of interest throughout most chapters in this thesis is introduced. The creatine kinase enzyme interconnects energy producing and energy consuming sites in (among others) muscle cells. To which extent the enzyme is involved in the energy transport process is subject to debate [Beard and Kushmerick, 2009]. The 'phosphocreatine shuttle hypothesis' states that most energy in muscle cells produced by the mitochondria in form of ATP is not transported to the energy consuming myofibril as such [Bessman and Geiger, 1981]. Instead, creatine kinase mediates the transport of energy via the high-energy metabolite phosphocreatine (PCr) [Bessman and Geiger, 1981]. We used a simple model which captures the essentials of the creatine kinase system to investigate the phosphocreatine shuttle hypothesis in cardiac muscle cells. Previous analysis had shown that the model contradicts an obligatory transport via PCr [van Beek, 2007]. In chapter 2, basic sensitivity analysis is applied to the model to investigate the robustness of the model response to changes in parameter values. Results show that even moderate changes in parameter values can yield a wide range of model

predictions. The conclusion of this analysis is that in order to challenge the phosphocreatine shuttle hypothesis using the simple model and the available data, a more sophisticated approach is desirable.

This analysis is described in chapter 3. The model of the creatine kinase system [van Beek, 2007] is calibrated with measurements of the response time of oxygen consumption in isolated perfused rabbit hearts under different conditions. In order to account for inaccuracies in parameter values, a random walk through parameter space was performed [Gutenkunst et al., 2007b]. As a result, ensembles of different parameter sets were generated. The ensembles were used to define confidence bounds on model predictions. Because the model is sensitive to variations of parameters as shown in chapter 2, prior information on parameter values is explicitly taken into account in the random walk. As prior information, measurements of the model parameters and their standard measurement errors were collected from the scientific literature. By incorporating the *a priori* knowledge on model parameters, the variation in model predictions could be significantly narrowed. The resulting model predictions are statistically sound because the possible measurement error in all data used in the analysis is taken into account. With the framework employed in chapter 3, the contribution of PCr to energy transport in heart muscle is predicted to be relatively low and most of the energy is transported as ATP, which contradicts the phosphocreatine shuttle hypothesis. Since the main function of the creatine kinase system does not appear to be the transport of energy from mitochondria to myofibrils, other known and possible functions were tested using the model. Simulations suggest that creatine kinase is responsible for the temporal buffering of energy resources. In the heart muscle, large amounts of energy are stored in form of PCr which can be released very quickly whenever the energy demand of the heart increases (e.g. when the heart rate increases).

Creatine kinase and its product PCr does not only play a role in heart muscle, but also in the energy metabolism in skeletal muscle tissue. In chapter 4, a whole body model for energy conversion is used to simulate energy conversion in a cyclist in a mountain time trial of the Tour de France. While the whole body model describes heat production and transport, the biochemical events during muscle contraction in the leg are described with the model used in chapters 2 and 3. Since energy demand during cycling is characterised by large bursts of periodic muscle work, the energy buffering functions of the creatine kinase system is investigated *in silico*. Since biochemical properties in the human leg muscle differ from the previous analyses in cardiac muscle, a new set of model parameters describing skeletal muscle energy turnover was assembled. The modified model was validated using data from human experiments on a bicycle ergometer in order to predict energy buffering during bicycle racing. It is shown that creatine kinase is essential for the allocation of energy during the bicycle race and that the buffering of mitochondrial energy production is lost if creatine kinase was inhibited.

In chapter 5, the ensemble modelling approach from chapter 3 is applied to carbon transition networks. A detailed model of carbon transitions in the citric acid

cycle, an essential pathway in muscular energy metabolism, was used to analyse data from carbon tracer experiments in porcine cardiac muscle. The approach is based on the incorporation of substrates which contain labelled carbon atoms into the energy metabolism of the heart *in vivo*. Nuclear magnetic resonance (NMR) spectra of intermediate metabolites can be obtained which allow the exploration of how fast the labelled carbon atoms travel through the pathway. As a result, the velocity of metabolic reaction fluxes can be quantified by analysing the data with a detailed model of carbon transitions within the citric acid cycle. The quantified rates of the metabolic reactions yield valuable information about the physiology of the heart in health and disease. As with metabolic models used in the previous chapters, carbon transition networks are subject to parameter uncertainty. In chapter 5, a large NMR data set from isotope labelling experiments under various experimental conditions was used. Since the noise level in the available data was very high, the determination of uncertainty in model predictions was in this case particularly crucial in order to still be able to extract valuable physiological information from the data. By sampling parameter ensembles that agree with the NMR data, two model parameters that can be quantified within reasonable precision limits could be identified. Also prior information from literature was used to narrow the confidence limits on the parameter estimates. The performance of the method was tested by estimating myocardial oxygen consumption from the two model parameters that were quantified. The estimates were compared to measurements that were performed independent from the carbon tracer method and therefore were independent from the model estimates. The agreement between both methods was reasonable, which validates the approach introduced in this chapter. It is therefore shown that combining prior information on the velocity of metabolic fluxes with a random walk to generate parameter ensembles is a fruitful approach in order to quantify metabolic fluxes.

Carbon labelling experiments are not always feasible in any organism or form of tissue. A different approach to quantify the rates of metabolic reaction fluxes is Flux Balance Analysis (FBA). In FBA, information about reaction fluxes entering and leaving a reaction system is integrated with a detailed stoichiometric description of all intermediate biochemical reactions within that system. Several biophysical assumptions can then be used in order to gain information about the fluxes within the system. A central assumption in this modelling framework is the steady-state assumption: if no (or a constant) outside stimulus is applied to a biochemical reaction system, the system will reach a state at which the concentrations of the intermediate metabolites will not change. Using the steady-state assumption, the reaction fluxes in the system of interest can be quantified using a mathematical optimisation approach called linear programming, which is the central part of FBA. There is a wealth of detailed stoichiometric information about reaction systems in various organisms which can be used to build computational models for FBA. A listing of all relevant metabolites and reactions is referred to as a metabolic reconstruction and over the last years, an increasing number of metabolic reconstruction databases became publicly available. Chapter 6 of this thesis aims to connect existing metabolic

reconstruction databases with analysis tools that can perform FBA. The software tool BiGGR was developed which allows for the assembling of computational models from several metabolic reconstruction resources and the subsequent analysis using the mathematical algorithms required for FBA. Further, the software tool introduced in this chapter allows for visualisation of selected parts of the assembled networks and their estimated reaction fluxes. It is important to note that a preliminary version of the BiGGR software was already introduced in chapter 4 (section 4.8.1). BiGGR was therafter re-implemented and new functionality was added. The new version features a flux estimation procedure which contemplates uncertainty in measured input data and resulting flux estimates by combining FBA with an ensemble modelling approach. A general description of BiGGR and its functionality can be found in the first part of chapter 6. The second part of chapter 6 consists of the BiGGR 'package vignette', which gives technical details and a step-by-step example of a flux estimation within BiGGR using a model of human brain metabolism.

To conclude, chapter 7 discusses all results obtained in this thesis and puts the obtained results into perspective.

Chapter 2

Robust modelling, measurement and analysis of human and animal metabolic systems

van Beek, J.H.G.M., Hauschild, A., Hettling, H. and Binsl, T.W.(2009) Robust modelling, measurement and analysis of human and animal metabolic systems. *Philosophical Transactions of the Royal Society A*, 367:1971-1992

2.1 Abstract

Modelling human and animal metabolism is impeded by the lack of accurate quantitative parameters and the large number of biochemical reactions. This problem may be tackled by: (i) study of modules of the network independently; (ii) ensemble simulations to explore many plausible parameter combinations; (iii) analysis of 'sloppy' parameter behaviour, revealing interdependent parameter combinations with little influence; (iv) multiscale analysis that combines molecular and whole network data; and (v) measuring metabolic flux (rate of flow) *in vivo* via stable isotope labelling. For the latter method, carbon transition networks were modelled with systems of ordinary differential equations, but we show that coloured Petri nets provide a more intuitive graphical approach. Analysis of parameter sensitivities shows that only a few parameter combinations have a large effect on predictions. Model analysis of high-energy phosphate transport indicates that membrane permeability, inaccurately known at the organellar level, can be well determined from whole-organ responses. Ensemble simulations that take into account the imprecision of measured molecular parameters contradict the popular hypothesis that high-energy phosphate transport in heart muscle is mostly by phosphocreatine. Combining modular, multiscale, ensemble and sloppy modelling approaches with *in vivo* flux measurements may prove indispensable for the modelling of the large human metabolic system.

2.2 Introduction

Metabolism consists of the biochemical reactions taking place in the body. It is the counterpart in the molecular domain of the energy transformations and signalling taking place in the cells and the body as a whole. Metabolism is therefore of enormous importance to human and animal physiology and pathology. For instance, instead of studying the tremors in patients with Parkinson's disease, it may often be more efficient to study the synthesis of the neurotransmitter dopamine, which is involved in this disease, in cultures of human cells.

One difficulty in the study of human and animal metabolic networks is that the total system is very large and strongly interconnected. In addition, some metabolic pathways overlap extensively [Ma et al., 2007] and hence influence each other's behaviour. This causes difficulties in studying them separately. Qualitatively, much is known about metabolism, from micro-organisms to higher life forms. This means that we often know which metabolites take part in metabolism and how they are connected via reactions in the metabolic pathways. For the most part, biochemical reactions are very similar throughout all kingdoms of life, which to some extent may simplify the modelling process: many reactions and parts of pathways can be transferred from models of one organism to another. The possibility for such simplification is limited because the regulation of metabolic pathways and the mechanisms of regulation of enzymes vary between species and must in many cases be adapted to build accur-

ate models for other species. Although there is a certain universality in metabolic systems, even between bacteria and higher organisms, regulation and quantitative behaviour of the models has to be adapted and fine tuned for each organism. This often means quantitative changes in pathways that remain qualitatively similar.

In contrast to qualitative properties, little is known about the quantitative parameters of these metabolic systems, especially not under *in vivo* conditions. Parameters are not yet available or perhaps not even measurable. This lack of accurate kinetic parameters is an important impediment for modelling human and animal metabolism. Here, we describe strategies for robust modelling of metabolic systems, which means that we deal with the limited accuracy of metabolic parameters. We try to obtain insight and useful predictions in the face of limited knowledge about the metabolic system.

The problem of insufficient quantitative knowledge may be tackled by: (i) modularisation of the network to study parts of the system one by one and independently; (ii) making use of ensemble simulations to explore the range of plausible parameter values in multidimensional parameter space; (iii) determining combinations of parameters that have either large or very small effects on model predictions when the parameters are changed in a correlated way ('sloppy' parameter behaviour, see below); (iv) multiscale analysis that combines information from the molecular, biochemical and organ levels; and (v) measuring metabolism under *in vivo* conditions, providing important data on biochemical network function as a whole.

While these are important strategies to bridge the gap between the qualitative knowledge of metabolism and a comprehensive quantitative analysis of the human and animal metabolic system, another important gap exists between mathematical modellers and experimentalists. Many experimental scientists have difficulties in understanding mathematical and computational models. Such models of biochemical systems are presently often formulated as systems of ordinary differential equations (ODEs). Each differential equation represents the rate of change of a metabolite pool caused by the chemical reactions transporting mass into and out of that pool. For many experimental scientists, this approach is probably not appealing because it is not in line with their intuition. For this reason, we will discuss a more graphical approach, provided by Petri nets, which is accurate enough for many purposes and close in structure and functioning to the metabolic pathways that it represents. This approach may help experimental scientists to formulate and use models of metabolic pathways more efficiently.

2.3 From qualitative knowledge on metabolism to large-scale quantitative modelling

Textbooks of biochemistry and an enormous number of publications in the literature describe the types of molecules that have been discovered in the bodies of humans and

animals. Many biochemical reactions, catalysed by enzymes, connect the metabolites as they react with each other or are formed from each other. Sequences of enzymatic reactions are connected in metabolic pathways. The Boehringer chart, which shows a large number of reactions and their connections, has decorated the walls of biochemical laboratories for many decades now. Recently, several metabolic system databases have been assembled. The KEGG [Kanehisa et al., 2012], Reactome [Croft et al., 2011] and PathwayCommons [Cerami et al., 2011] databases, for instance, list a great number of reactions in many pathways. Bernard Palsson and colleagues built a reconstruction of human metabolism *in silico* [Duarte et al., 2007]. A second reconstruction of human metabolism was published [Ma et al., 2007]. In addition, a database exists that contains the small metabolites found in the human body, the Human Metabolome database [Wishart et al., 2007].

There is a remarkable difference between metabolism as portrayed in textbooks, where dominant fluxes (rates of flow) in major pathways are treated, and the connectivity of metabolism as found in an *in silico* reconstruction of human metabolism [Duarte et al., 2007]. The textbook version of the tricarboxylic acid (TCA) cycle, also known as the Krebs cycle, usually describes the TCA cycle as an ordered series of reactions and does not emphasise the many existing side reactions. The textbook mental picture of the TCA cycle is almost that of a distinct module. A close look at the comprehensive *in silico* description shows that the TCA cycle intermediates are substrates and products of numerous side reactions, which connect the TCA cycle intermediates extensively with the rest of metabolism. Kinetically, it may be true that the TCA cycle reactions dominate, but in reality its intermediates are part of an extensive metabolic network. To understand this complex densely interconnected metabolic system under all conditions, it would be very useful to have a dynamic model of metabolism that addresses the full network connectivity. If such a model proves to predict metabolic network function reasonably, it could in the future also help to regulate metabolism under disease conditions.

The lack of accurate kinetic equations and parameters is clear, but even if all the parameters are known with reasonable accuracy, this does not mean that metabolite levels or metabolite fluxes can be calculated with good precision. Even for a relatively 'simple' organism such as yeast, with kinetic measurements collected to characterise all the enzymes of the glycolytic pathway [Teusink et al., 2000], it turned out that the prediction of pathway flux and metabolite levels in the pathway was of rather limited accuracy.

Given that there is so much information available on the connectivity of metabolism, the question may be posed whether it is timely to start building quantitative and dynamic models of human and animal metabolism. The answer is clearly positive. On the one hand, much experimental data have been gathered, but on the other, understanding of the metabolic system is still very limited. Gradually, by trial and error, building a more comprehensive model of the human metabolic system may help to integrate and understand the enormous amount of experimental information on human and animal metabolism. Building a valid model will be possible only by mak-

ing predictions on new experiments and correcting the model if the predictions prove wrong. In this way, dynamic models of metabolism, containing many pathways and metabolites, may prove their value to integrate the extensive knowledge on thousands of metabolites, with reactions catalysed and regulated by thousands of gene products. It is a major challenge to systems biochemistry to develop such a quantitative description of human metabolism, which integrates all biochemical and molecular biological knowledge and explains most experimental findings. However, because the mind of an ordinary scientist is, in general, insufficient to contain all the information available on the human biochemical system, robust modelling and analysis techniques making use of all available information are necessary. This becomes even more desirable when analysing and predicting metabolic system behaviour during a disease process or an experimental or therapeutic intervention. Because it is inevitable to face the challenge of developing models of human and animal metabolism, there is a need to deal with the large scale of the system and in particular with the substantial number of imprecisely known kinetic parameters.

2.4 Strategies for quantitative modelling of metabolic systems

There exist different strategies to deal with the large scale of the human biochemical system and with the incomplete quantitative knowledge of the parameters characterising metabolic processes. Some robust approaches are as follows.

- Modularisation of the network to study parts of the entire system in detail, which can be experimentally 'isolated'. Thereafter, the whole system is assembled from such well-characterised parts. This approach may not be possible in all or even most cases, but is rewarding if feasible because the biological function of a small group of molecules can be accurately characterised.
- Application of ensemble simulations to explore the range of plausible parameter values. To this end, many simulations are done with different parameter sets that cover the plausible part of multidimensional parameter space, taking correlation between the parameters into account. Importantly, such simulations can also be used to explore the effect of measurement noise on the confidence regions of parameters estimated from experimental data.
- Determination of parameter combinations that are shown by analysis to have large or small effects on model predictions when changing in a correlated way. The former represent 'stiff' directions in parameter space while the latter depict 'sloppy' directions in parameter space. The approach of finding those parameter combinations has been termed 'sloppy modelling' [Brown et al., 2004, Gutenkunst et al., 2007a,b]. This does not mean sloppy work by the scientists who design the model, but emphasises that some of the combinations of parameters

in the model exert a weak influence on the model prediction. Often strong correlations within groups of parameters are discovered, which together define the 'stiff' and 'sloppy' directions.

- Application of multiscale analysis, which means that measurements at various aggregation levels are taken into account: information measured at the level of molecules, biochemical pathways, organelles, cells and organs can thus be combined. For instance, not only the properties of enzymes or isolated mitochondria are incorporated in the analysis, but also the measured response of a whole pathway or network of connected enzymes and organelles in the cell. For instance, the time course of metabolite levels or metabolic flux (rate of flow) in the pathway in response to stimulation of the pathway provides data at the pathway level. Experimentally, this may be accomplished by measuring the time course of adaptation to altered cellular workloads, such as increased muscle contraction frequency, neural firing rate or secretion of hormone. Yet another example is the measurement of metabolic fluxes in specific pathways under various steady-state conditions in relation to the metabolite levels in the pathway.

- Measuring metabolism under *in vivo* conditions. We describe computational methods to quantify metabolic fluxes from experimental measurements under *in vivo* conditions. This can, among other ways, be done by providing substrates for metabolism that are labelled with stable isotopes, and subsequently measuring the incorporation of the label in the network's metabolites. The latter type of measurements define the operation of the metabolic system under the *in vivo* working conditions and thereby provide important information on the functioning of the intact living system.

We will start with a description of flux measurements because it provides a simple example of a model of a metabolic system. Such measurements provide an important input for the multiscale approach and show examples of the 'sloppiness' of parameter space.

2.5 Measuring metabolic fluxes *in vivo* with stable isotopes

Metabolic fluxes reflect cell function and dynamic adaptation of living organisms to their environment. Given the incompleteness of available kinetic parameters to dynamically simulate metabolism, it is very useful to determine the reaction rates in metabolic networks experimentally. Common isotope experiments are, for instance, done in experimental animals, human beings and micro-organisms, and entail the infusion of isotope-labelled substrate into the metabolic system. These isotopes are distributed by metabolic fluxes among the metabolites, usually until the isotopic

steady state is reached. The amount of isotope incorporation is measured by nuclear magnetic resonance spectroscopy (NMR(S)) or mass spectrometry (MS) in chosen metabolites. This allows us to quantify metabolic fluxes via computational modelling and analysis, even in cases where kinetic parameters are missing [Kelleher, 2001, Wiechert, Sherry et al., 2004, Kelleher, 2001, Antoniewicz et al., 2007]. However, measurements in tissue samples taken after a steady state of label incorporation is reached only give insight into the relative fluxes in different pathways. Especially in micro-organisms, substrates enriched with stable isotopes are often given for a long time until incorporation in small metabolites and proteins has reached a steady state. Hence, experiments become long and costly. By contrast, if steady states are short, as in humans and animals, this common approach will not be accurate due to continuous changes in the system. Yet another obstacle is destructive measurement techniques such as MS, which do not allow us to measure the time course of isotope incorporation in the same small tissue region. It is impossible to take multiple samples if there is little material available, as may be the case with tissue biopsies and in cell cultures.

Therefore, we have recently developed a method to define the time course of isotope incorporation and to quantitate metabolic fluxes by analysing a single sample taken at a single time point [van Beek et al., 1998, 1999]. For this method, it is essential that the sample is collected before the attainment of a steady state of isotope incorporation. Despite taking only a single measurement, several metabolic fluxes can be quantified by computational model analysis of the measured isotope incorporation in the metabolites in the tissue sample. This new experimental protocol to quantify metabolic fluxes is named 'labelling with isotope for pre-steady-state snapshots' (LIPSSS). The isotope incorporation data are analysed with a computer package called FLUXSIMULATOR [Binsl et al., 2010b]. The LIPSSS protocol has in common with other isotope methods that a substrate labelled with isotopes is infused into the metabolic system of interest. However, in contrast to many other experimental protocols, in the LIPSSS protocol a single snapshot of stable isotope incorporation in metabolites is taken before the steady state of isotope incorporation is reached. This makes experiments shorter and cheaper. Furthermore, it facilitates *in vivo* experiments in mammals where metabolic steady states persist only briefly.

The quantification of metabolic fluxes from single time-point measurements is possible since the time course of isotope incorporation into metabolites depends on (i) the amount of isotopic material given, (ii) the sizes of the metabolite pools through which the label flows, with larger transit times of label caused by larger pool sizes, and (iii) the flux sizes that transport the isotopic material between the metabolites, with larger fluxes leading to shorter transit times for given metabolite pool sizes. Hence, at every point in time during the dynamic phase before isotope steady state is reached, each metabolite pool contains a characteristic composition of different isotope combinations, called isotopomers. These isotopomers within the metabolite pool determine distinct peaks measured with NMRS or MS. These peaks are calculated from models that contain the carbon transitions, usually already known from earlier experiments. The connectivity of metabolites at the carbon level is found in the

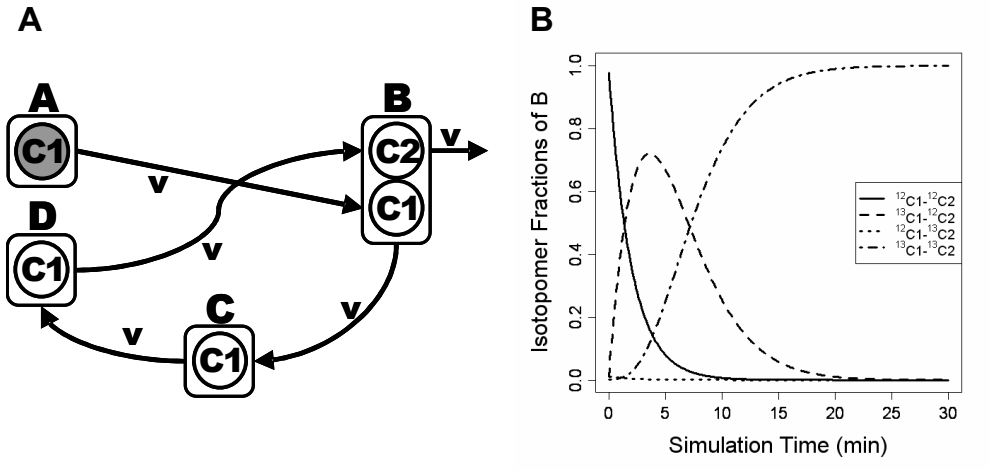

Figure 2.1: (a) Example of an artificial metabolic network model illustrated by its carbon transition network representation. The model consists of four metabolites A, B, C and D. To make the network as simple as possible to illustrate the principles, A, C and D are assumed to be compounds containing a single carbon only, while B is assumed to contain two distinct carbons, respectively. Metabolites are each represented by a rectangle, which contains one or more carbons represented by circles. The grey circle in metabolite A represents a ^{13}C isotope, which starts to enter the system after ^{13}C-enriched substrate is provided at t=0, and the white circles represent ^{12}C present at 99% natural abundance. The carbon atom in A enters metabolite B in the C1 position and after cycling through the system also provides the C2 atom of metabolite B. Isotopes in two different positions in the same molecule usually yield different signals in NMRS. Arrows indicate for all carbon atoms their origin and target position in the metabolites during carbon transitions mediated by the chemical reactions. (b) All possible isotopomers ($^{12}C1$-$^{12}C2$, solid line; $^{13}C1-^{12}C2$, long dashed line; $^{12}C1-^{13}C2$, short dashed line; $^{13}C1-^{13}C2$, dot-dashed line) of metabolite B and their time course after metabolite A is enriched with ^{13}C at t=0. The time course of fractions of isotopomers is dependent on the flux v in the metabolic pathway, and reversely the flux v may therefore be quantified from the isotopomer content.

scientific literature or databases.

An example of isotopomer fractions evolving dynamically over time in a simplified metabolic pathway is given in Figure 2.1. The measurements are analysed computationally with our computer package FLUXSIMULATOR to quantitate the metabolic fluxes. To this end, the experimental protocol is continuously simulated by FLUXSIMULATOR while different parameter estimates are explored by an algorithm to optimise the fit of the model to the measured data. In each step, using the simulated isotopomer fractions of the FLUXSIMULATOR simulation the corresponding nuclear magnetic resonance (NMR) multiplet intensities are calculated. The parameter estimates are provided by a nonlinear parameter optimisation routine [Nelder and Mead, 1965] and the calculated multiplets are compared in each optimisation step with the experimentally measured multiplets via the sum-of-squares criterion until a reasonable match between multiplets and model prediction is found (Figure 2.2).

FLUXSIMULATOR enables the user to implement metabolic models and to perform

flux estimation in an efficient and user-friendly manner. In many existing simulation software packages, metabolic models are specified directly by their mathematical representation via ODEs. By contrast, FLUXSIMULATOR uses a straightforward specification of the metabolic pathway via three plain text files. Based on this specification, the entire system of ODEs is assembled and simulated automatically [Binsl et al., 2010b]. This enables biomedical researchers to easily specify models that consist of hundreds of equations in a more flexible way than the error-prone implementation of the ODEs by hand. To illustrate the usefulness of automatic model assembly, the ODEs representing the dynamics of the simple metabolic model in Figure 2.1 are given below:

$$\frac{dA_{^{12}C1}}{dt} = \frac{dA_{^{13}C1}}{dt} = 0 \tag{2.1}$$

$$\frac{dB_{^{12}C1^{12}C2}}{dt} = \frac{(A_{^{12}C1}D_{^{12}C1} - B_{^{12}C1^{12}C2})v}{[B]} \tag{2.2}$$

$$\frac{dB_{^{13}C1^{12}C2}}{dt} = \frac{(A_{^{13}C1}D_{^{12}C1} - B_{^{12}C1^{12}C2})v}{[B]} \tag{2.3}$$

$$\frac{dB_{^{12}C1^{13}C2}}{dt} = \frac{(A_{^{12}C1}D_{^{13}C1} - B_{^{12}C1^{13}C2})v}{[B]} \tag{2.4}$$

$$\frac{dB_{^{13}C1^{13}C2}}{dt} = \frac{(A_{^{13}C1}D_{^{13}C1} - B_{^{13}C1^{13}C2})v}{[B]} \tag{2.5}$$

$$\frac{dC_{^{12}C1}}{dt} = \frac{(B_{^{12}C1^{12}C2} + B_{^{12}C1^{13}C2} - C_{^{12}C1})v}{[C]} \tag{2.6}$$

$$\frac{dC_{^{13}C1}}{dt} = \frac{(B_{^{13}C1^{12}C2} + B_{^{13}C1^{13}C2} - C_{^{13}C1})v}{[C]} \tag{2.7}$$

$$\frac{dD_{^{12}C1}}{dt} = \frac{(C_{^{12}C1} - D_{^{12}C1})v}{[D]} \tag{2.8}$$

$$\frac{dD_{^{13}C1}}{dt} = \frac{(C_{^{13}C1} - D_{^{13}C1})v}{[D]} \tag{2.9}$$

The metabolite symbols $A - D$ in these equations give the fractions of these pools taken up by the isotopomers given by the subscript. The metabolite symbols $A - D$ in square brackets indicate the amount per gram of dry tissue mass. The time course of the labelled substrate for the pathway (here, A) is determined by the experimenter who designs the label infusion protocol. Equation 2.1 indicates that the experimenter has arranged to keep the labelling state of pool A, which is infused, constant after stepping to a new value at $t = 0$. The flux v in $\mu mol * min^{-1} * g^{-1}$ of dry mass of tissue is constant throughout the circular pathway. This presents a simple model representation of a metabolic cycle, similar to, for instance, the TCA cycle.

It is of course not very difficult to implement the ODEs for the system in Figure 2.1 by hand in a computer program. However, this becomes a tedious task for

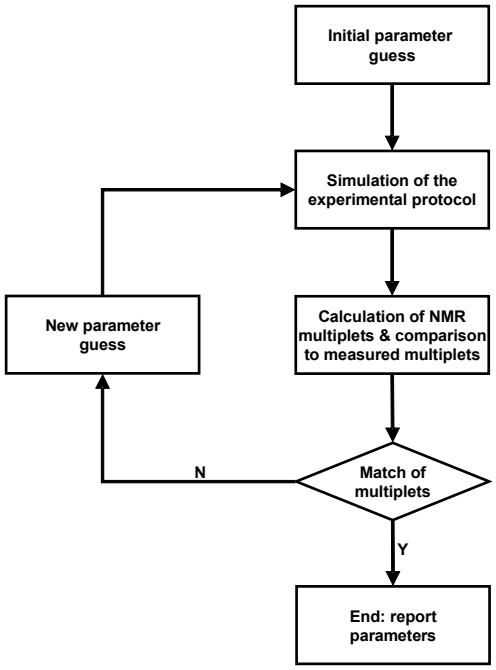

Figure 2.2: Flowchart of the algorithm used for parameter estimation.

extensive systems with metabolites containing more than two carbon atoms [Binsl et al., 2010b]. In particular, for each metabolite containing five carbon atoms, such as α-ketoglutarate in the TCA cycle and glutamate, 32 differential equations must be added to the ODE system. In FLUXSIMULATOR, these equations, similar to equations 2.1-2.9, are assembled based on simple textual information. In a text file, the brief statement A1:B1 D1:B2, for instance, indicates that atom 1 of metabolite B originates from atom 1 of metabolite A, and atom 2 of metabolite B originates from atom 1 of metabolite D.

FLUXSIMULATOR is implemented as a computer package realized in the 'R' programming language and environment, which is available for the most common operating systems such as MS WINDOWS, UNIX and Mac OS. Owing to the on-the-fly C code generation of the model, followed by compilation and back loading of the dynamic library, the simulation time for a model consisting of approximately 200 stiff ODEs is of the order of $0.1s$ on a $3GHz$ Pentium-based desktop computer. Parallelisation and execution on a cluster computer will enhance the computational speed further. 'Embarrassingly parallel' execution to analyse multiple datasets each on a single processor was easily accomplished. Parallelisation of parameter optimisation on one dataset has not been implemented yet but would be straightforward using, for instance, the R package snowfall.

The approach of performing a LIPSSS experiment in combination with analysis

and flux estimation with the FLUXSIMULATOR computer package and nonlinear parameter optimisation was experimentally validated. To that end, the flux in the TCA cycle and exchange with the directly related amino acids glutamate and aspartate was determined in tissue samples from porcine cardiac tissue *in vivo*. The TCA cycle flux represents the final pathway of aerobic energy conversion, tightly coupled to oxygen consumption in the mitochondria. Excellent correlation (r=0.90) was found between flux in the TCA cycle calculated from NMR measurements on extracts of the cardiac tissue samples after $5.5min$ infusion of ^{13}C labelled acetate in the coronary artery and "gold standard" measurements of oxygen consumption on the heart *in situ*.

Experimental protocols for isotope labelling experiments were designed based on computer simulation with the FLUXSIMULATOR package. Various experimental protocols, e.g. infusion time of the isotope-labelled substrate or the isotopic composition of the substrate, were simulated. The simulation shows the dynamic development of isotopomer fractions in one or more metabolite pools, and in that way demonstrates how absolute isotopomer composition and, perhaps equally important, the ratios of various isotopomer fractions at a certain point in time depend on the metabolic flux (Figure 2.1b). It is also possible to add realistic measurement noise to the simulated values at a certain point in time and then to simulate the process under more realistic conditions. This gives information about the model's complexity and parameter identifiability, and makes it possible to choose the experimental protocol with the most accurate and comprehensive measurement results.

Care must be taken that the model of the metabolic pathway and carbon distribution routes is compatible with the organ and cell type studied. The model for the cardiac study above was in particular designed and tested for heart muscle *in vivo*. Although the metabolic pathways of the TCA cycle and related amino acids incorporated in this model are almost universally found in other organs and cell types, the activity of various anaplerotic pathways varies under experimental conditions. Therefore, the models of the metabolic pathways and of carbon distribution must be adapted to accurately reflect the precise cell type under study. It is desirable that the model's suitability to accurately quantitate metabolic fluxes is examined on an organ-by-organ and cell-type by cell-type basis.

Using FLUXSIMULATOR, ensemble simulations to explore the ranges of parameter values (see section 2.9) or the assessment of sloppy parameter sensitivity patterns and spectra (see section 2.7) are easy. The analysis of parameter combinations by simulation enables the user to decide which parameter combinations can be estimated precisely and which are imprecise and can potentially be fixed at certain values or kept free to take care of sloppy directions in parameter space [Gutenkunst et al., 2007a,b].

2.6 Modelling metabolic systems modularly

To analyse complex metabolic systems, it is most useful if one can divide large networks into modules that can be isolated conceptually and experimentally to be studied

independently in detail. A module is an independently operable unit that is part of a whole. An example of the modular approach is the experimental isolation of the adenine nucleotide-creatine phosphate module in muscle metabolism [van Beek, 2007, Van Beek, 2008]. Although this module is small, not all the parameters of its components were reliably known *a priori* from the molecular and organellar level. The remaining parameters were estimated by combining information from the molecular and the whole-system levels (see section 2.8).

Models such as the carbon transition network (CTN) in Figure 2.1 are usually also part of a bigger system. If the fluxes in the pathway of Figure 2.1 dominate strongly over those entering and leaving via side paths, then it is possible to study this model as a module. It is probably difficult to divide the metabolic system completely into modules that can be studied independently, because metabolic pathways overlap and interact extensively. However, if it is possible to study a certain part in isolation, this makes it possible to define the function of that part of the system accurately. Connecting the modules again to form the whole system is an ideal strategy, which we must fear is only partially feasible.

2.7 Sloppy parameter behaviour

Quantitative mathematical models of metabolic or signalling pathways usually contain tens or even hundreds of parameters describing the kinetics of molecular interactions within the system. Sufficiently precise parameters are essential for useful predictions based on the model. Experimental determination of all kinetic parameters of the system, however, often proves to be difficult or even impossible. This is either because not all components can be isolated or because isolation procedures damage the component. Furthermore, intracellular operating conditions are virtually impossible to reproduce *in vitro*. Hence, the estimation of unknown parameters poses a major challenge in mechanistic modelling of biological processes.

The total set of unknown model parameters can be constrained collectively by fitting different parameter combinations to experimental data. This is accomplished by minimising a cost function, e.g. a least-squares distance measure between model prediction and experimental data points, which reports how well a set of parameters fits the available data. In many cases, a variety of different parameter combinations agree with similar cost with the available data. One might then often hear the criticism that the model is over-parametrised and that parameters cannot be determined with sufficient accuracy. This may, at first glance, seem to disqualify such an attempt to characterise the parameters, but we will show below that it is possible to live with uncertainty in some combinations of parameters and still obtain very useful predictions or useful quantification of some of the parameters.

It turns out that there are sloppy parameter combinations that do not significantly alter the simulation outcome of a model. System parameter combinations varying along sloppy directions in parameter space are therefore not well constrained

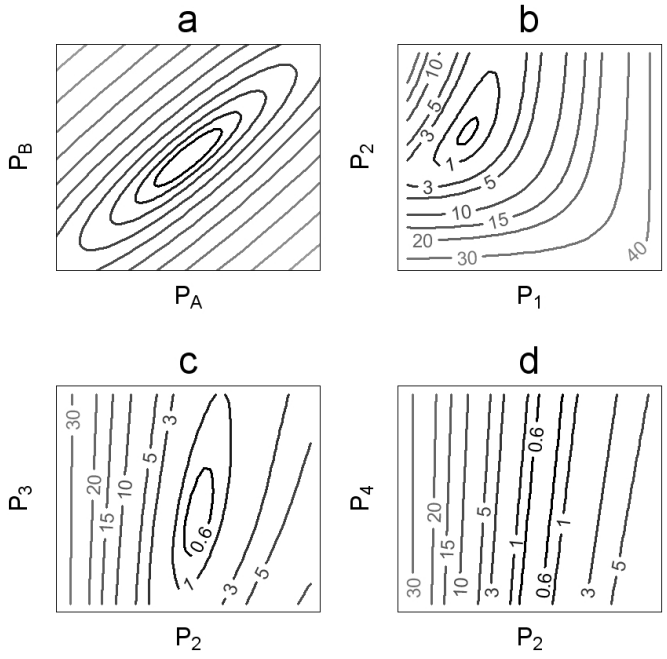

Figure 2.3: Sloppy parameter space. On the ordinate and abscissa are two model parameters. The contour lines give locations in parameter space with equal deviation from values predicted from the reference parameter set that best describes the available experimental data. This deviation is quantitated as the sum of squares (SSQ) between the predicted values for the indicated parameter location and the prediction from the reference dataset. (a) A linear model. (b-d) The results for a single nonlinear carbon transition model with approximately 50 carbon transitions and five parameters $P_1 - P_5$. The contours in (b-d) reflect equal SSQ for nine NMR data points reflecting the isotopomer composition of metabolites (Figure 2.1).

by experimental data. Other directions (stiff directions) can be well constrained, with small parameter variations yielding significant changes in predictions. An example of sloppy parameter behaviour is given in Figure 2.3.

The contours in this plot give the locations of equal deviation of model prediction from the result yielded by the best ('reference') parameter set for the data point. Figure 2.3a represents a model that is linear in the parameters P_A and P_B. Figure 2.3b-d represents a model of carbon transitions with five parameters, which is an extended version of the model of Figure 2.1. Note that parameters may show strong dependencies: in Figure 2.3a, a joint increase in P_A or P_B results in small changes in model prediction. This constitutes the long axis of the ellipse in Figure 2.3a, which is termed a sloppy direction in parameter space. In the case where P_A changes in the direction opposite to that of P_B, the prediction by the model changes briskly along the short axis of the ellipse. This is termed a stiff direction in parameter space. In the nonlinear case of Figure 2.3b the contours are not ellipsoidal, although the inner

contours can be approximated by ellipses. Note that the carbon transition model in Figure 2.3b shows a correlation between P_1 and P_2 with short and long axes of the same length order. The relationship between P_2 and P_3 in Figure 2.3c is closer to the linear case. Parameter P_4 in Figure 2.3d can vary widely with little effect on the prediction by the model, and there is a large difference in the length of the short and long axes of the contour. The ratio of the axes of ellipsoids approximating such contours can reach values of the order of 1000 in many systems biology models [Gutenkunst et al., 2007b]. Parameter sensitivity analyses on many metabolic and signalling networks taken from the BioModels database [Le Novere et al., 2006] showed that all these models have sloppy sensitivity spectra and that many parameters show a high degree of uncertainty [Gutenkunst et al., 2007b]. The dynamic behaviours of all analysed models depend on only a few stiff parameter combinations. Sloppy sensitivity spectra appear to be a universal property of systems biology models.

A simple example for a sloppy parameter combination can be given as follows. Consider a chemical side reaction of a linear metabolic pathway that follows the law of mass action with rate constants k_+ and k_- for forward and backward reactions, respectively. If the reaction is very fast relative to the rest of the biochemical network it is part of, it can be characterised solely by its equilibrium constant $K = k_+/k_-$ [Brown et al., 2004]. Both rate constants can be changed freely by the same factor without affecting K. In the model, the parameters k_+ and k_- represent removable degrees of freedom and they can be consolidated to one single parameter capturing the essential behaviour of the reaction.

The sloppy and stiff directions in parameter space are determined by calculating the eigenvectors of the Hessian matrix of the sum-of-squares function in the sloppy modelling approach [Brown et al., 2004, Gutenkunst et al., 2007b]. A similar approach to unveil more or less hidden correlations between parameters is principal component analysis (PCA) on a matrix, which describes parameter influence on model behaviour. Both methods identify linear combinations of parameter sensitivities that account for the major variation in the simulation outcome of a model. Those combinations that represent independent directions in parameter space can provide a basis for model reduction (e.g. [Gokulakrishnan et al., 2006]) in which negligible parameters are removed from the system in order to reduce complexity. A specific example of eigenvector decomposition of a model and what it reveals about the components of the model and their relations is given below (see section 2.8).

For a model of nerve growth factor signalling, [Brown et al., 2004] showed that individually determined experimental parameter measurements, e.g. obtained on isolated molecules, must be extremely precise in order to yield reasonable predictions of integrated network behaviour. They show that model parameters fitted collectively to time-course data on signal protein phosphorylation perform much better for model predictions of responses to new experimental interventions. Some parameter combinations can be determined accurately in mechanistic models of biological processes, while other combinations are ill determined but also do not matter so much for further predictions.

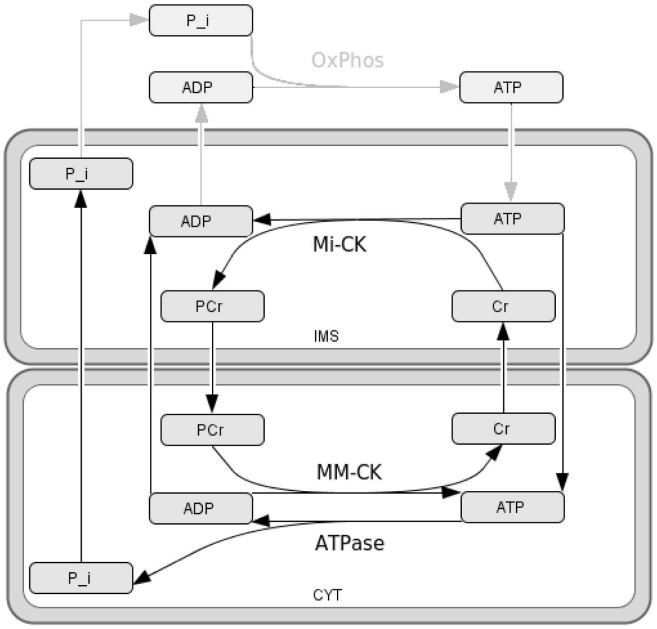

Figure 2.4: Scheme of the transport of high-energy phosphate groups from the mitochondrion into the cytosol in the muscle cell. The figure was generated using the software package CELLDESIGNER *[Funahashi et al., 2003].*

To make useful predictions, we can do a set of simulations with parameter combinations from the part of parameter space where it is plausible that parameter combinations are located. The part of parameter space considered should be compatible with the already available experimental data. This is termed an ensemble approach. It turns out that the ensemble prediction of the response in a new experiment can still be constrained to a useful small range.

2.8 Module for high-energy phosphate group transfer in heart muscle

In the following, we will illustrate the sloppy modelling ensemble approach by the analysis of a small model, published previously [van Beek, 2007]. This model describes the transport and buffering of high-energy phosphate groups in heart muscle and is used to analyse the dynamic adaptation of oxidative phosphorylation in the muscle cells to changing workloads. The model is publicly available in the BioModels database [Kongas and van Beek, 2007] and in the CellML model repository. A 'skeleton model' approach was chosen, meaning that only key elements that regulate the adaptation of adenosine triphosphate (ATP) synthesis in the mitochondrion to increased cytosolic

ATP hydrolysis in the muscle cell are incorporated (Figure 2.4).

After ATP has been synthesised in the mitochondrial matrix and exported to the mitochondrial intermembrane space (IMS), its high-energy phosphate group can be transported further into the cytosol (CYT) in two different ways, either by direct diffusion of ATP itself through the mitochondrial outer membrane (MOM) or by the so-called 'phosphocreatine (PCr) shuttle', which operates as follows. In the IMS, one high-energy phosphate group of ATP is transferred to creatine (Cr) to form PCr, which then diffuses through the membrane into the cytosol. In the cytosol, the phosphate group of PCr is then transferred to adenosine diphosphate (ADP) to form ATP. The reactions are catalysed by two isoforms of creatine kinase (CK), one mitochondrial form (Mi-CK) and one cytosolic form (MM-CK), respectively. The input of the model is ATP hydrolysis, given by a forcing function that represents the experimentally measured changing muscle workload. Inorganic phosphate (P_i) together with ADP is then transferred to the mitochondrial matrix to serve as the substrate for ATP synthesis. The model consists of 10 ODEs, one for each of the five metabolites in the cytosol and for the same metabolites in the IMS, and has 24 free parameters for enzyme kinetics and mitochondrial membrane permeability (see [van Beek, 2007] for details).

Metabolite and enzyme levels, as well as kinetic constants, had been experimentally determined and were collected from the scientific literature. For the permeability of the MOM to ATP and ADP, values from different experiments seemed to vary over a large range: its reported values differ from 0.16 to 85 $\mu M * s^{-1}$ [van Beek, 2007]. This may be due to damage to the interface of the mitochondria incurred during isolation from the cell or during removal of the outer cell membrane by chemical treatment. Simulation with the model showed that the dynamics of adaptation of oxygen consumption in the mitochondria depends strongly on MOM permeability. ADP permeability was therefore optimised to measurements with oxygen electrodes in the venous outflow from the heart as a whole corrected for the transport time in the coronary vessels, showing that the oxygen consumption in the mitochondria, which is directly linked to oxidative phosphorylation (OxPhos), increases with a generalised time constant of 3.7 s, in response to a step increase in electrically paced heart rate. With ADP permeability known, the behaviour of the system could be predicted. For instance, it was predicted that an increased expression of the Mi-CK leads to faster dynamic adaptation of oxidative phosphorylation, while an increased expression of MM-CK leads to slower dynamic adaptation of oxidative phosphorylation. Furthermore, it was predicted that the CK reaction leads to buffering of the oscillation of ADP in the cytosol and of oxidative phosphorylation in the mitochondria. However, very surprisingly, CK activity also leads to lower levels of inorganic phosphate, a breakdown product of ATP, which may inhibit contractility. The sloppy and stiff directions in parameter space (corresponding to the long and short axes of the ellipsoids; Figure 2.3) for this model were determined by calculating the eigenvectors of the Hessian matrix of the sum of squares. In the four eigenvectors with the largest eigenvalues, the coordinates of the parameters of several molecular processes

show up strongly. These are, respectively, oxidative phosphorylation, MM-CK, Mi-CK and the permeability to ADP of the MOM. An increased capacity for oxidative phosphorylation gives similar effects to increased affinity of oxidative phosphorylation to ADP and P_i in the eigenvectors with the three largest eigenvalues. A group of parameters for MM-CK and independently a group of parameters for Mi-CK are also found in the eigenvectors. The parameters of the equations of a single molecular process are usually coupled in a group in the eigenvectors. However, the eigenvectors tend to show combinations of molecular processes that are not necessarily directly connected in the scheme of the model (Figure 2.3). In this case, the eigenvectors therefore do not seem to provide a clearcut basis to divide the system into modules, although parameters in the nonlinear enzyme kinetic equations for a particular component are often found together. PCA of the parameter sensitivities of the model, an analysis done previously on a chemical reactor system ([Gokulakrishnan et al., 2006]), shows similar results. Interestingly, simulations with the system revealed that the PCr shuttle contributes only one-third to the total high-energy phosphate group transfer from the mitochondria to the sites of ATP hydrolysis to energise muscle contraction. This seemed to contradict the 'PCr shuttle hypothesis', which states that almost all high-energy phosphates in muscle cells are transferred as PCr [Bessman and Geiger, 1981]. However, our model prediction was based on a set of reference parameters on enzyme kinetics, and our statement was not supported by an extensive analysis of the experimental error levels that lead to imprecision in the parameters. Our aim in section 2.9 is therefore to investigate how uncertainty in the parameter values affects the predictions by the model.

2.9 Ensemble predictions of high-energy phosphate group transfer

The predictions derived from simulations of the high-energy phosphate transfer model give valuable insights into the functioning of energy metabolism in muscle cells. Nevertheless, the model relies strongly on the accuracy and precision of experimentally measured kinetic parameters for the components of the system. The parameter estimates are subject to experimental error and may differ between *in vitro* and *in vivo* conditions. In the following, we describe an ensemble approach to answer the question whether it is reasonable to state that the PCr shuttle accounts for only a small part of the total high-energy phosphate group transfer, given the limited precision with which model parameters are known. We chose the following ensemble method to evaluate the uncertainty of our prediction. We generated parameter sets for the ensemble simulation by drawing values from a Gaussian distribution with mean equal to the parameter value from the literature [van Beek, 2007] and a relative standard deviation of 0.1. In this way, we investigated what an error of 10 per cent for each measured parameter means for the predicted fraction of high-energy phosphate groups carried by PCr. The membrane permeability for ADP and ATP, assumed

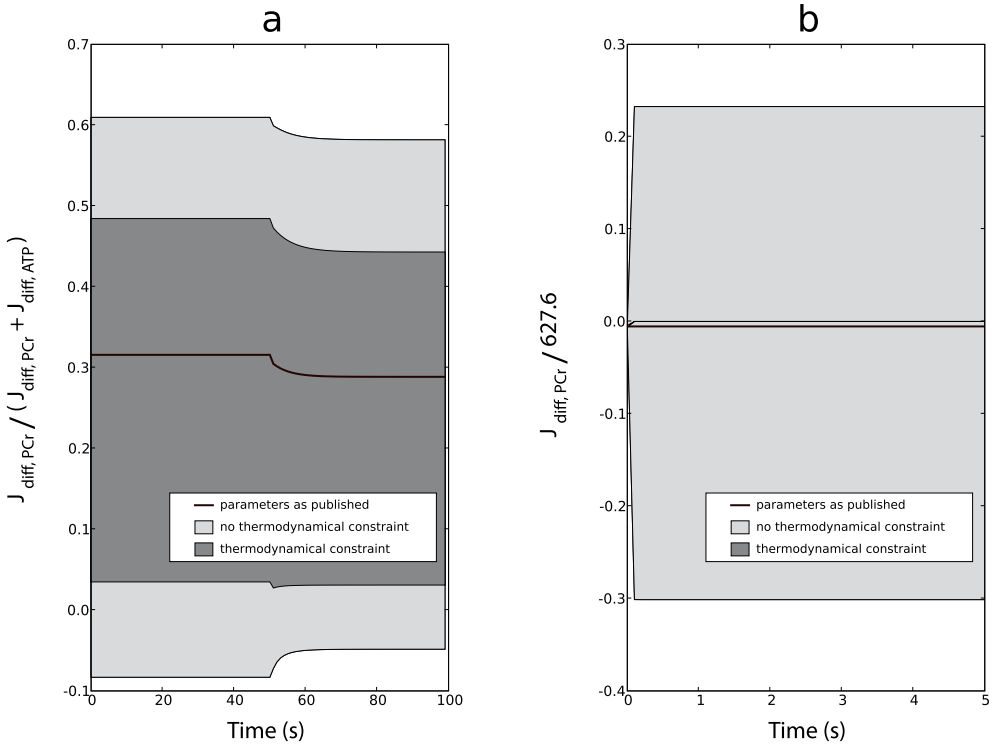

Figure 2.5: Fractions of high-energy phosphate group transfer carried by PCr. The light and dark grey shaded regions show the 95% central interval from an ensemble of size 1000 without and with thermodynamic constraint on system parameters, respectively. The fraction calculated from the reference parameter set [van Beek, 2007] is plotted in black. (a) Ensemble simulations over 100 *s with a step in ATP hydrolysis rate from* 486 *to 628* $\mu M * s^{-1}$ *at* 50 *s. (b) Ensemble simulations over 5âĂŁLs with ATP synthesis and hydrolysis rate of 0* $\mu M * s^{-1}$*. In this plot, the PCr flux values are divided by the maximum ATP hydrolysis rate in (a), 628* $\mu M * s^{-1}$*, to make the fractions comparable between (a) and (b). Please note the differences in scale between (a) and (b). Simulations were performed using the software package* SloppyCell*[Gutenkunst et al., 2007a].*

equal, is not randomly drawn directly, but optimised to exactly reproduce the randomly drawn generalised time constant of the response to an increase in the cytosolic ATP hydrolysis (3.7 s, s.e.=0.3 s), reflecting the variation found in the whole heart experiment. For each parameter set of the ensemble, the system is simulated over 100 s, with an increase in the ATP turnover rate by increasing the paced heart rate from 135 to 220 b.p.m. at 50 s. The fraction of high-energy phosphate being transferred by PCr is then calculated for each second of the resulting time series according to $J_{diff,PCr}/(J_{diff,ATP} + J_{diff,PCr})$, with J_{diff} giving the net diffusion flux of PCr and ATP from mitochondria to cytosol. The central 95 per cent upper and lower bound of the resulting trajectories from the ensemble calculation for the fraction of PCr phosphate transfer is plotted in Figure 2.5.

Strikingly, values for the calculated fraction drop below zero due to a net flux of PCr from the cytosol into the mitochondria, meaning that high-energy phosphate is transferred back into the IMS, which seems thermodynamically infeasible. This is caused by our scheme to generate random parameters, where the kinetic parameters of both CK enzymes were drawn independently. Six kinetic parameters for each enzyme determine the reaction's equilibrium constant, K_{CK}, via the Haldane relation. Drawing kinetic parameters independently therefore causes differences in K_{CK} in IMS and cytosol. This may cause a reversal of $J_{diff,PCr}$ (Figure 2.5a). Perpetual diffusion fluxes are calculated even for cases where there is no ATP hydrolysis and with oxidative phosphorylation being fully inhibited in the model (Figure 2.5b). Drawing kinetic parameters independently therefore causes a situation that is thermodynamically impossible. When there is no ATP hydrolysis and synthesis, all diffusion fluxes should become zero in the steady state. To prevent diffusion fluxes in the absence of any energy input, we add the thermodynamic constraint to the model that the K_{CK} is equal in IMS and cytosol. This is implemented by drawing one K_{CK} (with 10% random error added) and setting it equal in IMS and cytosol for both CK isoforms. For MM-CK and Mi-CK, five of the kinetic constants are drawn and the sixth constant is calculated to yield the correct K_{CK}. This is balanced such that all kinetic parameters are calculated in this way an equal number of times.

The resulting uncertainty range for the prediction of the contribution of PCr to the overall high-energy phosphate group transfer is shown as the dark grey region in Figure 2.5. Even though the ensemble prediction covers a broad range, Figure 2.5a suggests that PCr does not carry much more than half of the high-energy phosphate flux from mitochondria to contractile elements, which contradicts the PCr shuttle hypothesis posed more than 25 years ago in *Science* [Bessman and Geiger, 1981]. The prediction range is narrower if thermodynamically infeasible parameter combinations are annihilated. This example demonstrates that, for models relying on experimental parameter measurements, each prediction should be carefully investigated with respect to possible measurement errors.

2.10 Intuitive modelling for biologists: models of metabolism with Petri nets

Petri nets were developed in 1962 by Carl Adam Petri. The Petri net formalism describes a mathematical structure to construct directed bipartite graphs for modelling pairwise connections: a collection of two different node types is pairwise connected by edges. The nodes are called 'places', representing parts of a system, and 'transitions', representing interactions between the parts. In the common type of Petri nets, places contain a variable integer number of tokens that reflect the state of the place. The dynamic behaviour of the network is then described by the consumption of tokens from one place and the generation of tokens in the connected place when a transition fires during the execution of a Petri net (see [Murata, 1989] for a detailed

review). Identifying the places with metabolites and the transitions with biochemical reactions is natural and is probably more intuitive to non-mathematicians than ODE-based models. Hence, besides modelling the LIPSSS experiments described above by systems of ODEs, we modelled such metabolic networks also with Petri nets.

Petri nets can specify, verify and simulate a large amount of different systems (concurrent, asynchronous, distributed, parallel, deterministic and stochastic). They provide an intuitive way for users to deal with complex models and dynamic system behaviour due to their graphical representation. The user can easily switch between two points of view: (i) the transitions and the order in which they are executed and (ii) the states of the system that are reached after a given sequence of actions. Petri nets have been investigated for many years, leading to different Petri net types specialised for different applications.

Here, we present a new approach to simulate tracer and LIPSSS experiments using coloured Petri nets [Jensen, 1997]. Similar to Petri nets, metabolic networks are inherently bipartite due to the metabolic fluxes (transitions) that connect interacting metabolites (places) of the network. In addition, biological reactions take place independently or in parallel, which makes them easily representable by a concurrent system such as Petri nets. Furthermore, Petri nets provide random firing of transitions to simulate natural processes that are governed by stochastic laws. Their non-deterministic properties make them appropriate to model situations where numbers of molecules are low, because the randomness is expected to lead to small random deviations if many molecules react independently.

Coloured Petri nets contain distinct groups of tokens, each represented by a different colour. By assigning the colours to isotopomers, they are perfectly able to deal with the different labelling states of metabolites, i.e. isotopomers, occurring in tracer experiments. We developed a software package that enables one to easily specify metabolic networks at the level of carbon atoms and their connections via metabolic reactions (CTN). The user draws molecules and transitions in a graphical user interface. The CTN is then converted into a data structure representing the coloured Petri net.

During the conversion from CTN to Petri net representation, for each metabolic pool of the CTN, a place in the Petri net is created and the coloured tokens representing the isotopomers of the pool are determined. The user has the choice to simulate the Petri net's dynamic behaviour in either a continuous or a discrete (Figure 2.6a,b, respectively) manner [Chouikha and Schnieder, 1998].

However, the constitution and the number of tokens depend on the choice of Petri net, i.e. continuous or discrete. The continuous approach represents each isotopomer by a single coloured token with the fraction of the particular isotopomer attached as an attribute. In the discrete approach, however, the number of tokens for each isotopomer is proportional to the concentration of its metabolic pool. The number of molecules chosen to be represented by each token is the proportionality constant. Each token represents a fixed number of molecules of the metabolite, one or more, while

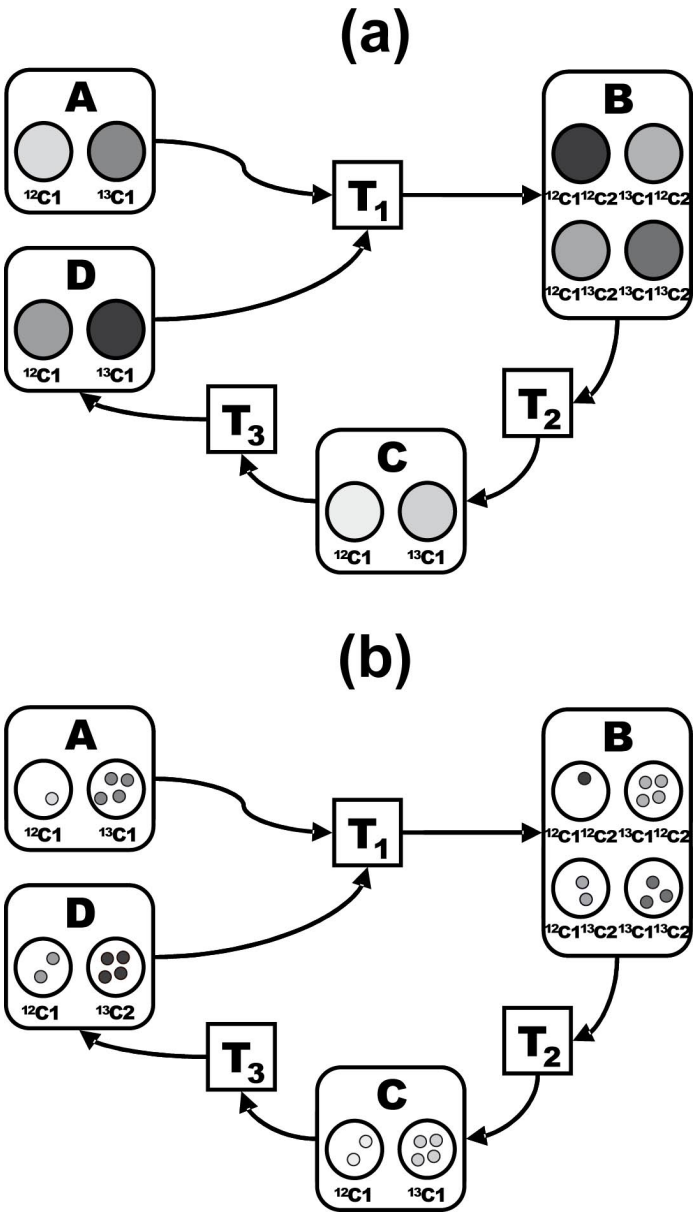

Figure 2.6: (a) Continuous and (b) discrete coloured Petri net representation of the CTN given in 2.1. The rectangles indicate 'places' that represent metabolites. The places are marked with tokens of different 'colours', each colour representing a different isotopomer. Please note that in the continuous case (a), only a single token for each isotopomer fraction is present. The token has a value attached to it to indicate the fraction. In the discrete case (b), the presence of an isotopomer is represented by a number of equally coloured tokens. The sum of tokens of an isotopomer in relation to all tokens present in the particular metabolite refers to the fraction of the isotopomer. The Petri net can be simulated either continuously or discretely (see text). The transitions T1-T3 in the Petri net represent a bundle of transitions between different isotopomers in the CTN.

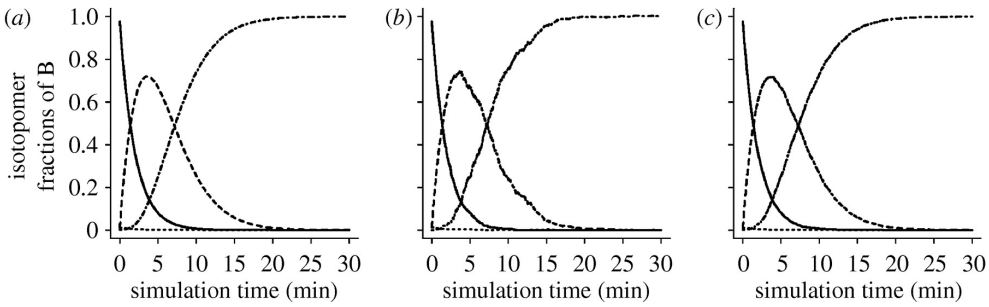

Figure 2.7: Simulation results of the metabolic network given in Figure 2.1 via Petri nets. The diagrams show the results of a simulation using (a) a continuous Petri net, (b) a discrete Petri net with a low number of molecules per metabolic pool (N=600 per pool) and (c) a discrete Petri net with a higher number of molecules per metabolic pool (N=6000 per pool). In the discrete cases, each molecule of a pool is represented by one token. Note that the simulation in (a) completely corresponds with the integration of the ODEs (Figure 2.1), while a simulation using discrete coloured Petri nets approaches such a continuous simulation better at a higher number of molecules. Isotopomers: $^{12}C1^{12}C2$*, solid line;* $^{13}C1^{12}C2$*, long-dashed line;* $^{12}C1^{13}C2$*, short dashed line;* $^{13}C1^{13}C2$*, dot-dashed line.*

the colour of the token still refers to its isotope composition. Finally, all individual carbon transitions present in the CTN are combined within the Petri net transitions. Therefore, for each reaction present in the metabolic network, it is determined which isotopomers in precursor pools give rise to each isotopomer of a particular metabolic pool and lumped within a Petri net transition. This information is subsequently used to perform the simulations. Both approaches offer the possibility to simulate the system step by step, giving the user the possibility to explore system behaviour interactively.

In the continuous case, a simulation of the Petri net closely resembles a simple Euler integration of the corresponding ODE representation of the model. At each time step, all Petri net transitions are fired simultaneously and the changes in the isotopomer fractions are updated in the numeric attribute attached to the tokens (one token per isotopomer). By contrast, during the discrete simulation, all Petri net transitions are fired in a random order and tokens representing 'real' molecules are transferred between the places. The results of a simulation of both the continuous and the discrete Petri net in Figure 2.6 are shown in Figure 2.7.

The results show that, although the discrete Petri net simulation is a completely different representation of the system, the overall behaviour of the system stays the same. The knowledge gained in this way for the experimentalist is very useful to develop metabolic models with desired properties and offers an entry point for a more detailed analysis using model representations via ODEs.

2.11 Future Directions

The human or animal metabolic network is extensive. Because it lies at the heart of physiology, it would be extremely useful to have a comprehensive understanding captured in computational models of metabolic functioning. Ideally, one would like to have access to an *in silico* model of metabolism, which covers the entirety of human metabolism without more or less arbitrary subdivisions and modularisations. However, owing to the complexity of metabolism, such subdivisions are necessary to study metabolism. Nevertheless, subdivisions can be quite arbitrary: a particular modularisation may be useful for one particular task and set of experiments, but insufficient for others. Metabolic pathways that interact little under one condition may become strongly linked under other conditions. The alternative is not very attractive: a large model constructed in one step is very difficult to test comprehensively. Partial reactions in the system may be in error. Modelling in one large step makes it hard to debug the system, while dividing the system into modules that are each tested and debugged is desirable.

Together, the approaches in section 2.4 may be helpful to tackle the modelling and analysis of the metabolic system, but the question is whether they are sufficient to allow us to model either the essence or the entirety of the complete metabolic system, despite the challenges posed by its large size and complex structure. The approaches of multiscale, modular, ensemble modelling, etc. are closely interwoven. Must we start by designing a grand scheme for modelling the system, providing a top-down very structured approach for modelling human metabolism? Even a grand design will work best if we proceed in relatively small steps. The difficulty for these steps will be that there are always boundaries to the modules into which we divided the system that may be hard to control.

A crucial aspect is organisation of the modelling process in such a way that, if inevitable mistakes are made in early model versions, the interplay between computer simulations and experimental tests results in a gradual improvement of the model. We must aim to make the modelling process the driving force behind metabolic experimentation and data collection, such that it becomes the vehicle for integration of knowledge and understanding of the complete metabolic system.

2.12 Acknowledgements

This work is part of the BioRange programme (project no. SP 2.2.1) of The Netherlands Bioinformatics Centre, which is supported by a BSIK grant through The Netherlands Genomics Initiative (NGI). The work is also part of the Centre for Medical Systems Biology, which is a Genomics Centre of Excellence funded by the Dutch Government via the NGI.

2.13 Author Contributions

Measuring metabolic fluxes *in vivo* with stable isotopes (section 2.5): TWB, JHGMvB. Sloppy parameter behaviour (section 2.7): HH, TWB, JHGMvB. Module and ensemble predictions for high-energy phosphate group transfer in heart muscle (sections 2.8 & 2.9): HH, JHGMvB. Models of metabolism with Petri Nets (section 2.10): AH, TWB, JHGMvB.

Chapter 3

Analysing the functional properties of the creatine kinase system with multiscale 'sloppy' modelling

Hettling H, van Beek JHGM(2011) Analysing the Functional Properties of the Creatine Kinase System with Multiscale 'Sloppy' Modelling. *PLoS Computational Biology* 7(8): e1002130.

3.1 Abstract

In this study the function of the two isoforms of creatine kinase (CK; EC 2.7.3.2) in myocardium is investigated. The 'phosphocreatine shuttle' hypothesis states that mitochondrial and cytosolic CK plays a pivotal role in transport of high-energy phosphate (HEP) groups from mitochondria to myofibrils in contracting muscle. Temporal buffering of changes in ATP and ADP is another potential role of CK. With a mathematical model, we analysed energy transport and damping of high peaks of ATP hydrolysis during the cardiac cycle. The analysis was based on multiscale data measured at the level of isolated enzymes, isolated mitochondria and on dynamic response times of oxidative phosphorylation measured at the whole heart level. Using 'sloppy modelling' ensemble simulations, we derived confidence intervals for predictions of the contributions by phosphocreatine (PCr) and ATP to the transfer of HEP from mitochondria to sites of ATP hydrolysis. Our calculations indicate that only 15 ± 8 (mean $\pm$ SD) of transcytosolic energy transport is carried by PCr, contradicting the PCr shuttle hypothesis. We also predicted temporal buffering capabilities of the CK isoforms protecting against high peaks of ATP hydrolysis (3750 $\mu M * s^{-1}$) in myofibrils. CK inhibition by 98% in silico leads to an increase in amplitude of mitochondrial ATP synthesis pulsation from 215 ± 23 to 566 ± 31 $\mu M * s^{-1}$, while amplitudes of oscillations in cytosolic ADP concentration double from 77 ± 11 to 146 ± 1. Our findings indicate that CK acts as a large bandwidth high-capacity temporal energy buffer maintaining cellular ATP homeostasis and reducing oscillations in mitochondrial metabolism. However, the contribution of CK to the transport of high-energy phosphate groups appears limited. Mitochondrial CK activity lowers cytosolic inorganic phosphate levels while cytosolic CK has the opposite effect.

3.2 Author Summary

Creatine kinase (CK) has several functions in cellular energy metabolism. It catalyses the reversible transfer of high-energy phosphate from ATP to creatine, facilitating storage of energy in the form of phosphocreatine. In muscle cells, this extra energy buffer plays a pivotal role in maintaining ATP homeostasis. Another proposed function of CK is the transport of energy from ATP producing to ATP consuming sites via a shuttle mechanism involving a mitochondrial and a myofibrillar isoform of CK. The extent to which this phosphocreatine shuttle mechanism is used in muscle and other tissues is hotly debated. We use a computational model of the CK system which can predict energy transport and buffering of high demand peaks to estimate the relative importance of both roles in heart muscle. We validate the model with multiscale data on the level of enzyme kinetic constants and with dynamic oxygen consumption measurements in rabbit hearts. Since model predictions can be strongly affected by changes in parameter values, we employ 'sloppy' ensemble modelling which allows to set confidence regions for predictions. Our results indicate that the main function

of CK in heart muscle lies more in temporal energy buffering of high peaks in ATP consumption during cardiac contraction than in energy transportation.

3.3 Introduction

It is well established that creatine kinase (CK) catalyses the reversible transfer of phosphate from ATP to creatine (Cr):

$$ATP + Cr \overset{CK}{\longleftrightarrow} ADP + PCr \tag{3.1}$$

However, how this biochemical function plays a role in cell functioning has been the subject of intense controversy [Beard and Kushmerick, 2009]. It is remarkable that two distinct isoforms of CK are expressed in muscle cells, one in the mitochondrial inner membrane space (IMS) and one in the cytosol where the contractile elements are located. This led to the idea of the 'phosphocreatine shuttle', proposed by Bessman and Geiger [Bessman and Geiger, 1981]: PCr formation from adenine nucleotide and creatine in the IMS is catalysed by the mitochondrial isoform of CK, Mi-CK, located in the IMS. PCr may then proceed to the cytosol, which constitutes a mechanism of facilitated diffusion of high-energy phosphate (HEP) groups. Re-transfer of HEP to adenine nucleotide to energise myofibrillar contraction is done by the muscular isoform of CK, MM-CK, located in the cytosol (see Figure 3.1).

Transfer of HEP was argued to be accomplished either by direct diffusion of ATP through the mitochondrial outer membrane (MOM) and cytosol or indirectly via the 'phosphocreatine shuttle'. The phosphocreatine shuttle hypothesis has led to extensive scientific debates on the role of CK, e.g. [Beard and Kushmerick, 2009, Meyer et al., 1984, Greenhaff, 2001].

Besides the energy transfer function, the creatine kinase system was thought to be responsible for (i) temporal energy buffering by maintaining an adequate ATP/ADP ratio during interruption of energy supply [Beard, 2006] or during changing energy demand [Meyer et al., 1984, Vendelin et al., 2000] and (ii) for regulation of oxidative phosphorylation [Saks et al., 1996]. The CK system, transporting creatine instead of ADP from the cytosol to the mitochondria, is a potential key regulator of oxidative phosphorylation. CK inhibition experiments on rabbit hearts [Harrison et al., 1999, 2003] and CK knockout experiments in mice [Gustafson and Van Beek, 2002] suggest that the creatine kinase system affects the dynamic adaptation of oxidative phosphorylation to energy demand.

Mathematical modelling has proven helpful to understand the CK system: several existing models account for a compartmentalised energy metabolism system in myocytes under various conditions [Vendelin et al., 2000, Aliev and Saks, 1997, Saks et al., 2000, Vendelin et al., 2004, Beard, 2005, Wu et al., 2008, Wu and Beard, 2009]. The main differences between the model analysed here and other models described in the

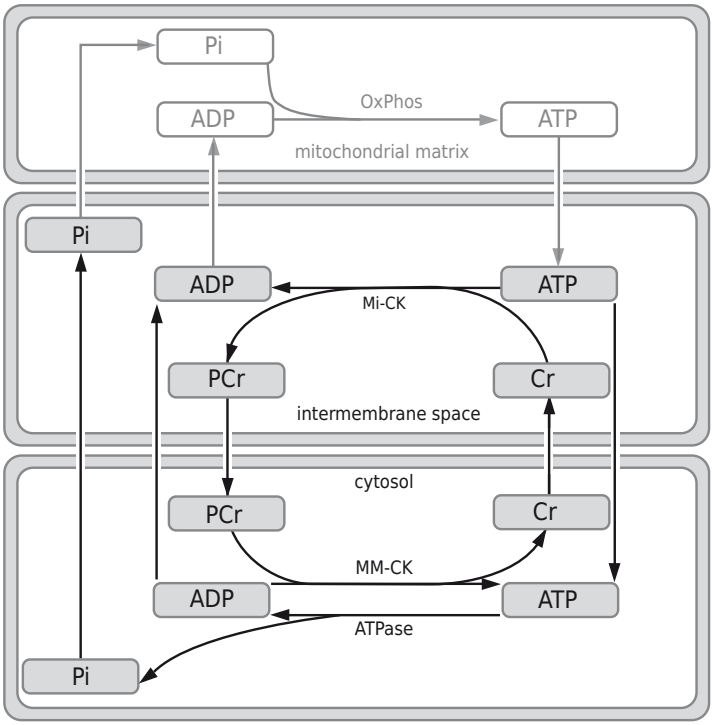

Figure 3.1: Scheme of model of the compartmentalised creatine kinase system. Main elements are ATP hydrolysis by ATPase, ATP synthesis by mitochondria and creatine kinase (CK) isoforms in the mitochondrial intermembrane space (Mi-CK) and cytosol (MM-CK). Oxidative phosphorylation (OxPhos) takes place in the mitochondrial matrix and responds to ADP and inorganic phosphate (P_i) levels in the mitochondrial intermembrane space. The concentrations of phosphocreatine (PCr), creatine (Cr), ADP, ATP and P_i are dependent on the rates of the enzyme reactions and transport. The figure was generated with CellDesigner *[Funahashi et al., 2003].*

literature are addressed in section 3.5. We build on a previously published computational model for the dynamic adaptation of oxidative phosphorylation to changing workloads which captures the key elements responsible for buffering and transport of HEP between IMS and cytosol [Kongas and van Beek, 2007, van Beek, 2007]. The model incorporates synthesis of ATP from ADP by oxidative phosphorylation in the mitochondria and ATP consumption in the cytosol, the reversible transfer of phosphate groups from ATP to creatine by CK enzyme reactions and metabolite diffusion between IMS and cytosol through the MOM (see Figure 3.1). The model's dynamic behaviour is affected by 22 free parameters for enzyme kinetics and membrane permeability, which had been determined experimentally and were collected from the scientific literature.

In recent work we investigated the sensitivity of the predictions of this CK model

with respect to possible error in the parameters using a simplified ensemble approach and found that even a modest error on each model parameter results in a broad range of possible predictions [van Beek et al., 2009]. However, models containing many molecular kinetic parameters, all known with little accuracy, can yield useful predictions as long as the correlation of these inaccuracies is taken into account. Brown et al. showed, using a computational model of nerve growth factor signalling, that viable model predictions can be achieved in spite of a high degree of uncertainty in all kinetic parameters [Brown et al., 2004, Brown and Sethna, 2003]. The approach identifies so-called 'sloppy' combinations of parameters, which, when jointly altered, do not significantly change the outcome of a model simulation, meaning that multiple combinations of parameters describe experimental data equally well. Gutenkunst et al. investigated a variety of metabolic and signalling networks and found these spectra of correlated parameter sensitivities to be universal in Systems Biology models [Gutenkunst et al., 2007a]. To use the information from these hidden correlations between parameters, a Bayesian ensemble of distinct parameter sets which agree with experimental data, can be sampled with Markov- Chain Monte Carlo (MCMC) methods. The likelihood of a parameter combination being included in the ensemble is proportional to the parameter combination's likelihood to predict the experimental input data set. Starting point for the walk through parameter space is the parameter set obtained from a least-squares parameter fit to experimental data. The resulting ensemble of parameter sets, constrained by the experimental data, allows a quantification of uncertainty not only of parameter values, but also delineates the uncertainty of model predictions for new experimental interventions. Below we demonstrate that combining molecular kinetic data, organellar data and whole organ response data with a sloppy modelling approach is feasible and fruitful.

We assembled a set of prior knowledge data on kinetic parameters of the CK enzymes and made use of measurements on the oxidative capacity and kinetics of isolated mitochondria and on metabolite transport across membranes and cytosol. These data at the molecular and organellar level were combined with experimental data on the response of the whole heart: for jumps to multiple heart rate levels the response time of the increase in oxygen uptake in the heart was measured. Based on model analysis of the oxygen transport system, the response time of oxygen uptake at the level of the mitochondria could be calculated from the whole heart level uptake [Harrison et al., 2003]. These response times for wild type CK levels and during CK inhibition played an important role as input data for the MCMC analysis. Based on these data from multiple levels in the system, we predict the contribution of PCr to HEP transport and the buffering capacity of the system toward the high-frequency high amplitude pulsations of ATP hydrolysis during the cardiac cycle. As a consequence, we determined that the functional role of the CK system in energy transport is limited and that high pulses in ATP hydrolysis are buffered by CK at order 100 millisecond time scales; both functions are presently not directly accessible to experimental measurement. Surprisingly, we also find that the mitochondrial CK isoform plays a role in regulating the cytosolic inorganic phosphate level.

3.4 Results

We employed experimental data from three scales: molecular kinetic parameters, organellar capacity parameters and whole organ dynamic response data. A priori experimental information about kinetic parameters was extracted from the literature (see Table 3.1). For nine of the 22 model parameters, standard measurement errors were reported. In order to constrain these parameters by their measurement errors, we added this molecular and organellar information as terms to a least-squares cost-function which also contained dynamic response times measured at the whole heart level (see section 3.6). In this way experimental data from the molecular, organellar and whole system level are treated in a unified way. For the MOM permeability to adenine nucleotides ($PS_{mom,AdN}$), a key parameter affecting the system's energy transport and buffering behaviour, values in literature were contradictory [van Beek, 2007]. The parameter $PS_{mom,AdN}$ was therefore not constrained. The cost function determines the probability that a parameter set is compatible with the observed data (see section 3.6). Using Markov Chain Monte Carlo, a distribution of parameter sets with high probability of agreement with the data is sampled. The resulting ensemble of parameter sets is therefore a multivariate posterior distribution, shaped by the cost function, which reflects the probability of individual parameter sets in a Bayesian sense [Brown and Sethna, 2003].

Data on the response times of the whole system level were taken from a study by Harrison et al., where electrically paced perfused rabbit hearts were exposed to a step increase in heart rate [Harrison et al., 2003]. After applying the challenge, the metabolic delay time t_{mito} was calculated from dynamic O_2 consumption measurements to estimate the generalised time constant of the ATP production time course. From a baseline level of 135 beats/min (bpm), heart rate was increased to 160, 190 and 220 bpm, respectively. Hearts were either exposed to iodoacetic acid (IAA) to block glycolysis or to iodacetamide (IA) to inhibit both glycolysis and CK activity, yielding in total 6 data points on the response time of oxidative phosphorylation, shown in Figure 3.2. Details on model, experimental data, cost function and the ensemble modelling approach can be found in section 3.6.

3.4.1 Parameter Estimation

Model parameters were estimated simultaneously to fit the t_{mito} values for all conditions using a least-squares optimisation procedure. Different optimisation algorithms (downhill simplex algorithm, Powell's method, Levenberg-Marquardt) gave similar quality of the fit. Initial and optimised parameter values can be found in Table 3.1. Figure 3.2 shows all t_{mito} values predicted by the model before and after parameter optimisation for all conditions. After fitting, the model correctly predicts a quicker energy supply-demand signalling when CK is inhibited by 98%, causing weaker ADP/ATP buffering by CK. In the optimisation procedure, the maximum velocities of the Mi-CK and the MM-CK enzyme were decreased by 12 and 36%, respectively,

Name	Description	Value	Unit	Reference	Optimised value	Prior $\sigma_{ln\theta}$	Ensemble mean ± SD	Posterior $\sigma_{ln\theta}$
$K_{eq,CK}$	Equilibrium constant for Mi-CK and MM-CK	152.0 ± 0.4		[Teague Jr. and Dobson, 1992]	151.95	0.026	152.23 ± 3.82	0.025
Parameters for the mitochondrial creatine kinase reaction								
$V_{max,Mi,f}$	Maximum velocity Mi-CK (PCr production)	882.0	$\mu M/s$	[van Beek, 2007]	775.05	0.336[a]	760.05 ± 264.39	0.333
$K_{ia,Mi}$	Binary dissociation constant ATP	750.0 ± 60.0	μM	[Jacobus et al., 1982]	751.32	0.081	754.79 ± 62.61	0.083
$K_{ib,Mi}$	Binary dissociation constant Cr	28800 ± 8450	μM	[Jacobus et al., 1982]	28733.44	0.336	29742 ± 10117	0.332
$K_{ic,Mi}$	Binary dissociation constant ADP	204.0	μM	[Aliev and Saks, 1997]	201.73	0.336[a]	221.03 ± 79.15	0.337
$K_{id,Mi}$	Binary dissociation constant PCr	1600.0 ± 200.0	μM	[Jacobus et al., 1982]	1597.69	0.128	1597.04 ± 190.54	0.118
$K_{b,Mi}$	Ternary dissociation constant Cr	5200.0 ± 300.0	μM	[Jacobus et al., 1982]	5209.08	0.058	5196.36 ± 302.73	0.058
$K_{d,Mi}$	Ternary dissociation constant PCr	500.0 ± 20.0	μM	[Aliev and Saks, 1997] [Saks et al., 1976]	499.51	0.040	502.19 ± 20.64	0.041
Parameters for the myofibrillar creatine kinase reaction								
$V_{max,MM,f}$	Maximum velocity MM-CK (ATP production)	11441.78	$\mu M/s$	[van Beek, 2007]	7373.07	0.336[a]	7769.77 ± 2591.30	0.308
$K_{ia,MM}$	Binary dissociation constant ATP	900.0	μM	[Aliev and Saks, 1997]	1026.24	0.336[a]	1033.59 ± 351.91	0.336
$K_{ib,MM}$	Binary dissociation constant Cr	34900	μM	[Jacobus et al., 1982]	34504.19	0.336[a]	36772 ± 12695	0.330
$K_{ic,MM}$	Binary dissociation constant ADP	222.4	μM	[Jacobus et al., 1982]	212.26	0.336[a]	225.49 ± 78.53	0.338
$K_{id,MM}$	Binary dissociation constant PCr	4730.0	μM	[Jacobus et al., 1982]	4516.55	0.336[a]	4955.05 ± 1692.93	0.329
$K_{b,MM}$	Ternary dissociation constant Cr	15500 ± 2500	μM	[Aliev and Saks, 1997] [Saks et al., 1976]	16744.44	0.167	16869 ± 2940	0.177
$K_{d,MM}$	Ternary dissociation constant PCr	1670 ± 40	μM	[Aliev and Saks, 1997] [Saks et al., 1976]	1669.76	0.024	1670.91 ± 38.38	0.023
Parameters for mitochondrial ATP production								
$V_{max,syn}$	Maximum ATP synthesis velocity	1503.74 ± 152.65	$\mu M/s$	[de Groot, 1999]	1332.64	0.103	1320.53 ± 113.50	0.085
K_{adp}	Apparent K_m mitochondria for ADP	25.0	μM	[van Beek, 2007] [Heineman and Balaban, 1990]	35.88	0.336[a]	34.61 ± 7.80	0.228
K_{pi}	Apparent K_m mitochondria for P_i	800.0	μM	[van Beek, 2007] [Heineman and Balaban, 1990]	346.57	0.336[a]	378.88 ± 118.91	0.296
Permeabilities of the mitochondrial outer membrane								
$PS_{mom,AdN}$	Membrane conductance ATP and ADP	13.3	s^{-1}	[van Beek, 2007]	23.64	None	31.74 ± 16.58	0.500
$PS_{mom,PCr}$	Membrane conductance PCr	162.5	s^{-1}	[van Beek, 2007]	155.49	0.336[a]	167.42 ± 57.39	0.334
$PS_{mom,Cr}$	Membrane conductance Cr	162.5	s^{-1}	[van Beek, 2007]	154.20	0.336[a]	163.06 ± 59.68	0.350
$PS_{mom,pi}$	Membrane conductance P_i	194.0	s^{-1}	[van Beek, 2007]	195.63	0.336[a]	199 ± 68.34	0.324

Table 3.1: *Shown are all model parameters describing the enzyme kinetics and transport across the mitochondrial outer membrane. The third column gives the parameter values obtained from the literature. If a standard measurement error could be obtained from literature, the value is given. We also give the parameter values after least-squares optimisation to experiment data. We finally give means and standard deviations of the parameter ensemble. Note that the model parameters for the maximum backwards velocity of both CK reactions, $V_{max,Mi,b}$ and $V_{max,MM,b}$ are not listed because their values are dynamically calculated from other parameter values via the Haldane relation: $V_{max,f}/V_{max,b} = (K_{ia} * K_b)/(K_{eq} * K_{ic} * K_d)$.*

[a] *$\sigma_{ln\theta}$ values which determine the spread of the prior distribution for parameters with standard errors not available from the experimental literature were set to 0.336, which is the maximum $\sigma_{ln\theta}$ value for parameters where the standard error is known from the literature.*

from their initial literature values. These literature enzyme activities for MM-CK and Mi-CK had been taken from the same experimental model, but without inhibition of glycolysis by IAA [Harrison et al., 1999].The experimental data used in the present analysis was measured in the presence of IAA which was found to impede CK activity by 20% [Harrison et al., 2003]. The drop in estimated CK activity is therefore plausible. Other parameters which are altered significantly by the optimisation are the apparent Michaelis constant for inorganic phosphate in the mitochondrion, K_{pi}, which drops from 800 to 347 μM, and the apparent K_M for ADP, K_{adp}, which increases from 25 to 36 μM. Both parameters occur in the model equation determining the rate of oxidative phosphorylation, which may explain the inverse variation. There exist *in vitro* measurements of K_{pi} that are lower than the initial value used in this analysis [van Beek, 2007]: Stoner & Sirak for instance measured K_{pi} to be 360 μM [Stoner and Sirak, 1979] which is close to our optimised value. Likewise, reported values for K_{adp} vary between 20 and 30 μM [Heineman and Balaban, 1990, Jacobus and Saks, 1982], corroborating the values obtained by the fit.

3.4.2 Monte Carlo sampling of parameter sets

Starting from the optimised parameter set (see Table 3.1), we sampled the parameter space to generate an ensemble of 658 independent parameter sets using the Metropolis-Hastings algorithm. The parameter set yielding the lowest cost in the complete ensemble was this optimised parameter set. The distributions of all parameters in the ensemble are shown in Figure 3.3.

The nine kinetic parameters which had known error values (see Table 3.1) show a mean value in the ensemble close to the measured value and a standard deviation close to their reported measurement error from the literature, which was to be expected given the prior information in the cost function. However, the parameters for which there was no standard error value available from the literature in general gave a standard deviation in the ensemble which was smaller than the default assigned large standard error (see Table 3.1). We tested the effect of different assumptions on the default prior standard deviations on posterior parameter distributions and ensemble predictions, reported in the supplemental material (section 3.7) which shows that the conclusions reported here are not changed by larger or smaller values on the default prior.

The mean value of $PS_{mom,AdN}$ in the ensemble is 31.7 s^{-1}, which is larger than the optimised value of 13.3 s^{-1} found previously [van Beek, 2007]. The distribution of $PS_{mom,AdN}$ shows substantial skewing with a minimum value of 7.4 s^{-1}, and a rather sharp exclusion of small values which give slow response times of the system. Based on experiments in isolated permeabilised cardiomyocytes, Sepp et al. estimated a value for MOM permeability to adenine nucleotides of 1833 nmol/min/mg protein per mM concentration difference [Sepp et al., 2010]. Converting this value expressed per mg tissue protein, assuming 150 mg protein per gram wet weight, this corresponds to $PS_{mom,AdN} = 7.45 \pm 1.89\ s^{-1}$. This is virtually the same as the minimum estimated

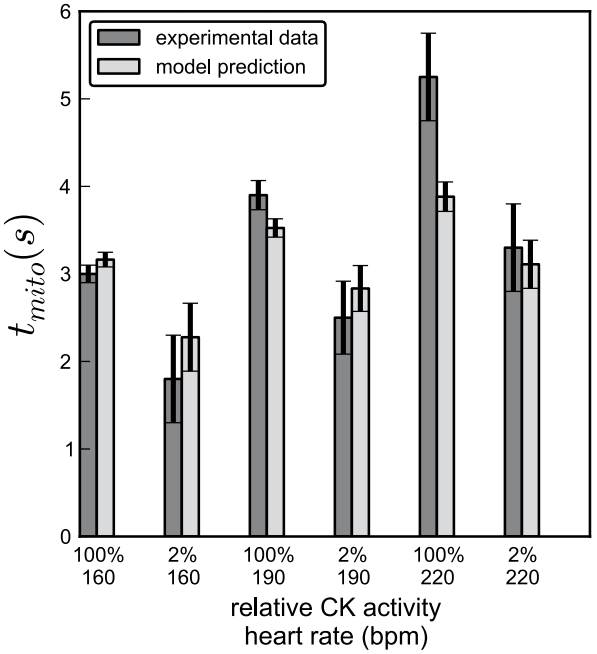

Figure 3.2: Fit by the model of measured response times to heart rate steps. The response times of oxidative phosphorylation (t_{mito}) were measured in isolated rabbit hearts [Harrison et al., 2003]. Model parameters were estimated using a modified Levenberg-Marquardt algorithm. Dark grey bars represent the t_{mito} values from the experiment, light grey bars represent the t_{mito} values predicted by the model after the fitting procedure. Data is available for six different conditions: three different amplitudes of heart rate jump (from 135 bpm to 160, 90 and 220 bpm heart rate), each one measured with full wild-type CK activity (100%) or with CK activity inhibited to 2% of wild-type value. The error bars reflect the standard error of the measurements and the standard deviation of the t_{mito} values in the ensemble, respectively.

in our ensemble analysis.

3.4.3 Predicting the contribution of PCr and ATP to energy transport

The contribution of PCr to intracellular HEP transfer, $R_{diff,PCr}$, is quantified by the ratio of PCr diffusion ($J_{diff,PCr}$) to the total phosphate group diffusion through the MOM:

$$R_{diff,PCr} = \frac{J_{diff,PCr}}{J_{diff,PCr} + J_{diff,ATP}} \tag{3.2}$$

An ensemble of simulations based on the parameter ensemble described above allows evaluation of the confidence region for the model prediction. In the ensemble,

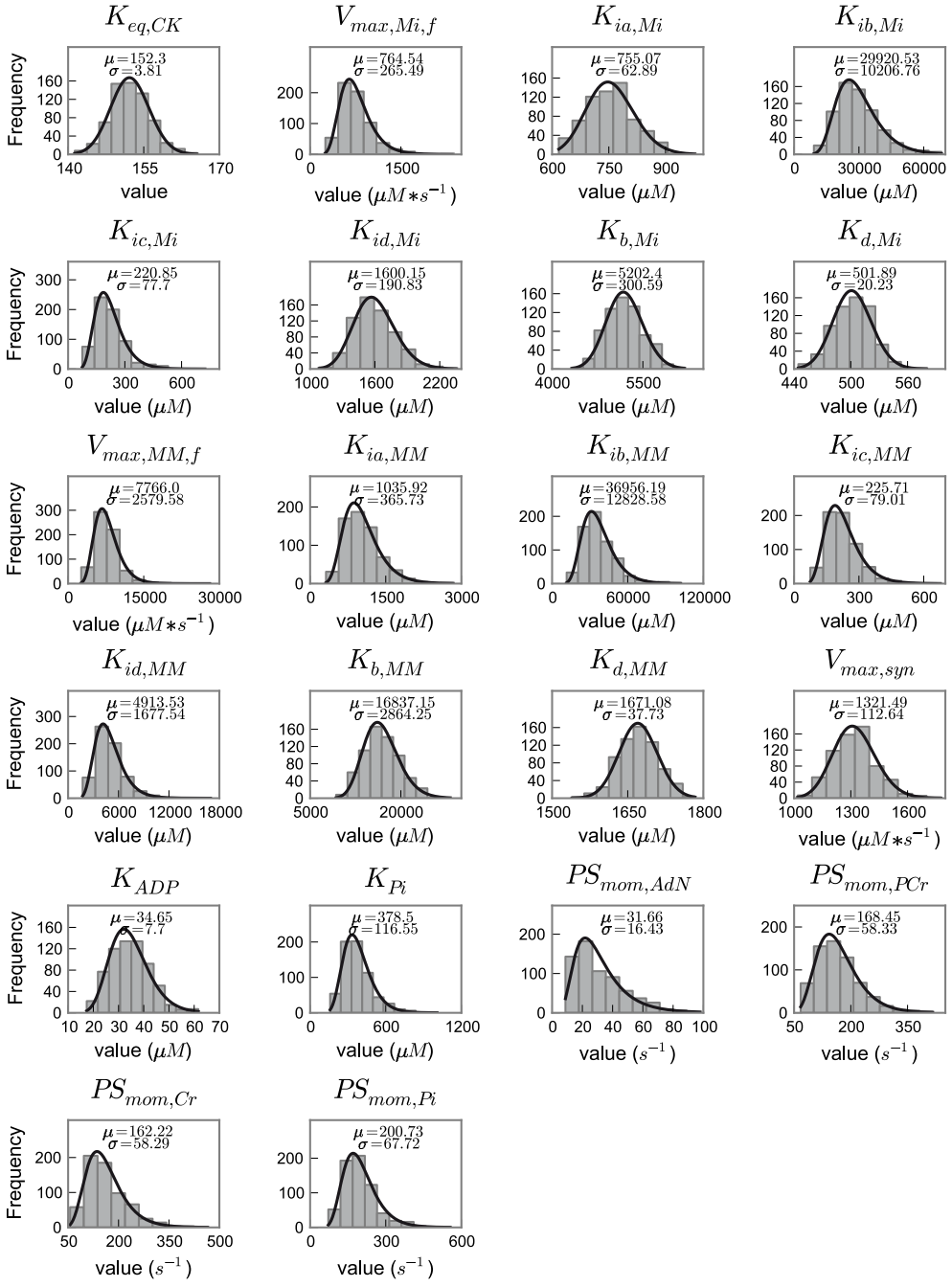

Figure 3.3: Distributions of individual parameters in the ensemble generated by the Metropolis-Hastings algorithm. Plots show histograms of all values in the ensemble for the given parameter. The ensemble consists of 658 parameter sets. Also plotted is the probability density function of the log-normal distribution with mean and standard deviation of each parameter scaled to the observed frequencies.

$R_{diff,PCr}$ is on average 0.17±0.09 (mean ± SD) at heart rate 160 bpm and 0.15±0.08 at 220 bpm in the case of active CK. Figure 4 shows the 95% confidence interval between upper and lower bound of the ensemble prediction for $R_{diff,PCr}$ for IAA and IA conditions in steady state at heart rate 220 bpm. The small oscillations during CK inhibition are due to the 2% residual activity of CK. The upper bound of the 95% confidence interval remains below 0.44 during the cardiac cycle for all simulated conditions.

$R_{diff,PCr}$ decreases during the peaks in ATP hydrolysis and even becomes negative for the lowest trajectories in the ensemble, which indicates that PCr diffuses back to the mitochondria at the end of systole (Figure 3.4). The simulations show for these cases that ADP diffuses into the IMS during the peaks of ATP hydrolysis, stimulating a reversal of the mitochondrial CK reaction to produce ATP from PCr, exactly as happens in the cytosol. For these lowest trajectories in the ensemble the CK activity per unit volume of the intermembrane space is higher than the CK activity per unit volume of the cytosol, causing the PCr to go down more steeply in the intermembrane space. This causes the cytosolic PCr concentration to exceed the PCr concentration in the IMS, and a negative gradient forces PCr to diffuse back into the IMS. However, when averaged over the cardiac cycle, $R_{diff,PCr}$ is always positive, indicating net flux of PCr from the mitochondria to the cytosol, and for the vast majority of the ensemble PCr diffusion flux never becomes negative during the entire cardiac cycle. Simulations suggested that the relative importance of the PCr shuttle becomes less with higher ATP hydrolysis at heart rates of 160, 190 and 220 bpm. We tested this hypothesis by predicting $R_{diff,PCr}$ for a range of heart rates from 60 to 300 bpm. The ensemble simulations reveal that $R_{diff,PCr}$ continuously drops for increasing heart rates for all sampled parameter combinations (see Figure 3.5). The predicted decline in $R_{diff,PCr}$ and increase in P_i concentration agrees with results of a recent study on perfused rat hearts [Vendelin et al., 2010]. Increased energy demand induces an increased ATP gradient between both compartments. At 160 bpm, the average difference between the ATP concentration in IMS and cytosol is 18.6 μM, at 220 bpm it becomes 22.3 μM for the optimal parameter set. The increased ATP gradient across the MOM induces direct ATP transport instead of facilitated transport via PCr.

In order to demonstrate the dependence of shuttle utilisation on the membrane conductance for adenine nucleotides, we predicted $R_{diff,PCr}$ as a function of $PS_{mom,AdN}$ for the ensemble. The predicted range shown in Figure 3.5B indicates that only for very small ATP permeability, PCr contribution becomes high. Even for the minimum value for $PS_{mom,AdN}$ in the ensemble (7.35 s^{-1}), the entire 95% confidence interval of $R_{diff,PCr}$ remains below 0.5. Low MOM permeability to adenine nucleotides causes high-energy phosphate group transport via PCr, and that $PS_{mom,AdN}$ is never lower than 7.35 s^{-1} therefore argues against a predominant phosphocreatine transport. Also when the value $PS_{mom,AdN} = 7.45\ s^{-1}$ estimated from Sepp et al. ([Sepp et al., 2010]), see above, is incorporated as prior knowledge, the analysis still yields similar predictions of $R_{diff,PCr}$, which stays with 95% confidence between 0.16 and 0.46 at heart rate 220 bpm.

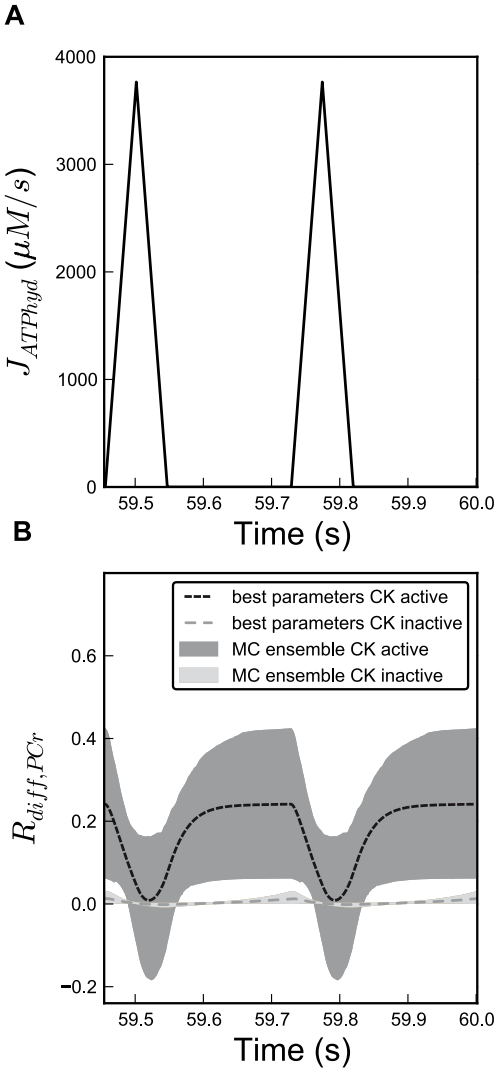

Figure 3.4: (A) Forcing function of pulsatile cytosolic ATP hydrolysis for the last two cardiac cycles of a simulation over 60s.(B) Prediction of the relative PCr contribution to high-energy phosphate flux across the mitochondrial outer membrane ($R_{diff,PCr}$) at heart rate 220 bpm. The shaded region gives the central 95% confidence interval of the $R_{diff,PCr}$ trajectories derived from ensemble simulations of 658 parameter sets. Solid lines depict a single simulation of the best scoring parameter set. Dark grey shading indicates the condition with CK active. Simulations with CK inhibited by 98% by IA are plotted in light grey. Note that two cardiac cycles are plotted after a steady state was reached.

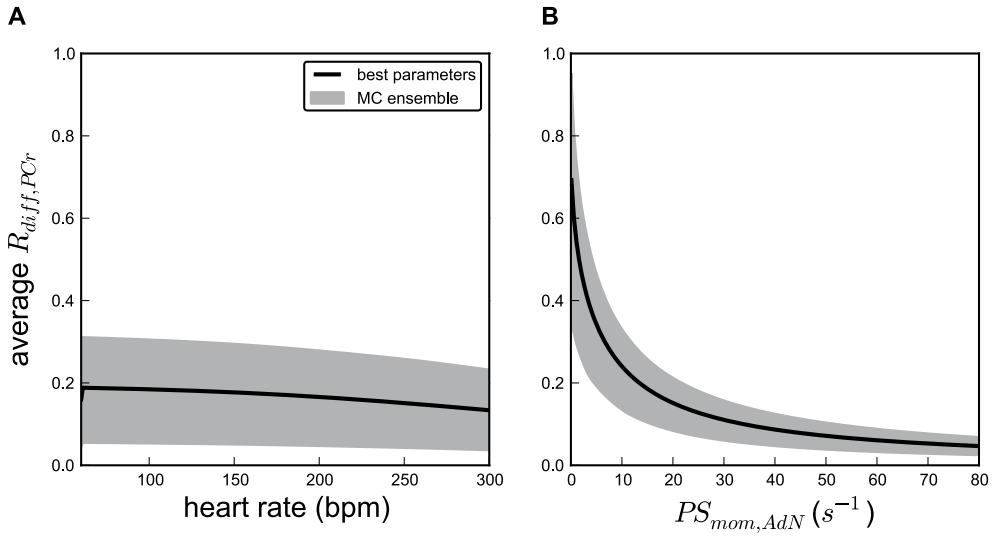

Figure 3.5: Dependence of PCr diffusion flux on heart rate and mitochondrial membrane permeability to adenine nucleotides. Prediction of the PCr contribution to high-energy phosphate flux across the mitochondrial outer membrane ($R_{diff,PCr}$), averaged over the cardiac cycle, as a function of (A) heart rate and (B) mitochondrial outer membrane permeability for adenine nucleotides ($PS_{mom,AdN}$), respectively. Values for (A) Steady state values for $R_{diff,PCr}$ as a function of heart rate (B) Steady state values for $R_{diff,PCr}$ as a function of $PS_{mom,AdN}$ at fixed heart rate of 220 bpm. We performed simulations for the ensemble of Figure 3, with the heart rate or $PS_{mom,AdN}$ set according to the x-axis. Shaded regions depict the 95% confidence interval of the prediction, black solid lines show the prediction for the optimised parameters (see Table 1).

It might be argued that the K_{ia} value of the mitochondrial CK should be set to 290 μM with oxidative phosphorylation active ([Jacobus et al., 1982]) to reflect functional coupling of CK to the adenine nucleotide translocator (ANT). Optimisation based on this K_{ia} value gives as result that on average 18% of the high-energy phosphate flux at a heart rate of 220 beats/min is transported in the form of PCr, the rest as ATP. The parameter values for $V_{max,Mi,f}$ calculated from rat heart mitochondria is 1609 ± 113 $\mu M/s$ in [Jacobus et al., 1982] and $V_{max,ATPsyn}$ is 2960 $\mu M/s$ which is about twice the value measured in the rabbit heart study analysed here. When using the rat heart parameters combined with $K_{ia} = 290$ μM, the contribution of PCr to high-energy phosphate transport is estimated to be 25%. Further analysis of a model which incorporates a microcompartment which functionally couples the mitochondrial creatine kinase to the adenine nucleotide translocator ([Vendelin et al., 2000]) shows that it is difficult to explain the response time and molecular kinetic parameters simultaneously with this model. The results of this analysis can be found in the supplemental material (section 3.8). The conclusion that the contribution of PCr to high-energy phosphate transport is relatively modest appears to be robust, because the contribution was estimated to be 15-17% in the ensemble study with rabbit heart parameters, see above, and does not become substantially higher in analyses with

other parameter sets.

3.4.4 Prediction of temporal energy buffering

The results described above indicate that direct ATP transport is predominant in working heart muscle. Given that PCr energy shuttling is of limited importance, we investigated another potential function of CK, i.e. temporal energy buffering. When ATP consumption by the myofibrils exceeds mitochondrial ATP production during muscle contraction, ATP homeostasis can be maintained by PCr [Greenhaff, 2001]. Ensemble predictions for $R_{diff,PCr}$, concentrations of cytosolic ADP and P_i and ATP synthesis rate at relative CK activity of 2, 100, and 300% of wild type levels are shown in Figure 3.6. Note that Mi-CK and MM-CK activities are both changed by the same factor in this set of simulations. Even at 3-fold increased CK activity, $R_{diff,PCr}$ does not increase dramatically (Figure 3.6). However, oscillations of cytosolic ADP concentrations are significantly affected by the CK activity. The amplitude of the ADP oscillation is 77 ± 11 μM at normal CK levels and becomes 146 ± 1 μM if CK is inhibited by 98%, as is the case for IA treated perfused hearts (Figure 3.6K,J). At threefold increased CK activity this becomes 36 ± 22 μM (Figure 3.6L). In simulations of a hypothetical case with 10000-fold increase of enzyme activity, oscillations of adenine nucleotide concentrations are almost fully damped to an amplitude of 2.6 ± 0.2 μM (data not shown).

The time course of mitochondrial ATP production oscillates with amplitudes of 566 ± 31, 215 ± 23 and 91 ± 14 $\mu M/s$ for 2, 100 and 300% relative CK activity, respectively (Figure 3.6G-I). The pulsatility of ATP and ADP concentrations and of ATP synthesis is synchronised to ATP hydrolysis in the myofibrils. The confidence regions for these trajectories are relatively narrow. By blocking CK by 98%, the average concentrations of ADP in the IMS increases to 64 ± 9 μM from 56 ± 9 μM at normal CK levels. In contrast to ADP, the amplitude of oscillations of cytosolic inorganic phosphate stays relatively constant at different CK activities at about 145 μM. This reflects that P_i is not directly buffered by CK. Surprisingly, average levels of cytosolic inorganic phosphate drop with CK activity. The average P_i concentration at 2% CK activity is 1618 ± 97 μM and becomes 1416 ± 80 μM for wild-type CK activity (Figure 3.6M,N). For all parameter sets in the ensemble the P_i concentration declines when CK activity is increased.

3.4.5 The specific role of the mitochondrial CK isoform

Transport of HEP by PCr from mitochondria to cytosol partially takes place via the circuit formed by both CK isoforms, but was predicted to be quantitatively not very important. On the other hand, temporal buffering of the systolic ATP hydrolysis burst needs only the MM-CK activity in the cytosol, which is much higher than the Mi-CK activity (see Table 3.1). It was therefore still unclear what the function of the mitochondrial CK isoform is.

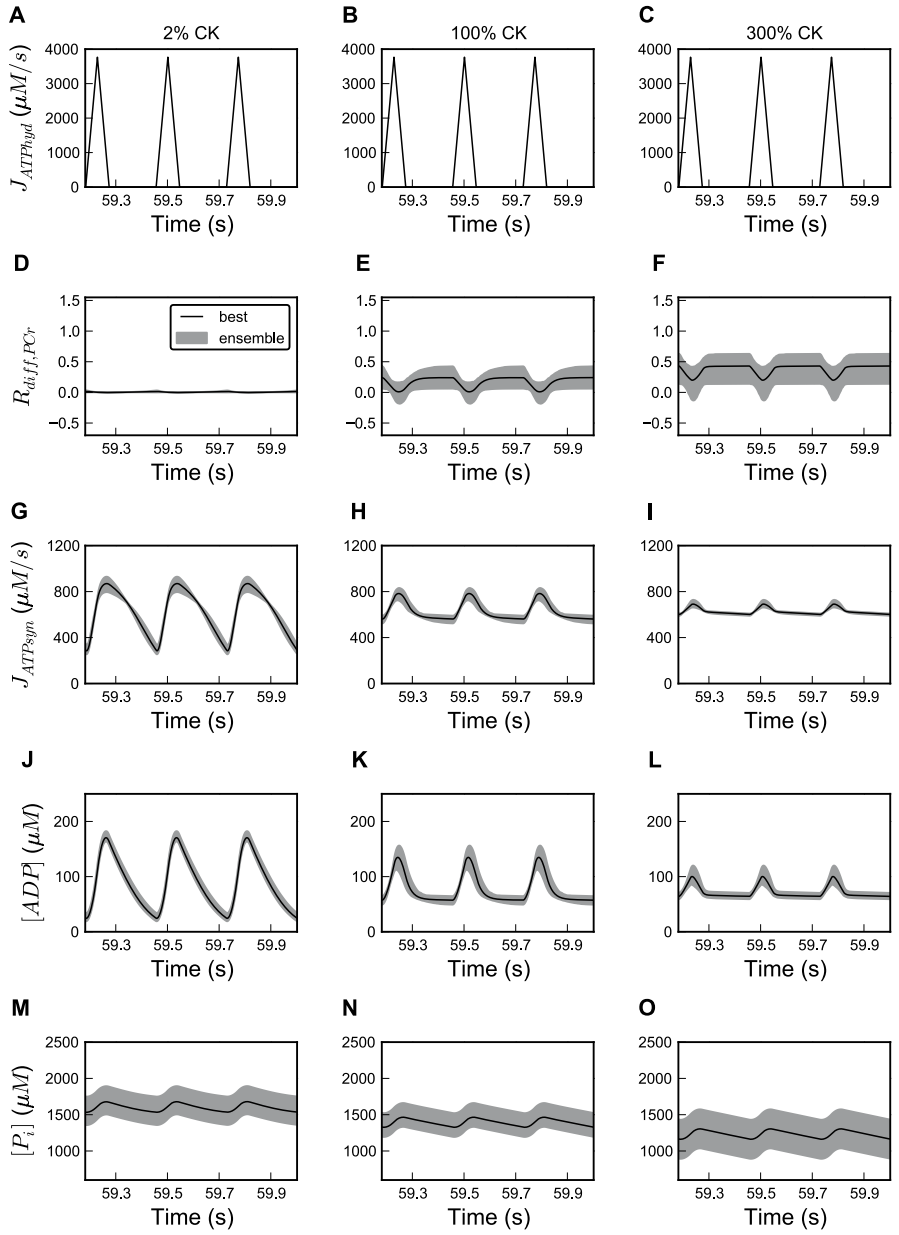

Figure 3.6: Fluctuations of metabolite concentrations and fluxes during the cardiac cycle at three levels of CK activity. Plots show (A-C) Trajectory of the forcing function of ATP hydrolysis and ensemble predictions of (D-F) $R_{diff,PCr}$, (G-I) mitochondrial ATP synthesis rate, (J-L) cytosolic ADP and (M-O) cytosolic Pi concentrations at heart rate 220 bpm. Mi-CK and MM-CK activities were set to 2, 100, and 300% of wild-type levels. Three cardiac cycles are shown at steady state. Solid lines show the simulated trajectory of the optimised parameter set (see Table 3.1). Shaded regions show the 95% confidence interval for all trajectories of the ensemble of 658 parameter sets. To alter CK activity, the parameters describing maximum enzyme velocity, $V_{max,Mif}$ and $V_{max,MMf}$, are changed in parallel to the indicated percentage.

In order to further elucidate the effect of the compartmentalised CK system on metabolism, we performed ensemble predictions with individual inhibition of both CK isoforms one by one. In Figure 3.7, we show the 95% confidence intervals of predicted metabolite concentrations and reaction fluxes. The amplitude of oscillations in mitochondrial ATP synthesis is predicted to rise from 215 ± 23 $\mu M/s$ at baseline CK activity to 278 ± 33 with 98% Mi-CK inhibition, compared to 375 ± 21 μM when MM-CK is inhibited by 98% (Figure 3.7I-K). Thus, despite its low activity, Mi-CK still has a small but clear effect on the ATP synthesis oscillation amplitude. Inhibition of Mi-CK has a larger effect when MM-CK is already inhibited (amplitude 565 ± 31 $\mu M/s$, Figure 3.7L). The damping of ADP oscillation is highly affected by MM-CK but not by Mi-CK: 98% inhibition of Mi-CK leads to an increase in the amplitude of systolic ADP oscillation from 77 ± 11 to 83 ± 11 μM (Figure 3.7M,N), whereas MM-CK inhibition doubles the amplitude to 146 ± 1 μM (Figure 3.7O).

Predictions of $R_{diff,PCr}$ illustrate that both Mi-CK and MM-CK are required for a functioning phosphocreatine shuttle. PCr diffusion averaged over the cardiac cycle makes a very small contribution to total HEP delivered from the mitochondria when either Mi-CK or MM-CK is inhibited by 98%. With 98% inhibited Mi-CK activity, $R_{diff,PCr}$ is even slightly below zero during diastole with low ATP hydrolysis, meaning that PCr is transported from cytosol to IMS (Figure 3.7F). Note that this situation is reversed with respect to normal Mi-CK and MM-CK activity where PCr diffusion is always positive during diastole and occasionally becomes negative during ATP hydrolysis peaks. For normal CK activity the explanation for reversed PCr diffusion during ATP hydrolysis (Figure 3.7E) was that the CK activity per unit volume is higher in the IMS than in the cytosol. During Mi-CK inhibition this is of course no longer the case and systolic PCr consumption in the cytosol leads to PCr diffusion from the IMS, explaining the reversal of PCr transport during systole. In contrast, with MM-CK inhibited, ATP is buffered by Mi-CK in the IMS and PCr diffuses to the IMS at the end of the ATP hydrolysis peaks. This explains why $R_{diff,PCr}$ goes more negative during ATP hydrolysis peaks with MM-CK inhibition and its oscillation is stronger than for normal Mi-CK and MM-CK activity (Figure 3.7E, G). When inhibiting Mi-CK activity, our model predicts an increase in the amplitude of [ADP] oscillation in the IMS from 57 ± 8 to 71 ± 8 μM. Mi-CK therefore has a damping effect on oscillations of ADP concentrations in the IMS, which contributes to the damping of mitochondrial ATP synthesis.

The concentration of cytosolic P_i is predicted to be lowered by mitochondrial creatine kinase activity. Blocking Mi-CK leads to a P_i increase by about 18% from 1416 ± 80 to 1670 ± 167 μM (Figure 3.7Q, R). If Mi-CK is inhibited by 100%, the steady state P_i concentration becomes 1678 ± 173 μM (data not shown). MM-CK inhibition decreases the P_i concentration; a combination of Mi-CK and MM-CK inhibition leads to a slightly higher P_i level compared to the wild-type (Figure 3.7S, T).

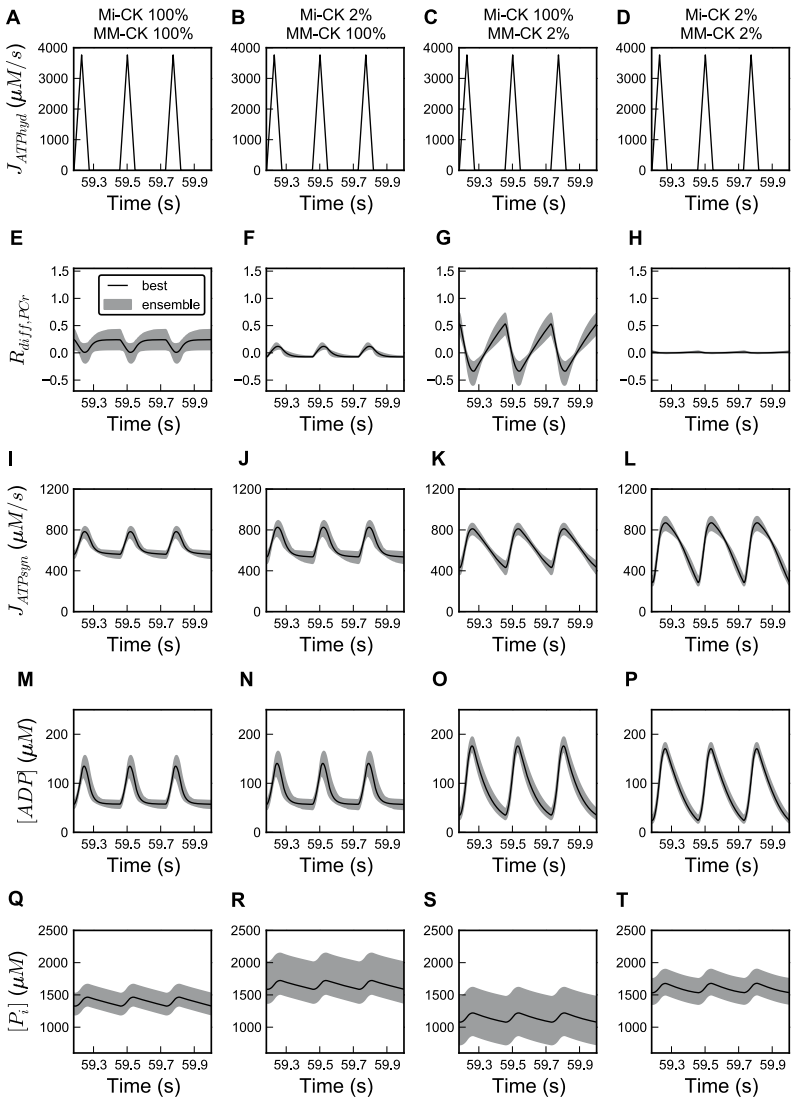

Figure 3.7: Ensemble predictions of metabolite concentration and flux oscillations during the cardiac cycle for selective CK isoform inhibition. In the first row (panels A-D), the pulsatile forcing function for ATP hydrolysis is plotted. Predictions of the time courses of (E-H) relative contribution of PCr to high-energy phosphate transport, $R_{diff,PCr}$, (I-L) ATP synthesis rate, (M-P) cytosolic ADP and (Q-T) P_i concentrations. Heart rate is 220 bpm. In the four columns we compare: no CK inhibition, 98% Mi-CK inhibition, 98% MM-CK, or both CK enzyme reactions inhibited by 98%. Black solid lines show the simulated trajectory of the optimised parameter set (Table 3.1). Shaded regions show the 95% central confidence interval for all trajectories of the ensemble of 658 parameter sets. To alter CK activity, the parameters describing maximum enzyme velocity, $V_{max,Mif}$ and $V_{max,MMf}$, are changed to the indicated percentage. Three cardiac cycles are shown after a steady state was reached. Note that the first and the last column also appear in Figure 3.6 and are shown here for ease of comparison.

3.5 Discussion

The relative importance of the different roles of the CK system in myocytes is still hotly debated [Greenhaff, 2001]. The present study was designed to investigate the function of CK in cardiomyocytes under varying workloads. In particular we sought to elucidate whether the phosphocreatine shuttle is the major pathway for HEP transfer from mitochondria to energy consuming myofibrils as stated in the phosphocreatine shuttle hypothesis or whether CK has other metabolic functions, e.g. the damping of swings in ATP and ADP concentrations and oxidative phosphorylation.

Various computational studies of cardiac energy metabolism have been published based on models which contained the creatine kinase reaction, ATP hydrolysis and synthesis. The model analysed in the present study is a subset of the model of Vendelin et al. ([Vendelin et al., 2000]) and was described previously [Kongas and van Beek, 2007, van Beek, 2007]. The diffusion gradients in the cytosol which had been shown to be very small ([Vendelin et al., 2000]) were replaced by a simple diffusion conductance. The adenine nucleotide translocator and phosphate carrier in the mitochondrial inner membrane and oxidative phosphorylation (OxPhos) reactions in the mitochondria in the model of Vendelin et al. were replaced by a Michaelis-Menten equation describing OxPhos flux as a function of ADP and P_i in the intermembrane space [van Beek, 2007]. The model was further modified in order to prevent thermodynamically infeasible loops by introducing constraints on the equilibrium of the CK reactions in IMS and cytosol [van Beek et al., 2009]. Some models in the literature implement substrate channelling between ANT and Mi-CK by a microcompartment which is located inside the intermembrane space [Vendelin et al., 2000, Aliev and Saks, 1997, Vendelin et al., 2004]. The performance of those models is discussed below. Other models exist for myocardial energy metabolism which do not consider the role of two creatine kinase isoforms connected via facilitated diffusion. For instance, Beard et al. integrated a detailed model of oxidative phosphorylation [Beard, 2005] with a model of spatially distributed oxygen transport and HEP metabolism to investigate the regulation of oxidative phosphorylation at different cardiac workloads [Beard, 2006] and HEP buffering in hearts at high workloads, acute ischemia and reactive hyperemic recovery.

In the present study we predicted the functions of the CK enzyme isoforms based on the following strategy. A set of experimental data from multiple scales was assembled. We based the analysis on our model which had been shown to contain the key elements of the CK system [Kongas and van Beek, 2007, van Beek, 2007]. The experimental data set allowed to estimate all parameters of this model. In order to set confidence regions for the predictions of CK function, the experimental errors for the data were taken explicitly into account. This was possible by generating an ensemble of model parameter sets. The probability of a set of parameters being part of the ensemble was determined based on the probability of the predicted experimental data set given the parameters. This approach was termed sloppy modelling [Brown and Sethna, 2003]. Brown et al. [Brown et al., 2004] and Gutenkunst et al. [Gutenkunst

et al., 2007b] applied it to time series of protein activity levels measured during dynamic responses of a system as a whole. The surprising finding in their studies was that responses of the system as a whole were predictable with acceptable confidence regions even if all parameters of the model were known with very poor accuracy. This is possible because the correlation between parameters is well defined by the behaviour of the system as a whole. A novel aspect in the present study is that we combined data taken from different integration levels in the system: kinetic parameters determined on enzymes in isolation, enzyme activity levels measured in tissue homogenates, organellar capacity levels measured on isolated mitochondria and dynamic response times determined on the heart as a whole. The whole organ response times were very important because they sensitively depend on the permeability of adenine nucleotides through the outer mitochondrial membrane, one of the organellar level parameters. This MOM permeability could not be determined experimentally with any degree of accuracy in isolated mitochondria. Combining some strategically important data from the whole system level with molecular parameters appears sufficient to predict system properties with acceptable confidence regions (Figures 3.4 - 3.7). Many of the experiments that are invoked to support high degrees of functional coupling between CK and ANT have been done in isolated mitochondria or in isolated myocytes and muscle fibres that were permeabilised. These were often studied at temperatures substantially below the physiological level. An important aspect of our analysis is that we try to estimate the functional roles of CK in the intact heart. To that end we combine the kinetic data from the molecular level with data obtained in isolated perfused hearts. It is important to realize that these hearts were intact, with contractility and cell membranes fully functional. Our model analysis explains the experimental data without invoking direct coupling of CK and ANT. However, the limited permeability of the mitochondrial outer membrane to adenine nucleotides, estimated from the response time in the intact heart, results in a certain degree of dynamic compartmentation of the adenine nucleotides. This approach helps to define the functional roles of CK in the intact heart at physiological temperatures. If CK-ANT direct coupling is the only way that ADP is delivered to the ANT, then the experiments with 98% inhibition of CK cannot be explained. It would then also be hard to explain that Mi-CK knockout animals still have substantial cardiac contractile function [Saupe et al., 1998]. Future CK-ANT interaction models need to address such experimental data sets with CK inhibition and also explain the phosphate-labelling data of Erickson-Viitanen et al. [Erickson-Viitanen et al., 1982].

Our findings suggest that the principal role of the CK system in heart muscle is to serve as a temporal energy buffer for ATP and ADP at the 100 millisecond time scale. CK's role in supporting transport of high energy phosphate groups seems of limited importance. If oxygen supply is interrupted, PCr will also buffer ATP and ADP for several seconds [Beard, 2006]. Temporal energy buffering therefore has a relatively large bandwidth. Joubert et al. experimentally investigated the role of the CK shuttle by ^{31}P NMR magnetisation transfer protocols in vivo and proposed the hypothesis of a versatile role of PCr on intracellular energy transport depending on cardiac activity

[Joubert et al., 2004, 2002]. Partial inhibition of ATP synthesis led to a decrease of indirect energy transport via PCr. This decrease is predicted by our model (data not shown). Some computational models on compartmentalised energy transfer in muscle, as for instance in Vendelin et al. ([Vendelin et al., 2004]), assume restricted diffusion of adenosine nucleotides to an extent where energy transport via PCr becomes essential. However, a large restriction of adenine nucleotide permeation of the cytosol and MOM is not compatible with the relatively fast responses of oxidative phosphorylation to cytosolic workload steps [van Beek, 2007].

The conductance parameter $PS_{mom,AdN}$ in our model reflects not only the permeation of the MOM proper but in series with that also diffusion in myofibrils and cytosol. The inverse of $PS_{mom,AdN}$ in our model is therefore the sum of the inverse of permeability-surface products (PS) for the MOM proper and cytosol, respectively [van Beek, 2007]. The present Monte-Carlo ensemble approach indicates that $PS_{mom,AdN}$ lies within a range from 7.4 to 115 s^{-1} (see section 3.3). Based on the transverse diffusion coefficient of 52 μm^2 for ATP in the myofibrillar space measured with fluorescently labelled ATP [Vendelin and Birkedal, 2008], the PS calculated for the cytosol is 216.7 s^{-1} [18]. Given an ensemble mean $PS_{mom,AdN}$ of 31.7 s^{-1} (see 3.1) we predict that about 15% of the total resistance to diffusion can be attributed to the cytosol. Note that the fluorescently labelled ATP has a higher molecular mass than ATP. The true diffusion coefficient of ATP is probably higher and the contribution of the cytosolic space to diffusional resistance is therefore probably overestimated in this calculation. The contribution of PCr to HEP transport predicted in the present study (section 3.4) is compatible with measured response times of the system (section 3.2). It has been suggested that in cardiomyocytes the density of mitochondria and their vicinity to myofibrils is sufficient to ensure energy transport via adenosine nucleotides [Meyer et al., 1984]. The prediction by our model that CK-facilitated transport of PCr is not obligatory for HEP transport is in line with the observation that CK knockout has relatively mild effects on cardiac function [Gustafson and Van Beek, 2002, Saupe et al., 1998, Veksler et al., 1995].

Activation of oxidative phosphorylation has been proposed to be strongly dependent on substrate channelling of ATP and ADP between the tightly coupled enzymes Mi-CK and ANT, meaning that ATP exported from the mitochondrial matrix via ANT is immediately available as a substrate for Mi-CK. The resulting ADP is then channelled back to stimulate oxidative phosphorylation in the mitochondrial matrix. However, the hypothesis of functional coupling is still debated [Beard and Kushmerick, 2009] and other studies seem to contradict it [Lipskaya and Savchenko, 2003]. In order to investigate the effect of functional coupling between the ANT and Mi-CK we implemented and analysed the model of Vendelin et al. ([Vendelin et al., 2000]), where the reactions are coupled by a microcompartment (section 3.8). The model, which contains constants which phenomenologically reflect the functional coupling of Mi-CK to the ANT is considered to give a good and computationally effective representation of the functional coupling between Mi-CK and oxidative phosphorylation [Vendelin et al., 2004]. It appeared to be rather difficult to fit the model of Vendelin et al. to the

given experimental data of mitochondrial delay times (t_{tmito}) when measurements on molecular kinetic parameters are taken into account in the cost function. Especially at low workloads, a quicker response to a step in ATP consumption rate after CK inhibition could not be predicted with this model. Even when all parameters from the model of Vendelin et al. were variable during the optimisation procedure, the quality of the fit is far from optimal despite the fact that the model of Vendelin et al. has about three times as many parameters as our present model. We therefore do not consider the microcompartment explicitly in our present study.

The present results suggest that most of the delay of the activation of oxidative phosphorylation after temporal changes in ATP hydrolysis is caused by the delay of changes in phosphate metabolite levels outside the inner mitochondrial membrane. To investigate whether processes inside the mitochondria delay the response further, we tested a model of the mitochondrial matrix including metabolite transport across the inner mitochondrial membrane with instantaneous step changes in ADP or Pi and also with ADP and Pi simultaneously outside the inner mitochondrial membrane. This corresponds to the model applied in section 3.8 with all processes outside the inner mitochondrial membrane removed and the ADP and Pi concentrations outside the inner mitochondrial membrane set as forcing function. After a 20% increase in ADP concentration, ATP synthesis in the mitochondria reached a steady higher level within one second. The response time, calculated as for t_{mito}, was 0.4 s. For a step in P_i the response was even faster with a negative value for the response time of 20.3 s because the response showed an overshoot. For a simultaneous change in ADP and Pi the mitochondrial response was essentially complete within half a second, with a response time of 0.08 s. When extra-mitochondrial ADP is changing, both mitochondrial oxygen consumption and ATP efflux via the ANT reacted even faster than the ATP synthase reaction. The fast response of mitochondrial metabolism predicted by the model is in agreement with spectroscopic measurements of the oxidation state of the electron carrier cytochrome b which was oxidised with a half-time of 70 milliseconds after a step in extra-mitochondrial ADP concentration at 26 $^{\circ}C$, and presumably much faster at the physiological temperature [Chance, 1965].

In studies on isolated rabbit cardiac muscle mitochondria the direct contribution of mitochondrial ATP to PCr formation by Mi- CK is low [Erickson-Viitanen et al., 1982]. It was shown with radioactively labelled phosphate groups that if the concentration of ATP in the environment of the mitochondria is larger than 0.2 mM, less than 6% of PCr synthesis uses ATP synthesised immediately beforehand in the mitochondrial matrix. This is incompatible with a model where a major part of PCr is synthesised from ATP directly transferred to creatine kinase via a very small compartment with limited exchange with its environment.

By *in silico analysis*, we inferred distinct roles for the mitochondrial and myofibrillar creatine kinase enzymes. MM-CK is mainly responsible for damping large swings in metabolite concentrations and large oscillations in the rate of oxidative phosphorylation which would otherwise be caused by the large peaks of ATP hydrolysis during the cardiac cycle. Mi-CK restricts high concentrations of inorganic phosphate, which

is surprising considering that inorganic phosphate is not handled directly by CK. Despite its low activity, Mi-CK also decreases oscillations of ATP synthesis, mainly due to the effect of Mi-CK on ADP oscillations in the intermembrane space.

The effect of the CK isoforms on the buffering of ADP oscillations and the prevention of high concentrations of inorganic phosphate may play a role in the prevention of formation of reactive oxygen species (ROS). ROS production highly depends on the mitochondrial membrane potential, which is increased at low ADP levels [Meyer et al., 2006, Korshunov et al., 1997]. The electric membrane potential in mitochondria can also be altered by inorganic phosphate, leading to enhanced ROS release [Oliveira and Kowaltowski, 2004]. Low ADP concentrations during diastole are prevented by MM-CK according to our predictions (see section 3.7). A protective role of Mi-CK against oxygen radical formation by preventing high inorganic phosphate concentrations is also predicted by our model. A function of Mi-CK to prevent oxygen radical formation has been found experimentally in isolated brain mitochondria [Meyer et al., 2006]. The energy buffering role of the CK system has been linked to the prevention of oxidative stress in neurons [Klivenyi et al., 1999, Brewer and Wallimann, 2000]. Creatine supplements to nutrition have also been shown to have a neuroprotective effect in models of Huntington's disease [Hersch et al., 2006, Matthews et al., 1998]. The effects of creatine as a nutritional supplement in health and disease have recently been reviewed by Wallimann et al. [Wallimann et al., 2011].

In conclusion, we showed that by using a relatively small 'skeleton' model we were able to explain the dynamic adaptation of cardiac energy metabolism to changing workloads and to discern different functions of distinct CK isoenzymes. The sloppy modelling approach enables to make useful predictions of CK system behaviour despite limited experimental input data and limited knowledge of kinetic parameters. The concept of sloppy modelling can also be used to find optimal experimental designs to further test the model [Casey et al., 2007]. We also demonstrated that combining a computational model analysis with experimental data on the level of cellular organelles and isolated enzymes and with the response of the heart as a whole provides a powerful combination that gives valuable insights in the functional roles of CK, such as regulation of oxidative phosphorylation, energy transport, inorganic phosphate levels and buffering of peaks of ATP hydrolysis at the 100 millisecond time scale.

3.6 Methods

3.6.1 Computational model

For our analysis, we employed a previously published computational model [van Beek, 2007]. It is available in various formats and can be found in the BioModels database [Le Novere et al., 2006] as well as in the CellML model repository [Lloyd et al., 2008]. The model incorporates the key elements of the CK system with ATP synthesis in the mitochondria and pulsatile ATP hydrolysis in the cytosol (see section 3.1). The input

of the model is a forcing function of cytosolic ATP usage catalysed by myosin-ATPase and ion pumps. The model contains ten ordinary differential equations (ODEs) describing the rate of change of each metabolite concentration (ADP, ATP, PCr, Cr, P_i) in two compartments over time. These equations were extensively described previously [van Beek, 2007]. Model dynamics depend on 22 kinetic parameters retrieved from the literature which are listed in section 3.1. In general the kinetic constants retrieved from the literature have relatively modest standard errors. However, for the permeability of the MOM to ATP and ADP (assumed to be equal in the model analysis; cf. [Vendelin et al., 2000]), reported values differed from 0.16 [Vendelin et al., 2000] to 85 $\mu m * s^{-1}$ in the model of Beard [Beard, 2005] based on measurements of Lee et al. [Lee et al., 1994]. This large variation is possibly due to mitochondrial isolation or cell membrane permeabilisation procedures.

The mitochondrial outer membrane permeability-surface product parameter $PS_{mom,AdN}$ influences the response time for dynamic adaptation of oxidative phosphorylation strongly. Therefore the dynamic measurements of venous oxygen outflow in the heart as a whole in response to an increase of heart rate allow estimating the mitochondrial membrane permeability at the organellar level. The whole heart measurements were corrected for oxygen transport delay to reflect events at the level of the mitochondria (see below). The mitochondrial response time t_{mito} is defined as the generalised time constant of the time-course of oxygen consumption (defined to be equivalent to the first central statistical moment of the impulse response function in case the system is linear), previously described in [van Beek, 2007, van Beek and Westerhof, 1990, van Beek et al., 1998, Van Beek and Westerhof, 1991]. From a model simulation, tmito is calculated as follows:

$$t_{mito} = \int_{t_{step}}^{t_{end}} \frac{J_{ATPhyd,test} - J_{ATPsyn}(t)}{J_{ATPhyd,test} - J_{ATPhyd,basis}} dt \tag{3.3}$$

Where $J_{ATPhyd,basis}$ and $J_{ATPhyd,test}$ are the values for the ATP hydrolysis rates for the two electrically paced heart rates at baseline and test level, averaged over the cardiac cycle; J_{ATPsyn} denotes the time course of ATP synthesis in the mitochondrion. t_{step} is the time point when the heart rate is increased and t_{end} is the time point of the final oxygen measurement. Note that the average J_{ATPsyn} in the steady state before and at the end of a test challenge equals $J_{ATPhyd,basis}$ and $J_{ATPhyd,test}$, respectively.

In order to correspond with the experimental conditions in [Harrison et al., 2003], t_{end} was set to 60 seconds with $t_{step} = 0$ seconds; an initial run for 40 seconds before the heart rate step ensures that ATP synthesis has adapted to ATP hydrolysis and is found to be in steady state at this stage. In order to investigate the damping capabilities of the modelled system, ATP hydrolysis is simulated as a pulsatile function representing the alternating nature of energy demand in systole and diastole as described in [van Beek, 2007]. Figure 3.8 shows the dynamic response of mitochondrial ATP production in response to a step in heart rate and ATP hydrolysis.

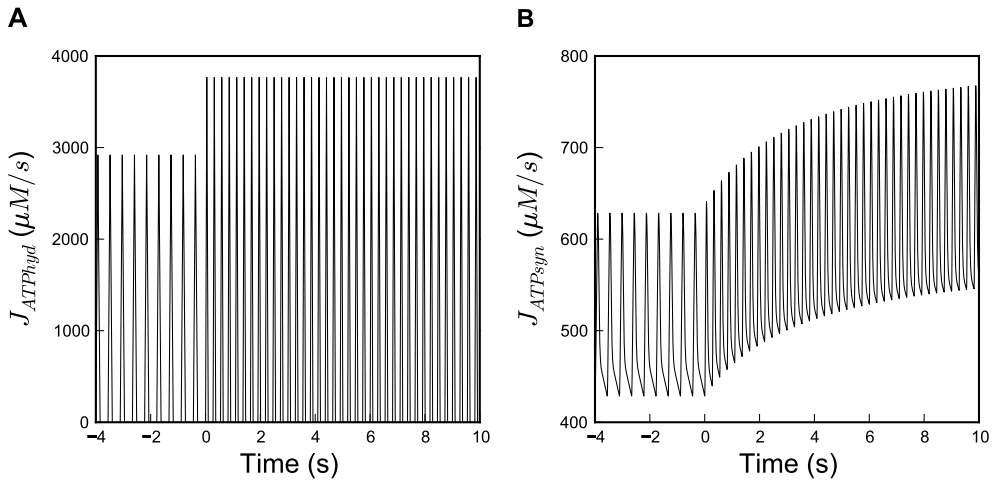

Figure 3.8: Pulsatile nature of energy production and consumption in the beating heart and the response to a step in heart rate. Shown are the time courses of (A) ATP hydrolysis and (B) synthesis simulated with the model of Figure 1. At time 0 s, average ATP hydrolysis rate was increased from 486.5 to 627.6 $\mu M * s^{-1}$ corresponding to an increase in heart rate from 135 to 220 bpm, as was imposed in the experiments which were simulated in this study. Please note the difference in scale of the y-axis between panels A and B.

3.6.2 Sloppy ensemble modelling

Almost all models in systems biology contain parameters that cannot be determined precisely. It is common practice to estimate missing parameter values by a parameter fit to experimental data. After the fit, one can make model predictions and analyse the underlying biological processes. This, however, is dangerous because a range of parameter combinations may agree with the available data equally well, potentially leading to deviating model predictions of new experimental situations. Directions in parameter space where parameter changes do change the simulation outcome very little were termed 'sloppy' by Brown et al., whereas directions where small changes in parameter values affect the dynamic behaviour of the system strongly were termed 'stiff' [Brown and Sethna, 2003]. Sloppy parameter sensitivity spectra have been identified for numerous biological models by the analysis of the eigenvectors and eigenvalues of a sensitivity matrix calculated from the chi-square cost function [Gutenkunst et al., 2007a]. Sloppy models exhibit a characteristic pattern with the logarithms of eigenvalues approximately uniformly distributed over a large range. A sensitivity analysis of the CK model revealed the presence of both stiff and sloppy parameter combinations and a 'sloppy' sensitivity spectrum [van Beek et al., 2008]. Since our model shows sloppy parameter sensitivities and is based on data subject to experimental variation, drawing predictions from an ensemble of parameter sets is preferable to merely relying on one parameter set fit to experimental data. According to the sloppy modelling paradigm ([Brown and Sethna, 2003, Gutenkunst et al., 2007a]), the probability of a set of model parameters θ to be included in the ensemble

is proportional to its likelihood to describe given experiment data D multiplied by the likelihood of prior experimental information about the parameter values themselves. The sampling process is thus based on Bayesian inference of a posterior distribution of parameter sets $P(\theta|D) = P(D|\theta) * P(\theta)$, where $P(D|\theta)$ is the likelihood of the data given a parameter set, $P(\theta)$ is the prior probability of the parameter set based on experimental prior knowledge on single parameter values and the posterior $P(\theta|D)$ is the probability of a parameter set to describe the given experimental data. The construction of the ensemble with a Markov-Chain Monte Carlo (MCMC) method was done with the Metropolis-Hastings algorithm [Gutenkunst et al., 2007b]. The Sloppy cell software environment, used for the analysis, was adapted to process all operators which were in the SBML file describing the model. The modified version is provided online[1]. To speed up convergence, Sloppy Cell takes larger steps along sloppy directions and smaller steps along stiff directions in parameter space; this 'importance sampling' is described in [Brown et al., 2004, Brown and Sethna, 2003].

3.6.3 Experimental data

Measured values of molecular model parameters and their provenance, extracted from the scientific literature, are listed in table 3.1. For nine of the 22 parameters reliable standard measurement errors could be found. In addition to the direct measurements on molecular parameters, we employ t_{mito} values from a study by Harrison et al. where the effects of inhibiting creatine kinase and different sizes of electrically paced heart rate jumps in rabbit hearts were investigated [Harrison et al., 2003]. Isolated hearts were perfused with Tyrode's solution containing among others glucose and pyruvate to provide substrates for energy metabolism. In our dataset we include two experimental conditions where hearts were exposed to either (i) iodoacetic acid (IAA) to block glyceraldehyde-3-phosphate dehydrogenase (GAPDH) or (ii) iodacetamide (IA) to inhibit both CK and GAPDH. In order to provide a sufficient amount of reducing equivalents to fuel aerobic respiration despite the inhibition of glycolysis, the buffer also contained pyruvate.

Adenosine was added to the Tyrode buffer to ensure that oxygen supply is non-limiting when oxygen consumption is recorded. The whole heart measurements were corrected for the O_2 transport time in the coronary vessels based on a model of oxygen transport by convection in blood vessels and diffusion through tissue. The t_{mito} therefore reflects the response time at the level of the mitochondria (cf. [Harrison et al., 2003] and references cited there).The mean response time was also corrected for a small deviation from an ideal step in beat-to-beat ATP hydrolysis measured as an initial overshoot in rate-pressure product [van Beek and Westerhof, 1990]. For each condition, steps in heart rate were imposed from 135 to 160, 190 and 220 beats per minute, respectively, using electrical pacing. Note that glycolysis is always inactive when the dynamic response is measured, which corresponds to the absence of glycolysis in the computational model. This approach made it possible to isolate the

[1] `http://www.ploscompbiol.org/article/info\%3Adoi\%2F10.1371\%2Fjournal.pcbi.1002130`

contribution of the CK system from the contribution of glycolysis, which removes substantial complexity from the model analysis.

A step in ATP hydrolysis from 486.5 to 627.6 $\mu mol*l^{-1}$ cell water $*s^{-1}$ corresponds to a step in the electrically paced heart rate from 135 to 220 bpm, as was estimated from measurements of myocardial oxygen consumption [van Beek, 2007]. From these values, we linearly interpolated hydrolysis rates of 531.4 and 579.5 $\mu mol*l^{-1}$ cell water $*s^{-1}$ for heart rates 160 and 190 bpm, respectively. To simulate CK inhibition by IA the model parameters for the maximum velocities of both enzyme reactions were set to 2.3% of their original values, corresponding to the CK activity measured for the inhibited hearts. Note that the enzyme activities, the mitochondrial capacities and the whole organ dynamic response times were all measured in the same experimental model by the same laboratory.

3.6.4 Cost function

Model parameters are fitted to experimental data using a modified Levenberg-Marquardt least-squares procedure in logarithmic parameter space, which is part of the Sloppy-Cell modelling environment. For our model and data we calculate the cost C for a given parameter set θ as follows:

$$C(\theta) = \frac{1}{2}\sum_{c}\frac{y_c(\theta) - d_c}{\sigma_{d_c}} + \sum_{i} prior(\theta_i) \tag{3.4}$$

with y_c being the model prediction of tmito (Equation 3.3) as a function of the parameter value θ and d_c the measured value for condition c with standard error σ_{d_c}. The first term of the cost function takes into account the experimental data on the whole heart level, whereas the second term represents prior experimental information about parameter values found in the literature or measured in conjunction with the modelled experiments. The prior cost, which gives a penalty for a parameter θ_i for drifting to far from its measured value θ_i*, is calculated as in [Gutenkunst et al., 2007a]:

$$prior(\theta_i) = \frac{1}{2}\left(\frac{ln\theta_i - ln\theta_i*}{\sigma_{ln\theta}}\right) \tag{3.5}$$

Note how the prior is used to enter experimentally measured information on parameters measured at the molecular level in the second term of Equation 3.4, while the first term contributes measured information on the whole system response. The deviations of the predicted response times from their measured values are penalised relative to their measured standard errors and the deviation of the molecular parameters from measured values are penalised relative to their reported standard errors. Values for molecular parameters reported in the literature are usually given as mean

and standard error. However, in the sloppy modelling framework, it is preferable to choose a normal distribution in log space

[Brown et al., 2004, Gutenkunst et al., 2007b,a]. A Gaussian distribution of logarithmic parameters has been proposed to be biologically plausible [Schaber et al., 2009]. This forms a convenient way to deal with dimensionless positive quantities as parameter values [Liebermeister and Klipp, 2005].

In order to calculate the $\sigma_{ln\theta}$ value for a parameter θ in log space from its reported standard error (considering the span of a 95% confidence region), we set the value as follows:

$$\sigma_{ln\theta} = \frac{1}{4} ln \left(\frac{\theta_i + 2SE}{\theta_i - 2SE} \right) \tag{3.6}$$

where SE is the absolute standard error of parameter θ_i. If the standard error is small relative to the mean of the parameter, the shapes of the prior distributions become approximately normal (see Figure 3.3). Since standard errors for only nine of all 22 system parameters could be found, we chose the default $\sigma_{ln\theta}$ value for the remaining parameters to be at the maximum of all $\sigma_{ln\theta}$ values for parameters with known error. This maximum was the error of the parameter for the binary dissociation constant for creatine from Mi-CK ($K_{ib,Mi}$, and see Table 3.1). In order to investigate the effect of altered default prior standard deviation on posterior parameter distributions and ensemble predictions, we performed several additional ensemble simulations with lower and higher default values. Results of these simulations can be found in section 3.7. The parameter describing MOM conductance for adenine nucleotides, $PS_{mom,AdN}$, could not be reliably determined by experiments on the organellar level and was therefore not constrained by a prior.

3.6.5 Determining prediction uncertainty: Ensemble simulations

A first estimate of parameter values was determined by a least-squares fit to the data, using the cost function of equation 3.4. This initial best parameter estimate resulting from the optimisation is used as the starting point for a walk through the parameter space using the Metropolis-Hastings algorithm. Starting the random walk from the optimised set of parameters made the algorithm converge more quickly to the posterior distribution. We use the algorithm's implementation in SloppyCell to sample parameter sets with probability density proportional to $exp(-C(\theta))$. All scripts to reproduce the presented calculations can be found online[2] To ensure that the members of the ensemble are statistically independent, we 'prune' the ensemble by taking only every n^{th} sample, where n is the maximum correlation time of all parameters.

[2] http://www.ploscompbiol.org/article/info\%3Adoi\%2F10.1371\%2Fjournal.pcbi.1002130

The correlation time of a parameter is defined as the time constant of its autocorrelation function. For our model, taking 50000 steps in the random walk is sufficient to obtain more than 600 independent parameter sets. The independent parameter sets in the ensemble provide the final estimate of the parameters, not only characterised by a mean but also by a standard deviation which reflects the spread of the estimation. Calculations were executed in parallel on a ClusterVision parallel machine with 16 nodes of four 3GHz processors with 4GB RAM. For computational performance reasons, we calculated model simulations for parameter estimation and ensemble sampling with an ATP hydrolysis rate averaged over the cardiac cycle rather than the pulsatile pattern shown in Figure 3.8. This reduced the time needed for calculations tremendously, making it feasible to do the ensemble calculations in several hours.

However, to investigate the damping characteristics of the system, we use a pulsatile forcing function of ATP hydrolysis (see Figure 3.8A) [van Beek, 2007]. To assess the differences in metabolite levels and fluxes caused by replacing the pulsatile function with a time-averaged continuous function, 1000 parameter sets were randomly drawn from all parameter sets tried in the Monte-Carlo random walk, to compare the values of model results between pulsatile and non-pulsatile simulations. The variables most affected by the pulsatile approximation are $R_{diff,PCr}$ and t_{mito}. The difference between pulsatile vs. non-pulsatile simulations of all 1000 parameter sets is 7.6±4.3 and 6.8±1.5% (mean±SD), respectively. tmito values from non-pulsatile simulations are always slightly smaller than values from a pulsatile simulation, but their deviation is smaller than the standard error of the experimental tmito data. The difference between pulsatile and non-pulsatile model results for other variables is below 4.5% of their average values in a non-pulsatile setting.

3.7 Supplemental text 1: Ensemble predictions with different default prior standard deviations

3.7.1 Ensemble Simulations

For 13 of the 22 model parameters, no reliable standard measurement error could be found in the literature. For those parameters, we assumed the default prior standard deviation corresponding to the highest available measurement error ($\sigma_{ln\theta} = 0.336$; reflecting the reported error of $K_{ib,Mi}$ which is about 30% [Jacobus et al., 1982]. In order to test how a different default prior affects posterior parameter distributions and predictions, we performed three additional ensemble predictions with a $\sigma_{ln\theta}$ for parameters with unknown errors of (i) 0.107, which is calculated from the mean of all $\sigma_{ln\theta}$ for reliably known measurement errors, (ii) 0.637 and (iii) 1.009, which is double and triple the $\sigma_{ln\theta}$ for Kib,Mi, respectively. Table 3.2 shows mean parameter values and the prior and posterior standard deviations for all three ensembles. In general, posterior standard deviations increase with a higher prior $\sigma_{ln\theta}$. For most parameters, the standard deviation of its posterior distribution in the ensemble (in log-space) is

Parameter	Ensemble mean ± SD default $\sigma_{ln(\theta)}$=0.107	Ensemble mean ± SD default $\sigma_{ln(\theta)}$=0.673	Ensemble mean ± SD default $\sigma_{ln(\theta)}$=1.009
$K_{eq,CK}$	152.27±3.91	152.47±3.9	152.22±4.13
$^{*}V_{max,Mi,f}$	886.61±97.52	762.75±393.77	668.78±524.62
$K_{ia,Mi}$	749.62±61.07	751.8±63.69	744.45±60.57
$K_{ib,Mi}$	30032.9±10318.4	30056.72±9295.1	31737.3+11647.7
$^{*}K_{ic,Mi}$	206.48±22.91	254.96±172.29	277.93±368.08
$K_{id,Mi}$	1604.3±213.01	1571.7±194.88	1619.82±214.37
$K_{b,Mi}$	5203.2±286.04	5213.49±296.48	5209.05±308.39
$K_{d,Mi}$	502.39±19.91	501.01±19.61	498.87±19.25
$^{*}V_{max,MM,f}$	11391.1±1209.4	8937.51±5186.11	10207.3±9436.2
$^{*}K_{ia,MM}$	900.34±97.96	1353.01±1030.03	1559.29±1574.26
$^{*}K_{ib,MM}$	35045.9±3818.3	44952.01±34756.14	54320.8±69811.6
$^{*}K_{ic,MM}$	224.17±24.6	256.4±185.15	348.16±467.35
$^{*}K_{id,MM}$	4786.54±481.17	5587.71±4428.64	6338.79±7203.39
$K_{b,MM}$	15868.9±2673.5	16076.45±2828.23	15980.34±2744.0
$K_{d,MM}$	1667.44±38.2	1673.59±37.56	1679.14±41.17
$V_{max,syn}$	1303.72±55.71	1399.88±132.62	1456.75±151.98
$^{*}K_{adp}$	25.8±2.96	35.12±8.91	33.2±13.52
$^{*}K_{pi}$	799.47±84.7	147.14±58.13	107.42±45.63
$PS_{mom,AdN}$	200.83±286.48	12.49±2.92	11.22±2.32
$^{*}PS_{mom,PCr}$	156.65±17.07	184.92±136.09	229.31±235.54
$^{*}PS_{mom,Cr}$	157.18±16.63	194.64±153.8	262.71±305.25
$^{*}PS_{mom,pi}$	196.77±20.96	243.11±168.61	304.62±317.18

Table 3.2: Shown are mean values and standard deviations for three parameter ensembles, generated with prior $\sigma_{ln(\theta)}$ values of 0.107, 0.673 and 1.009 for parameters with unknown standard error.
*Parameters for which the default prior was used.

close to the prior. However, some parameters, e.g. $V_{max,Mi,f}$, K_{adp} or K_{pi} only show a slight increase of the standard deviation in the posterior distribution with increased default prior (see Table 3.2), which indicates that they can be estimated relatively well, since they are more strongly constrained by the data. Remarkably, the posterior standard deviation for $PS_{mom,AdN}$ increases with decreasing default $\sigma_{ln\theta}$ for the other parameters. Since a smaller prior value decreases variability of the affected parameters in the sampling process, $PS_{mom,AdN}$, being the only unconstrained parameter, can compensate for the decreased flexibility when fitting the data and therefore its standard deviation in the ensemble may be increased.

3.7.2 Predictions

In order to test the effect of altered default $\sigma_{ln\theta}$ values on model predictions of the contributions of PCr to high-energy phosphate transport across the mitochondrial outer membrane and the buffering of ATP synthesis rate and ADP concentrations, we performed ensemble predictions for all three parameter ensembles with the altered default priors.

Figure 3.9 shows the prediction of the relative PCr contribution to high-energy phosphate flux across the mitochondrial outer membrane (Rdiff,PCr) for normal and inhibited CK activity. With a low default prior standard deviation, we predict PCr to contribute not more than 20% to energy transport (Figure 3.9A). As expected,

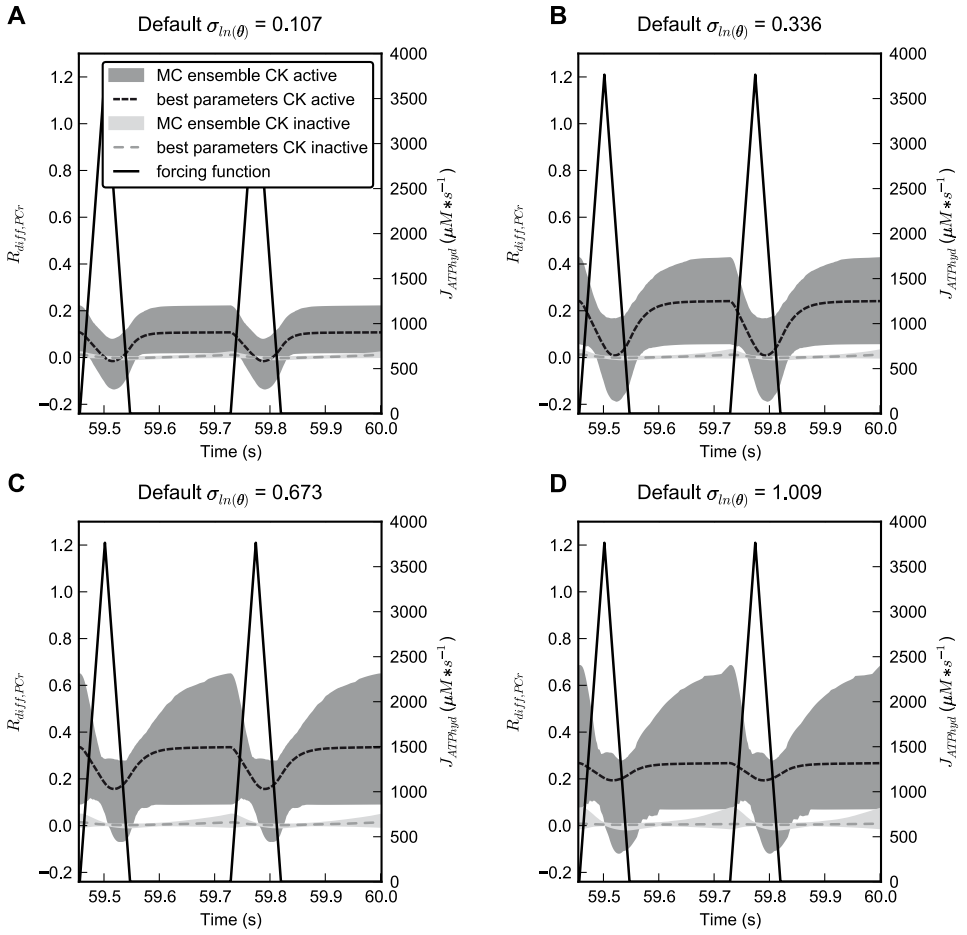

Figure 3.9: Ensemble predictions of energy transport from mitochondria to cytosol by PCr for different default priors. Prior $\sigma_{ln\theta}$ for parameters with unknown standard error was set to (A) 0.107, (B) 0.336, (C) 0.673 and (D) 1.009. Plotted are the relative PCr contribution to high-energy phosphate flux across the mitochondrial outer membrane ($R_{diff,PCr}$) for active and inactive (inhibition by 98%) creatine kinase and the forcing function of pulsatile ATP hydrolysis in the myofibril. Plotted regions represent the 95% confidence interval for all predictions in an ensemble. Please note the different scales for left and right y-axes in each panel.

the prediction becomes more uncertain when increasing the prior standard deviation. However, at a prior $\sigma_{ln\theta}$ of 1.009, although $R_{diff,PCr}$ slightly exceeds 60% during diastole (Figure 3.9D), the average predicted PCr contribution still is relatively low at about 25%. Ensemble predictions of the damping of oscillations in ATP synthesis rate and cytosolic ADP concentrations are shown in Figure 3.10.

The amplitude of the time course of ATP synthesis increases from 117.4 ±10.5 $\mu M * s^{-1}$ to 431.7±19.3 $\mu M * s^{-1}$ if CK is inhibited, given a low $\sigma_{ln\theta}$ of 0.107 (Figure 3.10A). With the highest prior standard deviation we tested ($\sigma_{ln\theta} = 1.009$), the amplitude of ATP synthesis is higher and the prediction more uncertain (309.9±68.1 $\mu M * s^{-1}$ at active and 720.1±63.6 $\mu M * s^{-1}$ at inactive CK, see Figure 3.10G). The effect of higher prior standard deviations on the amplitude of the ADP concentration is smaller than the relative effect on the amplitude of ATP synthesis. At the smallest tested default prior, the prediction of the amplitude in ADP concentration for active and inactive CK conditions is 57.8±5.7 and 144.5±2.5 μM, respectively (see Figure 3.10B). The ensemble with the highest default prior standard deviation predicts the amplitude in ADP concentration to be 81.3±24.7 μM (CK active) and 147.2±11.5 μM (CK inactive), which is shown in Figure 3.10H.

3.8 Supplemental text 2: Model analysis with additional microcompartment which couples CK to the adenine nucleotide translocator

The Mi-CK may be functionally coupled to oxidative phosphorylation via the adenine nucleotide translocator (ANT) [Jacobus et al., 1982, Vendelin et al., 2000, 2004]. Compartmentation of adenine nucleotides in a separate compartment between ANT and Mi-CK cannot completely explain the change in apparent ATP dissociation constants from Mi-CK during activation of oxidative phosphorylation, but phenomenological models and constants for dynamic compartmentation are considered sufficient to describe functional coupling in models of energy fluxes in the cell [Vendelin et al., 2004]. We therefore investigate the potential effects of functional coupling on our conclusions by analysis of the experiments with the model of Vendelin et al. [Vendelin et al., 2000]. In that model ATP and ADP exchange via the ANT between the mitochondrial matrix and a microcompartment which also gives access to Mi-CK. The extremely small volume of the microcompartment enables a direct substrate-product channelling between ANT and Mi-CK. We simulated the experimental data with the parameters in [Vendelin et al., 2000]. For the heart rate steps from 135 to 160, 190 and 220 beats/min t_{mito} was 0.19, 0.19 and 0.2 s, respectively, in the model simulation with CK fully active, to be compared with 3.0, 3.9 and 5.25 s in the experiments. With CK inhibited by 98%, tmito was 14.7, 11.3 and 10.2 s, while these values were 1.8, 2.5 and 3.3 s in the experiments. It is clear that the simulation results of the model of Vendelin et al. differed very much from the experimental values. We then

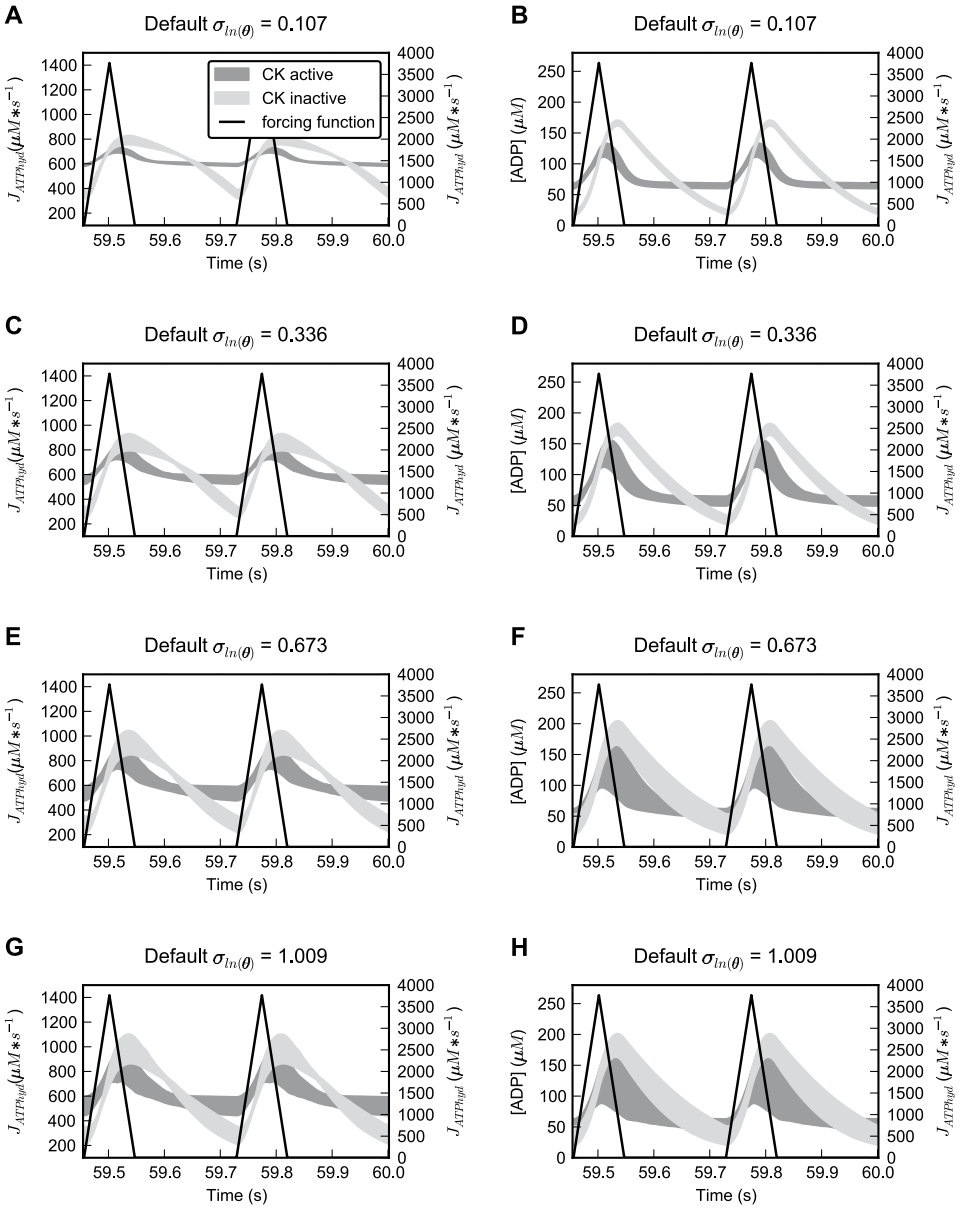

Figure 3.10: Ensemble predictions of oscillations of ADP levels and oxidative phosphorylation for different default priors. Predictions of ATP synthesis rate and ADP concentrations from ensembles sampled with prior $\sigma_{ln\theta}$ of (A, B) 0.107, (C, D) 0.336, (E, F) 0.673 and (G, H) 1.009. Dark and light grey regions show the 95% confidence interval for ensemble predictions with full and inhibited (2% activity) creatine kinase activity, respectively. The forcing function of ATP hydrolysis is plotted in black. Please note the different scales for left and right y-axes in each panel.

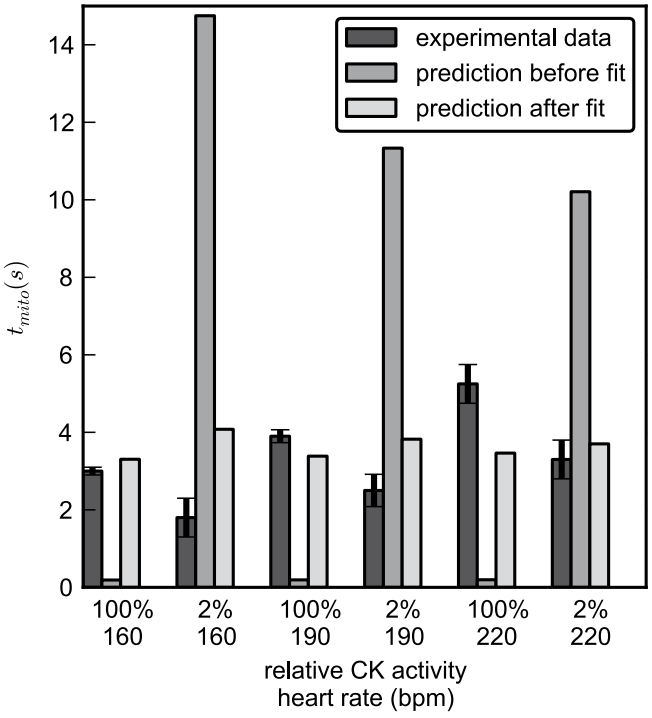

Figure 3.11: Parameter fit for the model of Vendelin et al. to response times of oxidative phosphorylation measured in isolated rabbit hearts. Model parameters were estimated using a modified Levenberg-Marquardt algorithm. Dark grey bars represent the tmito values from the experiment [Harrison et al., 2003], the lighter shaded bars bars represent the tmito values predicted by the model with the original parameters given in [Vendelin et al., 2000] and after parameter adjustment by the fitting procedure. Data is available for six different conditions: three different amplitudes of heart rate jump (from 135 bpm to 160, 190 and 220 bpm heart rate), each one measured with full wild-type CK activity (100%) or with CK activity inhibited to 2% of wild-type value. Note that in this optimisation, all model parameters where allowed to deviate during the fitting procedure.

applied our optimisation strategy and the tmito became 3.3, 3.4 and 3.5 s for the full CK activity for heart rate steps to 160, 190 and 220 beats/min, and 4.1, 3.8 and 3.7 s with CK inhibited by 98%.

Results of the initial simulations and model predictions after parameter fitting are shown in Figure 3.11. The correspondence with the experimental data is still unsatisfactory after parameter adjustment by the optimisation procedure.

3.9 Acknowledgements

We are very grateful to Ryan Gutenkunst for excellent advice on using the SloppyCell modelling environment and to Bernd Brandt for scientific advice and help with the computer cluster. We also thank Jaap Heringa for suggestions and comments on the manuscript.

3.10 Author Contributions

Conceived and designed the experiments: HH JHGMvB. Performed the experiments: HH. Analysed the data: HH JHGMvB. Wrote the paper: HH JHGMvB.

Chapter 4

Simulating the physiology of athletes during endurance sports events: modelling human energy conversion and metabolism

van Beek, J.H.G.M., Supandi, F., Gavai, A.K., de Graaf, A.A., Binsl, T.W. and Hettling, H. (2011) Simulating the physiology of athletes during endurance sports events: modelling human energy conversion and metabolism. *Philosophical Transactions of the Royal Society A*, 369(1954): 4295-4315.

4.1 Abstract

The human physiological system is stressed to its limits during endurance sports competition events. We describe a whole body computational model for energy conversion during bicycle racing. About 23% of the metabolic energy is used for muscle work, the rest is converted to heat. We calculated heat transfer by conduction and blood flow inside the body, and heat transfer from the skin by radiation, convection and sweat evaporation, resulting in temperature changes in 25 body compartments. We simulated a mountain time trial to Alpe d'Huez during the Tour de France. To approach the time realized by Lance Armstrong in 2004, very high oxygen uptake must be sustained by the simulated cyclist. Temperature was predicted to reach 39 °C in the brain, and 39.7 °C in leg muscle. In addition to the macroscopic simulation, we analysed the buffering of bursts of high adenosine triphosphate hydrolysis by creatine kinase during cyclical muscle activity at the biochemical pathway level. To investigate the low oxygen to carbohydrate ratio for the brain, which takes up lactate during exercise, we calculated the flux distribution in cerebral energy metabolism. Computational modelling of the human body, describing heat exchange and energy metabolism, makes simulation of endurance sports events feasible.

4.2 Introduction

The human physiological system is stressed to its limits during endurance sports competition events such as bicycle racing or marathon running. By computational modelling of the body during such challenging events, we can investigate adaptation under extreme stress, which is not only useful to understand the physiology of athletes but also of rescue workers, fire fighters and mountaineers, and others who must perform close to maximal capacity. Understanding the body's response under maximal stress may also help to understand disease processes. For instance, tissue hypoxia may be caused by great exertion rather than by ischaemia or tumour growth. Other examples are dehydration, high body temperatures caused by exertion rather than fever, and high blood lactate levels caused by physical exertion rather than hypoxia or shock [Alders et al., 2011].

Physical exercise affects human physiology at multiple scales. The physical work done by athletes is associated with force exertion, temperature changes in the whole body, sweat excretion and increased uptake of oxygen, water and food, all measurable at the whole body level. At the cellular scale, adenosine triphosphate (ATP) hydrolysis energises the interaction of actin and myosin molecules in the sarcomeres of the muscle cells. The response of the body involves an extensive interplay between various organs. The heart, for instance, starts to pump more blood to transport oxygen and carbon substrates to the muscle and remove carbon dioxide, lactic acid and heat. The brain participates, among others, by regulating lung ventilation, but is challenged with increased lactic acid levels and higher blood temperature, potentially altering

neuronal metabolism. The brain also controls muscle coordination and, importantly, determines the motivation of the athlete to perform maximally despite high brain temperatures, acidification, fatigue and sometimes even pain [St Clair Gibson and Noakes, 2004]. To prepare for the effort, the gut has to digest food to provide carbon substrates to the muscle cells via the blood stream. The whole body response is coordinated via hormones and the nervous system. In conclusion, modelling the human physiological system at macroscopic and microscopic scales during maximal athletic performance is an excellent test to demonstrate the possibilities of computational modelling for the physiome/virtual physiological human. The multi-scale aspects of modelling human physiology arise very naturally in this context.

Here, we will implement a whole body model for energy conversion and heat transport. The latter is very relevant because almost 80% of the metabolic energy during exercise is converted to heat. We will couple this whole body model to equations for the motion and external power requirements during cycling. At the sub-cellular scale, we will look at models for intracellular events, which have a much higher biochemical and/or time resolution than the whole body model. We discuss the supply of metabolic fuel for exercise by nutrition. In addition, we introduce a model of metabolic flux distribution to predict how brain metabolism may deal with higher blood lactate levels. Metabolic events in the muscle cells are analysed at high time resolution with a model for the buffering of bursts of high ATP hydrolysis caused by the rhythmic muscle contractions during bicycle racing.

To demonstrate the applicability of whole body modelling to world class athletic performance, we simulate a mountain time trial in the Tour de France from Bourg d'Oisans to Alpe d'Huez. As the basis for our virtual cyclist, we use published physiological data [Coyle, 2005] for Lance Armstrong, sevenfold Tour de France champion[1].

4.3 Implementing a whole body model of human energy conversion and heat transport

We will start with the description of a whole body computational model that describes the physiological system of an athlete performing external work at the macroscopic level. Although the model can be applied to various endurance sports, we will apply it to bicycle racing here.

The energy conversion relevant to propel the bicycle takes place in muscle. During cycling, less than a quarter of the energy obtained by burning nutrients with oxygen is converted to external work. The remainder of the energy is converted to heat, which is transported from the muscle by thermal conduction and circulatory convection. The blood also delivers oxygen and carbon substrates to the muscle and carries away

[1] Armstrong's seven Tour de France titles were revoked in August 2012, nine months after the publication of this article.

the breakdown products, such as carbon dioxide and lactic acid, in addition to heat. The heat is distributed through the body by the blood and raises the temperature in all organs, including the brain and skin. The heat transported to the skin can be dissipated for a major part by evaporation of sweat. Which muscle groups are active depends on the type of exercise. Because of the specific localisation of muscle activity and the extensive interplay between all organs, it is highly desirable that the whole body model contains multiple segments to represent distinct muscle groups and organs.

The whole body model can be viewed conveniently to consist of a controlled system, which contains metabolism and heat transport in blood and tissue, and a controlling system that regulates the response of the controlled system by neural and hormonal signals. Heart rate, cardiac output and lung ventilation increase under control of nerve signals from the brain. Sweat excretion and skin vasodilation are, for an important part, under central neural control. Therefore, the controller part of the model represents the action of the nervous and hormonal system on the heat transport in the controlled system, which consists of heat capacities, thermal conductance, tissue perfusion, etc.

A macroscopic model of human whole body physiology was described by Stolwijk & Hardy [Stolwijk and Hardy, 1966a] and originally implemented on an analogue computer. It provides a set of equations that fulfil the requirements outlined above. The model for the analogue computer was modified and implemented on a digital computer by Stolwijk for the National Aeronautics and Space Administration [Stolwijk, 1971]. The equations of the latter computational model were incorporated in the model presented here.

4.3.1 Representing compartments in the body: the controlled system

The controlled system of the body is divided into six segments in the model. The central segment of the body (the trunk) consists of the thorax and abdomen. Other segments represent the head, arms, hands, legs and feet (see Figure 4.1).

Each of these segments is divided into four nodes: skin, (subcutaneous) fat, muscle and a core node. The central blood volume in the large blood vessels and cardiac cavities forms the 25^{th} node. The cardiac output representing the total blood flow is partitioned to all the other 24 nodes and exchanges heat with the tissue. To each node, a certain mass, heat capacity and basal metabolic rate are assigned. For the head, for instance, the muscle mass is relatively small (about 0.4 kg), while in the legs, the muscle mass is large (about 10 kg). The core node of the head segment is relatively large (3 kg) because it represents the brain and the skull. The muscle nodes can contribute to external work by energy conversion from nutrients and oxygen. Below, we provide an outline of the model. Detailed equations for the model are given in the

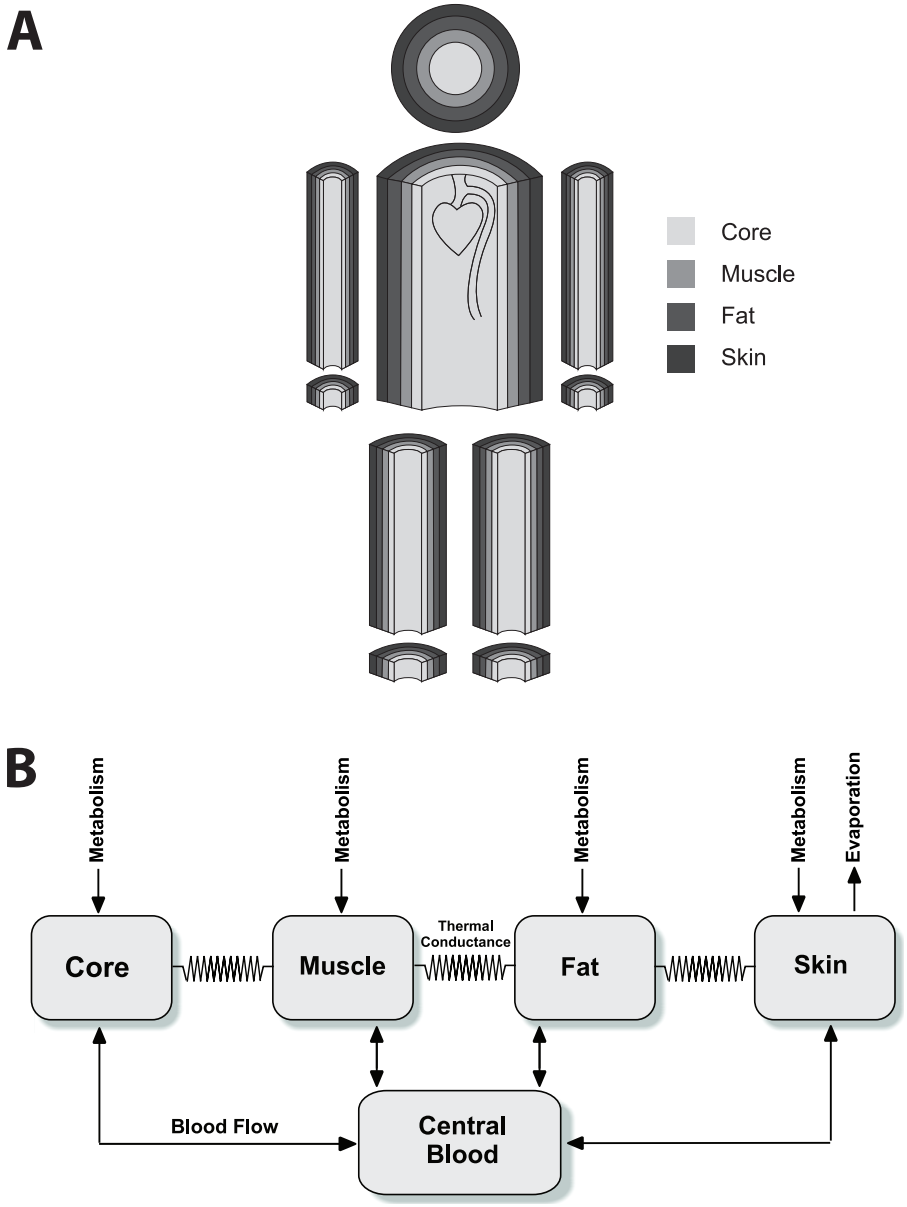

Figure 4.1: Scheme of the whole body model of heat transport. (a) The body is divided into six segments: head, trunk, arms, hands, legs and feet. Each segment consists of four layers approximated by concentric cylinders, except the head, which is a sphere. (b) Adapted from Stolwijk [Stolwijk, 1971]. Heat is generated by metabolism, exchanged with the blood perfusing the layer and conducted to adjacent layers. Heat is dissipated from the skin by evaporation, but also by convection and radiation.

online supplementary material[2].

Metabolism leads to substantial heat generation in the muscle nodes, which is calculated from the total metabolic rate by subtracting the contribution made to external work. The generated heat divided by the heat capacity of each node gives the potential rate of increase of the local temperature. Each node receives a fraction of the total cardiac output, and heat generated in the tissue is exchanged with this blood flow. Heat is also transported between the adjacent nodes by heat conduction through tissue. From the surface of the body inward, layers of skin, fat and muscle are encountered until the core of the segment is reached. This is represented by the four nodes per segment. The simple equation for conductive heat transfer TD from node N to node $N+1$ is

$$TD[N] = TC(T_N - T_{N+1}) \tag{4.1}$$

where TC is the thermal conductance, which is multiplied by the temperature difference between the nodes. For the skin nodes, this equation is replaced by an equation that takes radiative, convective and evaporative heat loss to the environment, including the effects of humidity and ambient temperature, into account (see the online supplementary material). The convective heat transport $BC[N]$ from a node by blood flow BF is given by

$$BC[N] = BF[N] * C_{blood} * (T_N - T_{centralblood}) \tag{4.2}$$

with C_{blood} the heat capacity of blood. This equation is simple because the temperature in blood flowing through small arteries and arterioles becomes almost equilibrated with tissue, even before it reaches the capillaries [van Beek, 1996].

The rate of change, dT/dt in temperature of node N is given by

$$C_{node} * \frac{dT_N}{dt} = Q[N] - TD[N] + TD[N-1] - BC[N] \tag{4.3}$$

The metabolic heat generation $Q[N]$ consists of a basal heat production rate $QB[N]$ plus a potentially large additional heat production in muscle during exercise. C_{node} is the heat capacity of the node. Even if a node has low $Q[N]$, its temperature may rise considerably because of heat conduction from adjacent nodes and heat transfer from the blood. Heat may, for instance, be transferred from leg muscles to the brain and skin via the blood.

A large quantity of heat is transported to the environment via the skin by radiation, convection and evaporation of water. Water is excreted from the sweat glands under central and local control. The skin also loses heat by convection, which is very much increased by high air velocity. This velocity is not only caused by the wind,

[2] http://rsta.royalsocietypublishing.org/content/369/1954/4295/suppl/DC1

but also by the motion of the cyclist. A small fraction of heat is given off to air in the airways and lungs by conduction and evaporation. Detailed equations are given in the online supplementary material.

4.3.2 Sweating, shivering and increasing cardiac output: the controlling system

The body senses the local temperature in various locations, among others in the brain and in the skin. Signals from the sensors are sent to the brain. If temperature is increased, the central controller sends signals to cause skin vasodilation and sweat production. If body temperature falls, signals are generated to produce heat by shivering and to constrict the blood vessels that control skin blood flow. Relatively little is known about the nervous mechanisms that sense tissue temperature and which form the control system in the brain, generating efferent signals to effector mechanisms that control heat generation and transport in the body [Werner, 2010]. Peripheral nerve signals have been measured that are responsive to temperature, and the anterior hypothalamus region of the brain is sensitive to temperature changes. However, there is insufficient information about the nervous regulation of temperature to model the controller in mechanistic detail.

Therefore, in the computational model, a simple black box approach was chosen: the deviation of node temperatures from a reference temperature constitutes the input to the central controlling mechanism [Stolwijk, 1971]. In particular, the central controller uses the temperature of the brain and of the skin nodes as input. Under thermoneutral conditions, the error signal is assumed to be zero, and no efferent signals are generated to stimulate sweating, skin vasoconstriction, vasodilation or shivering. It is unlikely that the temperature is explicitly compared with a set point in the real body [Werner, 2010], but the simple black box formulation adequately describes thermoregulatory responses.

The equations describing the controller were parametrised based on extensive measurements of the relation between input signals (brain and skin temperatures) and output signals (rate of sweat evaporation, blood flow to the skin) [Stolwijk and Hardy, 1966a, Stolwijk, 1971]. Therefore, the simple equations describe the reaction of the human body to temperature change adequately, based on a solid experimental basis. The equations and parameter values for the control of vasodilation, vasoconstriction in the skin, sweating and shivering are given in the online supplementary material.

4.3.3 Representing voluntary control of muscle activation

For an athlete who tries to win a race, not only the physical work capacity counts, but also the motivation to go to the limits of performance despite exhaustion and perhaps even pain. The key parameter for the physical capacity is the maximal oxygen uptake

of the athlete that can be measured in the laboratory. However, during a long effort, oxygen uptake cannot be sustained at the same level as during a relatively short test in the laboratory. The available mechanical power tends to show a downward trend during a long race [St Clair Gibson and Noakes, 2004]. Because the complex neural processes that determine the sensation of fatigue and exhaustion and the voluntary control of muscle contraction during exertion are insufficiently known, in the present model, these mental processes are enormously simplified by a 'brain push' factor, which gives the fraction of maximal oxygen uptake that will be used by the athlete.

4.3.4 Integrating the model

The ordinary differential equations were implemented and integrated in R computing language using the lsoda routine, which is based on algorithms by Hindmarsh [Hindmarsh, 1983] and Petzold [Petzold, 1983]. R is a high-level computer language and computational environment, which is open source and provides a large collection of open-source analysis tools.

4.4 The power required for bicycle racing

The external mechanical power available to increase and maintain the speed of the cyclist-bicycle system is delivered by the muscle nodes in the computational model of the cyclist. Physical power equals the force propelling the bicycle times the velocity. The opposing forces that must be overcome to propel the bicycle are (i) the frictional force of the air (wind drag), (ii) the frictional force of the tyres on the road (rolling resistance), (iii) the component of the force of gravity, which is parallel to the road surface when climbing, and (iv) the force used to accelerate or decelerate. The external power delivered by the cyclist's muscles is balanced by the power of the opposing forces. A voluntary decision to exert more force is associated with more external power and may lead to acceleration.

Equations describing the components of the resistance to motion have been derived and their coefficients measured. The rolling resistance (R_r) is a force (in N) given by

$$R_r = C_{R_r} cos[arctan(S)](M + M_b) \tag{4.4}$$

where S is the gradient of the road (i.e. vertical distance/horizontal distance), M is the mass of the rider and M_b is the mass of the bicycle [Olds et al., 1993]. This expression gives the normal force exerted by the combined mass multiplied by a constant C_{R_r} that depends on wheel radius and road surface and is approximately inversely dependent on tyre pressure. The friction of the chain and other moving parts of the bicycle are usually negligible. The measured values of the parameters in the power equations and further equations are given in the online supplementary material.

The air resistance given as a force in N is

$$R_a = k(v_{ss} + v_w)^2 \tag{4.5}$$

where v_{ss} is the speed of the bike relative to the road and v_w is the speed of the head wind relative to the road (negative for a tail wind) [Olds et al., 1993]. The coefficient k is proportional to the frontal projected area and air density. This dependency is given quantitatively in the online supplementary material.

If the road is sloping, the component of the weight of rider plus bicycle parallel to the road, F_g, is needed to account for gravity. An additional force F_{acc} can be used to accelerate. The instantaneous power needed to propel the cyclist at a certain moment in the race equals the instantaneous velocity multiplied by $(R_r + R_a + F_g + F_{acc})$. The power to propel the bicycle is balanced in the model by the power of muscle work made available from metabolic processes. This power is determined not only by the cyclist's physical capacity, but also by the motivation to win and the pacing strategy that is consciously followed (e.g. starting as fast as possible versus somewhat slower to save energy for later). The motivational factors are represented by the brain push factor mentioned above. The balance of the power required and delivered by the cyclist determines whether he accelerates, decelerates or maintains a steady speed. The acceleration is integrated to obtain the speed of the cyclist, which determines the road distance covered. The equations in the computational model are given in the online supplementary material.

4.5 Simulating a time trial to Alpe d'Huez

Here, we apply the model to simulate a mountain time trial in the Tour de France. The time trial starts in the town of Bourg d'Oisans at an altitude of about 720 m above sea level, and ends after 15.5 km in the town of Alpe d'Huez at an altitude of 1850 m. The conditions are similar to the mountain time trial in the Tour de France on 21 July 2004, although we emphasise that exact correspondence to the parameters of the time trial and the cyclist could not be established. The altitude of the road at 1 km intervals was obtained from published profiles. From these altitudes, we calculated the slope of the road with a 1 km resolution. The climb starts 1.5 km from the start and the slopes vary between 8 and 11% for three quarters of the distance covered. For the anthropometric and physiological characteristics of the rider in the model, we use published data for Lance Armstrong [Coyle, 2005], who won the stage with a time of about 39 min 41 s, more than 1 min faster than Jan Ullrich, who finished second.

In the simulation, the body mass of the rider was estimated to be 72 kg during the race, and his maximal oxygen uptake determined in the laboratory is 6.1 l O_2 per minute, which is the highest value recorded during repeated laboratory tests for this individual [Coyle, 2005]. An important parameter determined in the laboratory is the Δefficiency, which is defined as the increase in work accomplished on the bicycle

ergometer per unit time divided by the energy expended during the same time. The energy expended was calculated from the measured oxygen uptake and reflects the energy available from the oxidation of metabolic fuel, mainly carbohydrate and fat, with oxygen. The work accomplished per unit time is measured by determining the mechanical power delivered by the cyclist to the bicycle ergometer. The increase in oxygen consumption from the resting level the cyclist manages to accomplish is multiplied by the Δefficiency, 0.231 in this case, and by the total energy expended per amount of oxygen to calculate the energy available for propelling the cycle-cyclist system.

For the hypothetical case that the cyclist would be able to use 100% of his maximal oxygen uptake, as measured during a short bout of intense exercise in the laboratory at lower altitude, the simulation predicts that the time to complete the 15.5 km up to Alpe d'Huez is 38 min 31 s. Here, it was assumed that the decrease in barometric pressure during the climb to 1850 m leads to reduced air resistance because of lower air density. For two important reasons, it is not realistic to assume that 100% of maximal oxygen uptake can be sustained. Firstly, it is generally found that cyclists can sustain less than 100% oxygen uptake if the effort lasts longer than 10 min [Olds et al., 1993]. The second reason is that the progressively lower oxygen partial pressure in the air at higher altitude will lead to lower maximal oxygen uptake. The relative reduction of maximal oxygen uptake with a lower concentration of oxygen in the air exists in untrained subjects, but is even stronger in trained athletes [Ferretti et al., 1997], where an average decrease of about 13 % in oxygen uptake is found if oxygen partial pressure in the air falls by 24%. At an altitude of 1850 m, oxygen partial pressure is expected to be decreased by about 20% relative to sea level. If the simulation includes this effect of hypoxia on oxygen uptake, the predicted time becomes 41 min 22 s.

The actual time of Lance Armstrong in 2004 was about 39 min 41 s. This time is found in the simulation of the virtual cyclist if 96.5% of the maximal uptake of 6.1 l O_2 per minute can be sustained during the climb to Alpe d'Huez, assuming no further reduction owing to altitude hypoxia. A sustained fraction of 0.965 during about 40 min is very high relative to trained (but non-professional) cyclists [Olds et al., 1993]. To obtain the time realized by Lance Armstrong during the mountain time trial, it is necessary to assume an oxygen uptake of about 5.9 $l * min^{-1}$ for the virtual cyclist, despite the reduced oxygen partial pressure encountered when climbing from 720 to 1850 m. Given that professional cyclists often aim for peak performance at the time of decisive races, it is possible that a higher value of maximal oxygen uptake (VO_2max) was valid during the Tour de France in comparison with the value of 6.1 l O_2 per minute reported for the laboratory tests, which were carried out in earlier years.

Assuming a very high will power of the simulated cyclist such that he is able to maintain 95% of the maximal O_2 uptake and assuming that the reduction of oxygen uptake by altitude hypoxia is as measured by Ferretti et al. [Ferretti et al., 1997] for trained endurance athletes, we estimate that the maximal oxygen uptake measurable at sea level should have been 6.74 l O_2 per minute to complete the time trial in 39 min 41 s. We emphasise that it was not possible to obtain all parameters for the cyc-

list and the time trial with great accuracy, implying that we most probably obtained incomplete correspondence between the virtual cyclist in the model and Lance Armstrong during the 2004 time trial. Our estimates of oxygen uptake must, therefore, be treated with great caution. For instance, the sensitivity of real professional cyclists to altitude hypoxia may differ from that assumed in the model. We assumed that the simulated cyclist would be able to maintain 95% of maximal O_2 uptake, i.e. the brain push factor mentioned above. If this value is actually higher or lower, the estimate of the maximal O_2 uptake corresponding to the performance during the time trial will have to be adapted.

Simulation results for tissue temperatures and power dissipated by air resistance, rolling resistance and gravity for the cyclist with 6.74 $l*min^{-1}$ maximal O2 uptake are given in Figure 4.2. About 77% of metabolic energy was measured to be converted to heat in this Tour de France champion [Coyle, 2005], which leads to significant warming of leg muscles and brain to over 39 °C.

The cyclist who finished second in 2004 was reported to be 5 cm taller than Lance Armstrong. If the body height of the virtual cyclist is increased from 179 cm to 184 cm, the model simulation predicts that the time needed for the time trial becomes about 3 s longer. If the total weight of the rider and bicycle is 2.4 kg larger, with muscle mass and oxygen uptake constant, the predicted time becomes 40 min 42 s, which is about the time of Jan Ullrich, who finished second in 2004. Larger body weight may, among others, be caused by greater fat or skeletal mass. The slowing effect of the added weight is for the largest part caused by the increase in the component of the gravitational force parallel to the road surface. However, smaller effects are due to the component of the force normal to the road surface increasing the rolling resistance, and to increased air resistance owing to increased size of the body. This illustrates that small differences in body size can have significant impact on athletic performance.

4.6 Validating the whole body model

The parameters of the controlled (e.g. metabolism and heat transport) and controlling systems (neural and hormonal regulation, see detailed explanation above) were measured on human subjects [Stolwijk and Hardy, 1966a, Stolwijk, 1971]. For new subjects with other characteristics than the original study group, the model may reproduce steady-state thermoregulatory responses less accurately. The transient response to sudden changes in thermal stress forms a more stringent test than steady-state responses. To this end, an increase in air temperature from 30 °C to 48 °C for 2 h was simulated. The temperature of the walls was equilibrated with the air temperature during the experiment. The response of average skin temperature was measured to increase from 34 °C to 36.5 °C [Stolwijk and Hardy, 1966b], while the model predicted an increase from 33.4 °C to 36.3 °C. Both the present model implementation and the original model of Stolwijk [Stolwijk, 1971] predict a temperature increase in the core

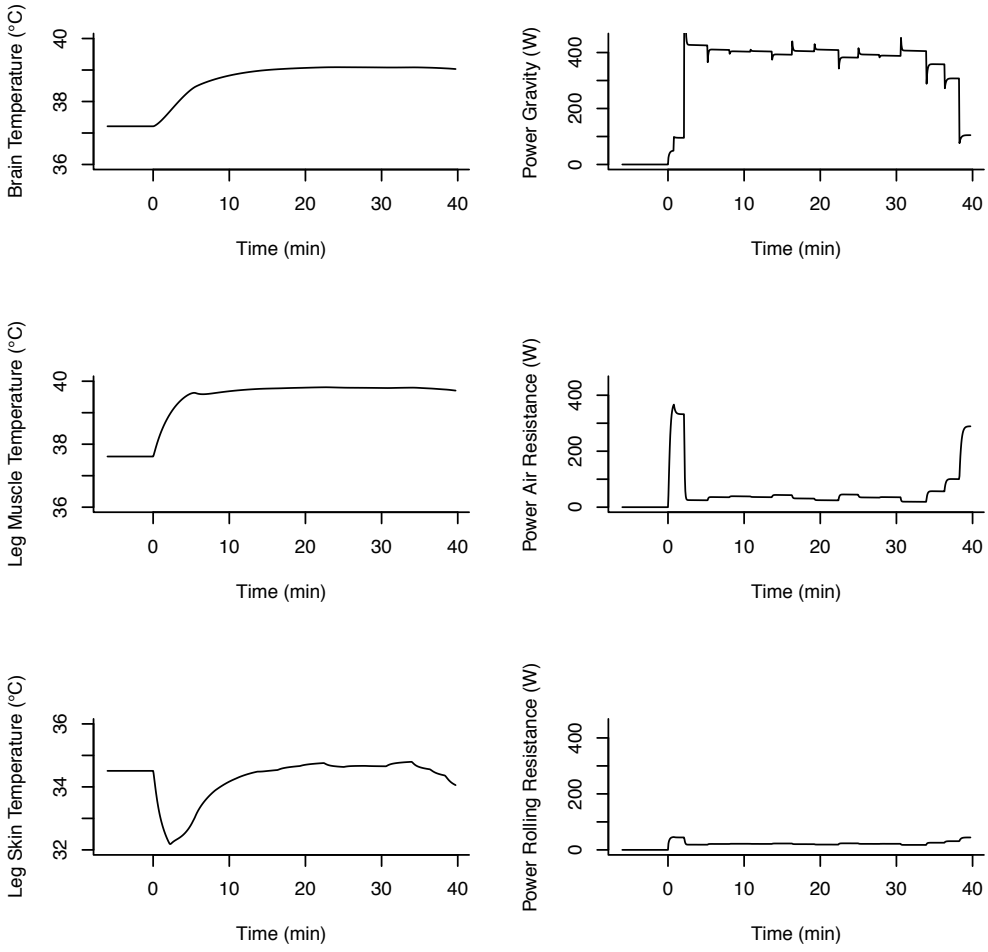

Figure 4.2: Predictions of the main fates of metabolic energy converted in the muscle during the mountain time trial. (a) The increasing temperatures of leg muscles and brain reflect heat generation. The increasing air velocity across the skin during the time trial leads to significant cooling, but this effect becomes less at steeper slopes, where speed is diminished. (b) Power (in Watts) expended against gravity, air resistance and rolling resistance. The sharp spikes in the power against gravity are caused by the sharp transitions to different grades caused by the low resolution (1 km) by which the slope of the road to Alpe d'Huez was known to us: the cyclist enters a section with a steeper grade with high velocity and is subsequently slowed down by the increased opposing gravitational force. The cyclist finishes after 15.5 km in 39 min 41 s.

of the trunk, which was significantly smaller than the increase in rectal temperature measured in a group of young subjects [Stolwijk and Hardy, 1966b].

We further simulated experiments by Nybo et al. [Nybo et al., 2002], where subjects started to exercise at 174 W on a bicycle ergometer at 20 °C ambient temperature. In the simulation, the brain temperature increased from 37.28 °C to 37.96 °C after 45 min. In the experiments, these numbers were 37.2 °C and 38.0 °C for the jugular venous temperature at the base of the cranium, which is close to the brain temperature. The time course of brain temperature was quite similar in the simulation and Nybo's experimental study.

The model equations of Stolwijk & Hardy [Stolwijk and Hardy, 1966a] are quite extensive. Many measured trends and variables are predicted remarkably well, although significant deviations from the measurements are also found [Stolwijk, 1971, Stolwijk and Hardy, 1966b]. Indeed, Stolwijk discussed a number of shortcomings of the model [Stolwijk, 1971]. Although the model in most cases predicts the right order of the response, it is desirable to further validate and improve this extensive model. Parameters may be fine tuned to fit special groups of subjects, for instance, trained professional athletes. The model parameters are representative of a group of human subjects studied in the laboratory, but were not yet adjusted for any individual, in this case, a world class cyclist. The parameters can, for instance, be further fine tuned by imaging the body of the modelled individual using magnetic resonance imaging and obtaining the VO_2max closer in time to the real time trial. The frictional resistance of the bicycle and air resistance in a wind tunnel could also be measured. It would then be possible to test the model by measuring the speed and delivered power of the cyclist and the slope of the road at greater resolution than has been possible here. Heart rate, skin and rectal temperatures could be measured non-invasively during a time trial. By weighing the body as well as fluid and food intake, the net evaporation during the time trial can be assessed.

What is still hard to quantify is the athlete's determination to voluntarily go as fast as possible. The effort made by an athlete will vary during a race: often, it is very high at the start of a time trial, falling quickly and sometimes showing periodic behaviour [St Clair Gibson and Noakes, 2004, Jeukendrup et al., 2000]. In short, although the model shows responses that are in line with existing experimental measurements for groups of subjects, the model parameters may be adjusted to better approach individual characteristics and may be tested in the field on individuals during a time trial out of competition.

4.7 Time-resolved model of events at the muscle cell level

In the whole body model discussed so far, metabolic processes inside muscle cells were not modelled in detail. In this section, we focus on how energy metabolism at the

cellular level can be described and predicted by means of mechanistic modelling with much higher time resolution than feasible for the whole body model. Endurance exercise requires permanent energy supply to enable contractile work in skeletal muscle. However, the amount of ATP in muscle cells is merely sufficient to energise two or three maximal contractions in skeletal muscle under normal conditions [Lieber, 2002]. Therefore, aerobic and anaerobic ATP production must be tightly coupled with the energetic demands of the muscle cell. Creatine kinase (CK) enzymes provide a buffer between energy producing and consuming pathways. CK facilitates additional energy storage in the form of phosphocreatine (PCr) by catalysing the reversible transfer of a phosphate group to creatine (Cr),

$$ATP + Cr \overset{CK}{\longleftrightarrow} ADP + PCr \tag{4.6}$$

where ADP stands for adenosine diphosphate. Previously, long ago, PCr was suggested to maintain intracellular ATP homeostasis during muscle contraction [Bergström, 1967]. The discovery of two distinct isoforms of CK, Mi-CK in mitochondria and MM-CK in myofibrils, led to the hypothesis of a role of CK as an energy shuttle from ATP producing to ATP consuming sites across the mitochondrial outer membrane (MOM) [Bessman and Geiger, 1981]. In previous work, we investigated the functions of CK using a mathematical model of energy metabolism in a cardiac myocyte integrated with data from various scales [van Beek, 2007, van Beek et al., 2008, Hettling and van Beek, 2011]. Our model predictions emphasised the role of CK in damping the rate of oxidative phosphorylation and ATP and ADP levels during high-amplitude bursts of ATP consumption during the cardiac cycle. In this section, we show that our model can describe experimental data from human skeletal muscle collected during cycling. As a result, our model predicts buffering of dynamic ATP synthesis and metabolite levels.

The model consists of four major components in two compartments that capture the key elements of the CK system (Figure 4.3a): (i) mitochondrial ATP production by oxidative phosphorylation (OxPhos), (ii) high-energy phosphate transfer by Mi-CK in the mitochondrial intermembrane space, (iii) the MM-CK reaction in close vicinity to myofibrils, and (iv) ATPase activity, represented by a forcing function of ATP hydrolysis simulating pulsatile energy consumption of the contracting muscle.

Parameter values for Mi-CK reaction, kinetics of ATP production and MOM permeability to metabolites were collected from van Beek [van Beek, 2007]. Kinetic constants for skeletal muscle MM-CK were reported in Vicini & Kushmerick [Vicini and Kushmerick, 2000] based on earlier measurements [McFarland et al., 1994]. To fine tune all 22 kinetic model parameters, we use data from an experiment in which 10 well-trained subjects performed an incremental challenge on a bicycle ergometer at submaximal workloads of 40 and 75% of their maximal aerobic capacity (VO_2max). After each workload challenge, a muscle biopsy from the quadriceps femoris was taken and analysed for intracellular concentrations of ATP, PCr and Cr (Figure 4.3b) among others [Sahlin et al., 1987]. The $[H^+]$ concentration reported for this study ranged

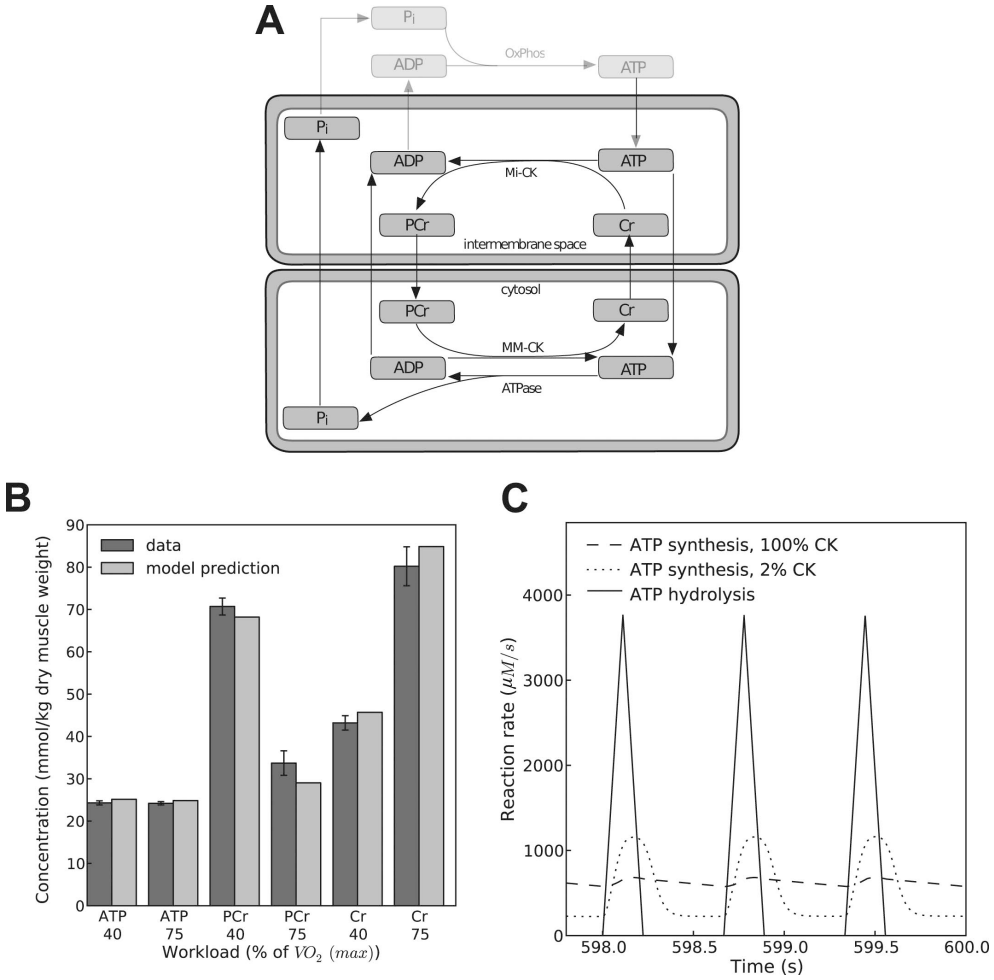

Figure 4.3: (a) Scheme of the model of the creatine kinase (CK) system. Cell and mitochondrial membranes are indicated. Cr, creatine; PCr, phosphocreatine; Pi, inorganic phosphate; OxPhos, oxidative phosphorylation. (b) Steady-state metabolite concentrations at the end of a 600 s simulation compared with data from cycling exercise experiments at different submaximal workloads [Sahlin et al., 1987]. For conversion of metabolite concentrations between model ($\mu mol * l^{-1}$ cell water) and data ($mmol * kg^{-1}$ dry weight), we assumed an intracellular water content of $3\,l * kg^{-1}$ dry muscle weight [Sjogaard et al., 1985]. (c) Model prediction of the ATP synthesis time course at 100% (dashed line) and 2% (dotted line) of normal CK activity at a workload of 40% of VO2max. The forcing function of pulsatile ATP hydrolysis is plotted as a solid line. Note that the last 2 s of a simulation over 600 s are shown ensuring a steady state. (b) Dark grey, data; light grey, model prediction. (c) Dashed line, ATP synthesis, 100% CK; dotted line, ATP synthesis, 2% CK; solid line, ATP hydrolysis

from $0.89 * 10^7 M$ at rest to $1.20 * 10^7 M$ at 75% of $VO_2(max)$ [Sahlin et al., 1987], and was taken into account in the calculation of the CK reaction [Lawson and Veech, 1979]. Maximal aerobic work rates were excluded from this analysis because anaerobic metabolism starts to play a major role and is not yet included in the model.

To estimate the rate of ATP hydrolysis powering contractile work in leg muscle, we use the average VO_2max of 4.1 $l * min^1$ of the subjects in the bicycling exercise [Sahlin et al., 1987], measurements of aerobic capacity in skeletal muscle [Gnaiger, 2009], and assume 10 kg active leg muscle mass for well-trained athletes [Tipton, 2005]. Further, we assume a P/O ratio of 6, meaning that six molecules of ATP are produced per molecule of oxygen consumed [Tipton, 2005]. From this, we estimate an average ATP hydrolysis flux, $J_{ATP,hyd}$, of 628 and 1175 $M * s^1$ for 40 and 75% of VO_2max, respectively. Parameter estimation was done with the downhill simplex algorithm implemented in the SloppyCell modelling environment [Gutenkunst et al., 2007a]. Because the kinetic parameters were taken from solid experimental data [McFarland et al., 1994], we constrained these parameters based on their reported measurement error using Bayesian prior terms added to a least-squares cost function. Parameters for which no measurement error was reported in the literature were constrained by a default prior term as in Hettling & van Beek [Hettling and van Beek, 2011]. Taking this a priori information into account yields more robust parameter values [Gutenkunst et al., 2007b]. Results of the parameter fit to the metabolite concentrations for two experimental conditions are shown in Figure 4.3b. The model predicts a decline of PCr, the high-energy phosphate store, with increasing workload, while ATP levels stay relatively constant. The model predictions for PCr, Cr and ATP are relatively close to the measured values, despite the fact that the kinetic parameters in the model were kept close to their reported literature values by constraining them with their measurement error using Bayesian priors.

A limitation of the present model analysis is the absence of glycolytic ATP formation. Tissue lactate levels were reported to be substantially elevated going from rest to 75 per cent of $VO_2(max)$ [Sahlin et al., 1987] at the time point of the phosphate metabolite measurements (Figure 4.3b). Because lactate production may change at later stages of exercise, the incorporation of dynamic changes in lactate production and intracellular pH in future modelling work is desirable.

For many professional cyclists, mechanical efficiency is maximised at pedal cadences around 90 r.p.m. [Joyner and Coyle, 2008]. In order to simulate alternating energy usage of the leg muscle at this frequency, we used a pulsatile forcing function of ATP hydrolysis. The shape of the ATP hydrolysis pulse is triangular (Figure 4.3c), as used previously for a beating heart [van Beek, 2007]. The leg muscle is assumed active during one-third of the pedal stroke cycle. This time course and the fractional value of one-third is close to predictions based on biomechanical calculations and to electromyographic measurements of the activity pattern in the quadriceps muscle during one pedal stroke [Redfield and Hull, 1986]. Given this triangular activation curve during one-third of the total cycle, the peak value of the ATP hydrolysis rate is sixfold the average ATP rate, which in turn is estimated from oxygen consumption

data.

To investigate the effectiveness of CK in buffering these pulses in ATP demand, we simulated conditions of normal and impaired enzyme activity (Figure 4.3c). With normally active CK, the amplitude of ATP synthesis is almost fully buffered. Blocking CK by 98% results in a ninefold increase of the amplitude of the ATP synthesis oscillations from 106 to 938 $\mu M * s^{-1}$. These results demonstrate the importance of the CK system in maintaining cellular ATP homeostasis and buffering of high amplitudes in mitochondrial ATP production in skeletal muscle during cycling.

4.8 Detailed models of carbon metabolism in the cell

Energy for muscle contraction is derived from the oxidation of carbohydrates and fatty acids derived from the food. The energy is used to synthesise ATP. The pathways for the metabolism of these carbon substrates may be analysed with computational models, which represent the biochemical reaction networks consisting of enzymes and of metabolites. A certain metabolite can be produced by one particular type of enzyme and be used by the next enzyme in the metabolic network. The enzyme content determines the maximal rate at which a reaction can take place. The precise distribution of the flow of carbon in the intracellular metabolic pathways can be predicted by models of interconnected metabolic reactions. The calculation of the flux distribution in the network resembles calculations based on Kirchhoff's law in electrical circuits: the amount of material flowing into a node in the network must equal the amount of material flowing out. For instance, if reaction 1 produces a certain metabolite X at flux V_1, while reactions 2 and 3 consume one and two molecules, respectively, of X at fluxes V_2 and V_3, the balance equation for metabolite X is

$$V_1 - V_2 - 2V_3 = 0 \tag{4.7}$$

Additional constraints on the flux distribution can be derived from biochemical knowledge. The enzyme cytochrome oxidase that transfers electrons to oxygen operates, for instance, in an irreversible way and will never produce oxygen. The first enzymatic reaction in the glycolytic pathway after the uptake of glucose in the cell is an irreversible reaction catalysed by the enzyme hexokinase,

$$glucose + ATP \rightarrow ADP + glucose\text{–}6\text{–}phosphate \tag{4.8}$$

Some, but not all, reactions in the biochemical networks have such an irreversibility constraint. Measurements on input and output fluxes from cells or measured capacities of enzymes may form additional constraints. In this way, flux balance analysis (FBA) [Duarte et al., 2007] derives the possible distribution of metabolic fluxes

in the network. If some of the most relevant central pathways of metabolism are included (e.g. the glycolytic pathway, the pentose phosphate pathway, the tricarboxylic acid cycle and transamination of carbon skeletons to amino acids), a model of flux distribution often contains more than 50 enzymatic reactions.

4.8.1 A computer package to calculate flux distributions

Most of the relevant reactions in the cell can be found in databases of metabolic pathways. In order to facilitate the building of metabolic pathway models, we developed the computer package BiGGR, which, among others, downloads the reactions from the Biochemical, Genetic and Genomic (BiGG) knowledgebase of metabolic reconstructions, which contains several thousand manually curated enzymatic reactions for the human system [Duarte et al., 2007]. The name BiGGR is derived from the BiGG database [Duarte et al., 2007, Schellenberger et al., 2010], while R stands for the R programming language. The R and Bioconductor framework are widely used among the genomics community [Gentleman et al., 2004]. In BiGGR, the user can choose the reactions for the model analysis. The computer program then builds the stoichiometric matrices from the chosen set of reactions describing all balance equations such as equation 4.7. The first matrix contains the set of linear equations, where metabolites are represented by the rows and reactions are represented by the columns. Linear inequalities that represent the constraints are represented by a second matrix. The systems of equations for metabolic systems are usually underdetermined [Duarte et al., 2007]. Despite this situation, it is often possible to derive the range of values that are feasible for the metabolic reactions. Sampling the flux space, for instance, by simple uniform sampling, makes it possible to explore which combinations of fluxes are compatible with the constraints. Alternatively, one can sample a probability distribution in flux space by a Markov chain Monte Carlo procedure: the probability that a point in flux space is included in the sample is proportional to the probability that this point represents measured data. If one can design a plausible linear combination of fluxes, which forms a cost function to be minimised (e.g. the total production of metabolic waste products) or forms a gain function to be maximised (e.g. the production of ATP), it is often possible to determine a unique flux distribution although the system of linear equations is underdetermined.

BiGGR facilitates FBA, building on a range of linear optimisation routines, which were already available for R [Soetaert et al., 2009]. In addition, BiGGR also attempts to automatically visualise the metabolic pathways using a 'hypergraph' format [Murrell, 2013], plotting the numerical results of a flux analysis on the pathway graph. More details about the BiGGR package can be found in chapter 6.

4.8.2 Simulating brain metabolism during exercise

Glucose is the main substrate for fuelling cerebral metabolism under normal physiological conditions when blood lactate concentration remains low. The ratio of O_2 up-

take/glucose uptake is close to 6 [Dalsgaard, 2006]. During maximal exercise, glucose uptake increases and the brain may also use arterial lactate released by the muscles [Dalsgaard et al., 2004]. However, oxygen uptake does not increase proportional to carbohydrate uptake, indicated by a cerebral metabolic ratio (O_2 uptake/(glucose+1/2 lactate) uptake), which falls below 3. The fate of carbohydrate in the brain during exercise remains unknown [Rasmussen et al., 2010]. With nuclear magnetic resonance spectroscopy, no glucose or lactate accumulation is found and the capacity to synthesise glycogen may be too low to account for non-oxidised glucose and lactate [Dalsgaard et al., 2004, Quistorff et al., 2008]. Immediately after maximal exercise, a small increase of free fatty acid release is detected, while no change was measured for other substrates including glutamine, glutamate and alanine [Dalsgaard et al., 2002].

To investigate whether excess carbohydrate uptake can be explained by fatty acid synthesis, we developed a computational model that represents pathways of central metabolism in the brain and simulated the fluxes of glucose and lactate metabolism during rest and maximal exercise. Given the high lipid content of the brain, this is a plausible hypothesis. The reactions in our model comprise central energy metabolism and fatty acid synthesis in their appropriate compartments (cytosol and mitochondria), and metabolite transport across the mitochondrial membranes. The inclusion of the malate-aspartate and glycerol-3-phosphate shuttles allows for transport of reducing equivalents between the cytosol and mitochondria. Amino acid synthesis of alanine and glutamate/gamma-amino butyric acid (GABA) cycling were included. Reactions are taken from the BiGG database, and measurements of substrate uptake in the brain during rest and exercise were taken from Quistorff et al. [Quistorff et al., 2008]. We use the linear inverse modelling routines [Soetaert et al., 2009] linked to the package BiGGR for the calculations. The system is assumed to be in a steady state, with a cost function to maximise ATP production. A list of full reactions and constraints used in this model, input and output fluxes assumed for the cell, as well as the flux distribution resulting from the analysis are given in the online supplementary material.

At rest with no lactate present, metabolism of 0.4 $mmol * min^{-1}$ glucose in the whole human brain is in agreement with the O_2/glucose ratio and yields a maximal mitochondrial ATP synthesis of 9.28 $mmol * min^{-1}$. No flux through the pentose phosphate pathway is predicted (Figure 4.4a). During 15 min of exercise, glucose uptake doubles, and, in addition, 0.53 $mmol * min^{-1}$ lactate and 3.2 $mmol * min^{-1}$ oxygen are taken up. We hypothesise that given the low ratio of oxygen to glucose and lactate uptake, the carbon in these carbohydrates may be used to synthesise and store fatty acids. We calculate that 0.14 $mmol * min^{-1}$ of fatty acid (equivalent to palmitate with 16 carbon atoms) must be synthesised in the brain to balance glucose and lactate uptake with the limited availability of oxygen to oxidise carbon (Figure 4.4b). The pentose phosphate pathway is now activated to produce nicotinamide adenine dinucleotide phosphate, which is required for the synthesis of free fatty acid. The result above suggests that fatty acid synthesis, perhaps followed by incorporation into lipids that are abundantly present in the brain, is a possible pathway for the excess

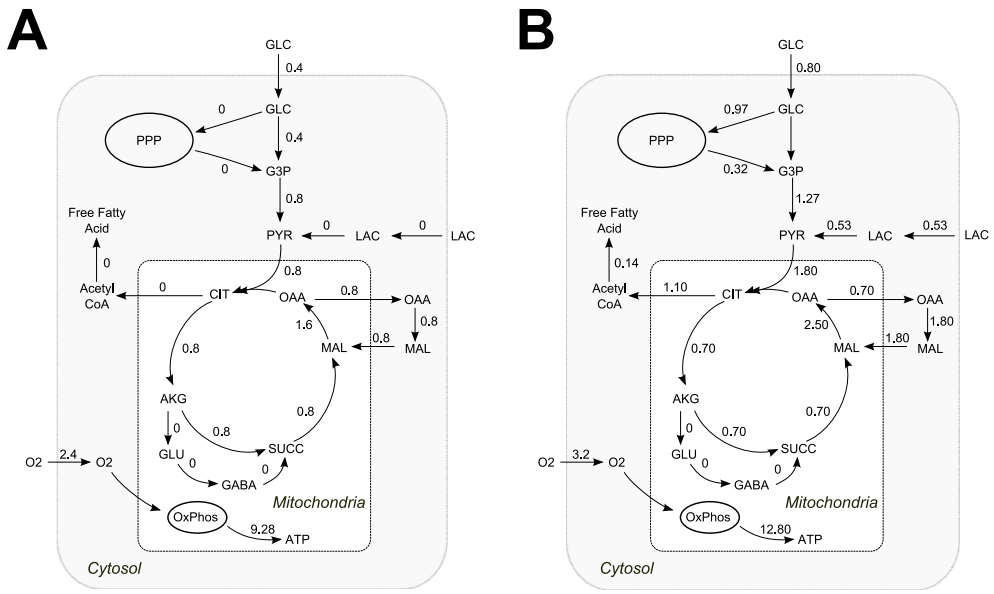

Figure 4.4: Models of carbon metabolism in brain. (a) Simulation of cerebral energy metabolism at rest and (b) during 15 min maximal exercise. Substrate uptake is from Quistorff et al. [Quistorff et al., 2008]. Flux values are in $mmol*min^{-1}$ for the whole brain. PPP, pentose phosphate pathway; OxPhos, oxidative phosphorylation; GLC, glucose; PYR, pyruvate; LAC, lactate; GLU, glutamate; GABA, gamma-amino butyric acid; MAL, malate; OAA, oxaloacetate; SUCC, succinate; AKG, alpha-ketoglutarate; CIT, citrate; CoA, coenzyme A; G3P, glyceraldehyde-3-phosphate. Please note that for clarity not all reactions are shown and the metabolites are sometimes balanced by reactions that are not plotted. There is, for instance, a small backflux towards glucose in the upper part of the glycolytic chain in (b).

carbon. Future experiments to measure storage of cerebral fatty acid, possibly using stable isotope labelling of precursors [van Beek et al., 2009, Binsl et al., 2010a], allow this hypothesis to be tested.

4.9 Future directions

The model approaches discussed address various spatial, temporal and biochemical resolutions and are conceptually complementary. However, they were not yet linked in an integrated computational framework and executed simultaneously. For the buffering of the ATP hydrolysis peaks in the CK model, we calculated the amplitude of ATP hydrolysis off-line. Yet, it seems feasible to couple the CK model to the whole body model. In that way, the energy turnover in the leg muscles would be coupled directly to the amplitude of ATP hydrolysis during cyclical work on the bicycle. The levels of inorganic phosphate and the maximal ADP concentration during the cycle could be used to indicate the fatigue status of the muscle and might be used to

determine the obtainable level of exertion. The minimal concentration of ADP may additionally indicate the risk of oxygen radical formation because of low stimulation of the mitochondria.

To develop computational models of endurance sports performance further, it might be very efficient if a framework existed where the modelling approaches at various scales could be seamlessly integrated. In addition, if one builds and links models of metabolic pathways, it is very efficient if one can integrate biochemical reactions available from databases. A prerequisite for such an exchange of reactions is a uniform standard nomenclature and symbols for metabolites, enzymes and other entities in physiological models. For the present work, we used the nomenclature of the BiGG database [Schellenberger et al., 2010], but it is clear that an ontology which is generally accepted by the biomedical research community or enforced by leading journals is needed.

Given the extensive research literature on sports physiology and the vast knowledge on basic biology, it is timely to start to develop quantitative models of human physiology. Quantitative integration of physiological and biochemical knowledge in computational models may facilitate the understanding of human performance, and may increasingly become a central research tool for integrative physiology. Because professional world class athletes push their performance to the very limits, model analysis of their performance during sports events may provide a unique window on human physiology. By demonstrating the application of virtual physiological modelling to sports or space travel, we may arouse the interest of the general public and young scientists. The benefits for medical applications and enhanced safety surveillance of athletes will no doubt follow.

Our model mainly predicts the maximal VO_2 value achieved during the race. The result suggests that VO_2max explaining the performance during the race is higher than the maximum found in laboratory tests earlier in time. Athletes use all sorts of methods to prepare optimally for the decisive race. Many athletes expose themselves to low oxygen concentrations, for instance, by sojourn at high altitude or by breathing gas mixtures with decreased oxygen content during sleep, to increase natural erythropoietin production. The winner of the 2004 time trial to Alpe d'Huez may have used hypoxic exposure as reported in Coyle [Coyle, 2005]. In addition, athletes use training schemes aimed at obtaining peak performance during the decisive stages of the event.

Our model does not yet contain oxygen transport to make it possible to analyse oxygen delivery and uptake. It is, therefore, not yet possible to analyse the effects of illegal (recombinant erythropoietin and autologous blood transfusion) or legal methods (hypoxic exposure) to support increased oxygen uptake with our model, but this may be an interesting future expansion of the model. An important aspect of preparation for a long endurance sports event is optimal nutrition to supply sufficient carbon fuel for long, intense exercise. The effect of nutrition on glycogen and fat stores in the body and the interaction of carbohydrates and fat during exercise are aspects that require extensive model analysis in the future. The glycaemic index or

the glycaemic load, which describe the effect of various carbohydrate formulations can be incorporated in the models [O'Reilly et al., 2010]. The interaction of protein given in combination with carbohydrates may also be incorporated in the model description of carbon substrate handling before, during and after exercise [Zawadzki et al., 1992, Ivy et al., 2002]. An overview of nutritional support for athletic performance is given in the online supplementary material.

A mathematical or computational model always represents a simplified reflection of the real complex system. Even if we build increasingly complex models, we will never completely approach reality. However, this poses no problem because we already have the real system. Restricting ourselves to capture only the essential processes in mathematical equations helps us to understand and predict the system's behaviour. Even if models gradually become quite complex, we can still track all interactions and how they synergize to yield the overall result. The easy accessibility of variables in computer models provides an advantage over the real system. Computational modelling is, therefore, a powerful tool to integrate all the knowledge of human scientists gathered in studies that, by themselves, consider only parts and details of the system. By doing this, we also produce powerful tools to predict the performance of the human body and learn how human body function may be improved in health and disease.

4.10 Conclusions

We conclude that simulation of athletic performance with a computational model of energy turnover and heat transport in the whole human body is feasible. The maximal oxygen uptake capacity of the athlete can be estimated based on performance in a race. The temperatures in various parts of the body such as brain and active muscle groups can be predicted. The changes in the phosphocreatine energy buffer status can be predicted, although glycolytic ATP production and pH regulation will have to be added. The model of carbon flux distribution in central metabolism allows us to investigate hypotheses on the apparent dysbalance of glucose and lactate uptake with cerebral metabolic needs.

4.11 Acknowledgements

This work was supported by a BSIK grant through The Netherlands Genomics Initiative (NGI). It is part of the BioRange programme (project number SP 2.2.1) of The Netherlands Bioinformatics Centre (NBIC), of the Centre for Medical Systems Biology and of The Netherlands Consortium for Systems Biology, which are three Genomics Centres of Excellence funded by the Dutch Government via the NGI. We are very thankful to Daniela Muente for helping us with the figures in this paper.

4.12 Author Contributions

Implementing and simulating and validating the whole body model (sections 4.3-4.6): JHGMvB. Time-resolved model of events at the muscle cell level (section 4.7): HH and JHGMvB. Model of carbon metabolism in brain (section 4.8.2): FS, JHGMvB. Model of gut metabolism (provided in the online supplementary material): TBW,AAdG,JHGMvB. Computer package to calculate metabolic fluxes (section 4.8.1): AKG, JHGMvB.

Chapter 5

Computational estimation of tricarboxylic acid cycle fluxes using noisy NMR data from cardiac biopsies

Hettling, H., Alders, D.J.C, Heringa, J., Binsl, T.W., Groeneveld, A.B.J. and van Beek, J.H.G.M. (2013) Computational estimation of tricarboxylic acid cycle fluxes using noisy NMR data from cardiac biopsies *BMC Systems Biology* 7:82

5.1 Abstract

Background

The aerobic energy metabolism of cardiac muscle cells is of major importance for the contractile function of the heart. Because energy metabolism is very heterogeneously distributed in heart tissue, especially during coronary disease, a method to quantify metabolic fluxes in small tissue samples is desirable. Taking tissue biopsies after infusion of substrates labelled with stable carbon isotopes makes this possible in animal experiments. However, the appreciable noise level in NMR spectra of extracted tissue samples makes computational estimation of metabolic fluxes challenging and a good method to define confidence regions was not yet available.

Results

Here we present a computational analysis method for nuclear magnetic resonance (NMR) measurements of tricarboxylic acid (TCA) cycle metabolites. The method was validated using measurements on extracts of single tissue biopsies taken from porcine heart *in vivo*. Isotopic enrichment of glutamate was measured by NMR spectroscopy in tissue samples taken at a single time point after the timed infusion of ^{13}C labelled substrates for the TCA cycle. The NMR intensities for glutamate were analysed with a computational model describing carbon transitions in the TCA cycle and carbon exchange with amino acids. The model dynamics depended on five flux parameters, which were optimised to fit the NMR measurements. To determine confidence regions for the estimated fluxes, we used the Metropolis-Hastings algorithm for Markov chain Monte Carlo (MCMC) sampling to generate extensive ensembles of feasible flux combinations that describe the data within measurement precision limits. To validate our method, we compared myocardial oxygen consumption calculated from the TCA cycle flux with in vivo blood gas measurements for 38 hearts under several experimental conditions, e.g. during coronary artery narrowing.

Conclusions

Despite the appreciable NMR noise level, the oxygen consumption in the tissue samples, estimated from the NMR spectra, correlates with blood-gas oxygen uptake measurements for the whole heart. The MCMC method provides confidence regions for the estimated metabolic fluxes in single cardiac biopsies, taking the quantified measurement noise level and the nonlinear dependencies between parameters fully into account

5.2 Background

Metabolic fluxes in animal tissues can be identified by measuring the incorporation of stable isotopes in intracellular metabolite pools. To quantify metabolic fluxes, isotope label incorporation is usually measured at several time points [Sauer, 2006], among others in heart tissue [Chance et al., 1983, Malloy et al., 1990, Schroeder et al., 2009]. Heterogeneity of metabolism inside the heart often confounds time series of small tissue samples, therefore a single time point protocol to quantify metabolic fluxes has been developed [van Beek et al., 1998, 1999]. Such single time point measurements in individual samples allow to define spatial profiles of metabolic fluxes in heterogeneous organs [Alders et al., 2007].

The incorporation of stable isotopes (e.g. ^{13}C) in metabolic intermediates can be detected by nuclear magnetic resonance (NMR) spectroscopy or mass spectrometry (MS). The data is then analysed with computational methods that require (i) detailed mathematical models of carbon transitions between the metabolites in the system and (ii) sophisticated optimisation procedures for estimating the flux parameters. In the past, we have developed a bioinformatics method to estimate metabolic fluxes in aerobic metabolism from very noisy NMR measurements resulting from the Labelling with Isotope for a Pre-Steady-State Snapshot (LIPSSS) protocol [Binsl et al., 2010a]. For LIPSSS, isotope labelled substrate for a metabolic pathway is infused for a short, definite period of time, and the metabolism is stopped before a steady state of label incorporation is reached. Finally, pathway metabolites are extracted and measured. Although the original computational analysis method [Binsl et al., 2010a] explores parameter space extensively to avoid local minima, only a rough estimate of parameter confidence regions was obtained by assuming local linearity. Here we introduce a Markov chain Monte Carlo (MCMC) parameter estimation strategy which allows a full description of the confidence regions of the estimated metabolic fluxes, including correlations and nonlinear dependencies between parameter estimates.

Brown et al. [Brown et al., 2004] and Gutenkunst et al. [Gutenkunst et al., 2007a] sampled ensembles of parameter sets for systems biology models with MCMC. Correlations between model parameters were taken into account and confidence bounds for parameters and model predictions were defined [Brown et al., 2004, Gutenkunst et al., 2007a]. Monte Carlo methods have previously been applied to metabolic flux analysis (MFA) in order to handle inaccuracies in data and model [Schellenberger and Palsson, 2009]. Sensitivity analysis by Monte Carlo sampling is also implemented in a ^{13}C MFA analysis software package [Weitzel et al., 2012]. In ^{13}C MFA, MCMC sampling has been used for uncertainty analysis [Kadirkamanathan et al., 2006, Quek et al., 2009], for flux estimation with noisy data [Yang et al., 2005], and for *in silico* experimental design to determine optimal substrate labelling protocols [Schellenberger et al., 2012]. Antoniewicz et al. proposed a different approach of determining confidence bounds on fluxes by calculating the agreement between model and experiment data as a function of the flux of interest [Antoniewicz et al., 2006].

We developed and applied an MCMC procedure to estimate the TCA cycle flux,

carbon substrate uptake, and oxygen consumption from NMR spectra of ^{13}C enriched glutamate sampled at a single time point. For the computational analysis, we expanded the R-package FluxEs [Binsl et al., 2010a]. This analysis was applied to cardiac tissue biopsies flash-frozen 5.5 minutes following ^{13}C acetate infusion in porcine hearts *in vivo*. The method was validated experimentally for a range of cardiac stress conditions. Our first goal was therefore to determine the uncertainty in the estimation of metabolic flux parameters based on the quantified uncertainty in the NMR measurements and in the prior knowledge. The second goal was to validate the computational estimations in experiments *in vivo*.

5.3 Methods

5.3.1 Ethical Statement

The study was approved by the Advisory Board for the Use of Experimental Animals of the Vrije Universiteit Amsterdam. The procedure is in accordance with the American Physiological Society 'Guiding Principles in the Care and Use of Animals', which state that muscle relaxants may be used in conjunction with drugs known to produce adequate anaesthesia.

5.3.2 Experimental Strategy

In this study the metabolic flux in the TCA cycle was measured in tissue biopsies taken from cardiac tissue via the LIPSSS experimental protocol which consists of a brief, timed infusion of ^{13}C labelled acetate in the left anterior descending (LAD) coronary artery of anaesthetized pigs [Alders et al., 2004]. We began with unlabelled acetate which was infused for 30 minutes, in order to establish a stationary metabolic state, followed by [2-^{13}C] acetate for 4 minutes and [1,2-^{13}C] acetate for 1.5 minutes. After exactly 5.5 minutes of ^{13}C enriched acetate infusion, metabolism was arrested by freeze-clamping part of the left ventricular wall of the heart before the isotopic steady state was reached. Biopsies from different regions of this part of the left ventricular wall were cut from the tissue slab after freeze-drying, and divided into approximately nine samples per heart with around 0.1 g dry mass per sample. After extraction with perchloric acid, the ^{13}C NMR multiplets of glutamate were measured. ^{13}C-NMR spectra were obtained at 100.62 MHz and analysed with the MRUI/AMARES software package (more information about tissue preparation, NMR measurement and the package can be found in reference [Alders et al., 2004]).

Up to nine separate multiplet intensities were detected for glutamate. For independent testing of the LIPSSS method and the associated parameter estimation procedures, 'gold standard' myocardial oxygen uptake was calculated from blood flow, hemoglobin content and blood-gas measurements taken before and during acetate infusion [Alders et al., 2004]. Note that these classic oxygen uptake measurements are

entirely independent of the LIPSSS method. We analysed data from LIPSSS samples taken from N = 38 porcine hearts divided into 6 different experimental groups: (i) basal state of the heart (control group, n = 7), two groups with constriction (see below for method) of the coronary vessels to reduce blood flow ((ii) mild stenosis group, n = 7 and (iii) a moderate stenosis group, n = 6), (iv) peripheral venous infusion of dobutamine to induce cardiac stress (dobutamine group, n = 6) or (v) infusion of adenosine for cardiovascular dilatation (adenosine group, n = 4) and (vi) finally, a combination of stenosis and adenosine administration (stenosis + adenosine group, n = 8). In the mild and moderate stenosis groups, LAD blood pressure was adjusted with an occluder to amount to about 70 and 35 mmHg downstream of the occluder, respectively. In the adenosine and stenosis + adenosine groups, adenosine was infused into the LAD at a rate of 100 $\mu g/kg/min$. In the stenosis + adenosine group coronary blood pressure was reduced to about 45 mmHg. In the dobutamine group, dobutamine was infused at a rate of 10 $\mu g/kg/min$. Note that the dobutamine group initially contained 8 hearts from which two were excluded from further analysis, due to a low mean arterial blood pressure and insufficient NMR signal for parameter estimation (see [Binsl et al., 2010a]), respectively.

5.3.3 Anaesthesia and animal experimental procedures

In all groups, sedation was performed with ketamine 15 mg/kg and midazolam 1 mg/kg intramuscularly, and anaesthesia was maintained by continuous infusion of sufentanil (4 $\mu g/kg/hr$), midazolam (0.5 mg/kg/hr), and pancuronium (0.2 mg/kg/hr). The trachea was intubated and the lungs were ventilated with a mixture of 60% oxygen/40% air. Fluid-filled catheters were introduced and hemodynamic parameters collected as previously described (see [Alders et al., 2004]). A continuous infusion of lidocaine was started to help prevent cardiac arrhythmias (9 mg/kg/hr, with an initial bolus injection of 50 mg). Five cm H_2O of positive end-expiratory pressure (PEEP) was applied before opening the thorax. The thorax was opened via a mid-sternal incision and the heart exposed by opening the pericardium. The left hemiazygos vein was tied off to prevent mixing of noncoronary venous blood with coronary venous blood. The LAD was dissected free over a distance of about 2 cm and was catheterised with a 24G catheter. In the stenosis and adenosine + stenosis groups, a custom-made adjustable aluminium occluder was placed around the artery, and LAD pressure was measured.

After finishing instrumentation the animal was allowed to stabilise for at least 15 minutes, the first batch of microspheres (labelled with ^{141}Ce or ^{103}Ru, in random order) was injected into the left atrium for baseline blood flow measurements. The intervention was performed and 30 minutes later a second batch of microspheres was injected for final blood flow measurements. Throughout the procedure hemodynamic data were recorded continuously.

Experimental procedures have been described more extensively previously [Alders et al., 2007, 2004].

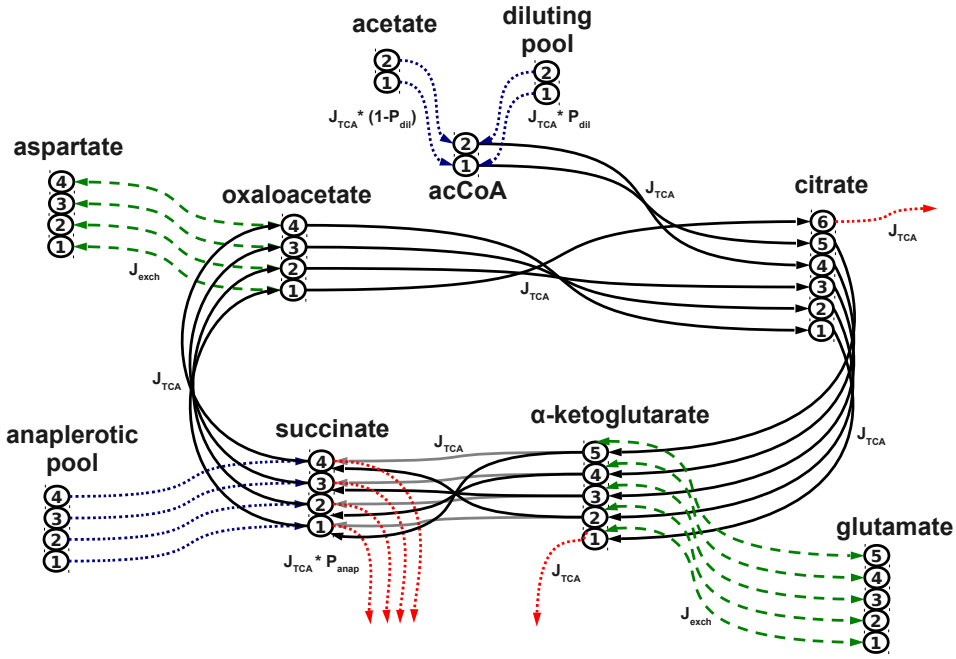

Figure 5.1: Computational model of carbon transitions in the TCA cycle. The numbered circles connected in a string represent single carbon atoms of the corresponding metabolite. Fluxes between carbon atoms of the metabolite pools are indicated by arrows. Blue and red dotted arrows stand for carbon atoms entering and leaving the system, respectively. Green dashed arrows indicate bi-directional exchange fluxes of carbon atoms with amino acids. The parameters determining the conversion rate are shown next to the arrows. Note that there are two possible transitions between α-ketoglutarate and succinate, indicated by arrows of different grey shade. The figure was adapted from Binsl et al. [Binsl et al., 2010a].

5.3.4 Computational Model

The NMR measured enrichment of glutamate with isotopes was analysed with a computational model. The model of carbon transitions in the TCA cycle used in this study was described previously in detail [van Beek et al., 1998, 1999, Binsl et al., 2010a] and is therefore only described here in brief. The model contains ten metabolite pools, consisting of metabolites which contain two to six carbon atoms, and 50 transitions of carbon atoms between the metabolites (Figure 5.1).

Isotopically labelled substrate enters the system via the acetate pool. Acetate is then converted into acetyl coenzyme A (acetyl-CoA), which then enters the TCA cycle. Since acetyl-CoA can also be formed from endogenous unlabelled substrates such as glucose, glycogen, or fatty acids, a diluting pool was introduced to account for dilution of the labelled acetate. The intermediates of the TCA cycle are rep-

resented by the 6-carbon metabolite pool labelled as citrate (which also comprises cis-aconitate and isocitrate), α-ketoglutarate, succinate (including succinyl-CoA) and oxaloacetate (representing a 4-carbon metabolite pool which also comprises malate and fumarate). Glutamate and aspartate are amino acids synthesised by transamination from α-ketoglutarate and oxaloacetate, respectively. The replenishment of TCA cycle intermediates was modelled by an anaplerotic influx connected to succinate. Malloy et al. have given detailed descriptions of the equations for anaplerosis [Malloy et al., 1990, 1996]. The metabolite concentrations were given as fixed parameters in the calculations: the glutamate pool size was measured by biochemical assay in each sample [18], because sensitivity analysis showed that results are sensitive to its value. However, the same sensitivity analysis showed that the metabolite pool concentrations of citrate, α-ketoglutarate, oxaloacetate and aspartate had little effect on the results and these concentrations were taken from previous studies [van Beek et al., 1999, Alders et al., 2004]. All metabolite concentration parameters were therefore fixed and all flux parameters estimated during the Markov chain Monte Carlo procedure (see below). More information about the model and a listing of all model equations can be found in the online supplemental material[1].

The dynamic behaviour of the model is affected by five system parameters (Figure 5.1). The flux parameters J_{TCA} and J_{exch}, were expressed in μmol/(min*g dry weight [dw]) and represent reaction fluxes through the TCA cycle and exchange reactions with amino acids, respectively. The dynamics of incorporation of ^{13}C label from acetate into the acetyl-CoA pool depends on transport in the blood vessels, permeation of the cell membrane, the flux of the conversion of acetate into acetyl-CoA, the flux of acetyl-CoA into the TCA cycle and the acetate and acetyl-CoA pool sizes. Fortunately, the time course of incorporation of ^{13}C label into the acetate pool is almost mono-exponential [Randle et al., 1970] and can be represented by a single time constant which we term T_{trans}. We incorporated this efficient way to represent acetyl-CoA dynamics into our model [van Beek et al., 1999]. The two parameters P_{dil} and P_{anap} account for the degree of dilution of labelled acetate and the rate of anaplerosis relative to TCA cycle flux, respectively. Both are flux parameters which are expressed as fractions of J_{TCA}. J_{TCA} and P_{dil} describe energy and substrate turnover which are our targets to measure and are therefore labelled 'primary parameters'. On the other hand, J_{exch}, T_{trans} and P_{anap} are constrained during parameter estimation by Bayesian priors (see below) and because they are not our primary target parameters they are termed 'auxiliary parameters' which are allowed to vary to determine the uncertainty which they cause in the primary parameters, The LIPSSS estimate for myocardial oxygen consumption is calculated from the primary parameters, (see Equation 5.5 below). Note that primary and auxiliary parameters are estimated together in the same procedure.

[1] http://www.biomedcentral.com/1752-0509/7/82/additional

5.3.5 Matching model simulations to NMR measurements

The computational model described above accounts for all possible carbon isotope labelling states (isotopomers) of each of the metabolites. The system is described by 132 ordinary differential equations (ODEs) to calculate the rate of change of each isotopomer over time. For instance, the metabolite glutamate, which contains 5 carbons, is represented by $2^5 = 32$ ODEs. The isotopomer composition is expressed as fractions of the metabolite concentration of the corresponding pool. Thus, at each time point, the sum of all isotopomer fractions in a pool is 1. All ODEs are then integrated over time to yield the simulated isotopomer fractions. For comparison with the ^{13}C NMR measurements (m_{exp}), simulated NMR multiplet intensities (m_{sim}) were calculated from the simulated isotopomer fractions for the time point at which the sample was taken in the experiment [Binsl et al., 2010a]. To this end all isotopomers contributing to a particular NMR intensity were added. The simulated multiplet intensities are dependent on the values of the five model parameters. To quantify the agreement between model simulation and experimental data we define a least-squares cost function C as a function of the parameter vector θ, in which the squared residuals for all multiplets are weighted by their standard deviations and summed. Additionally, we include Bayesian prior terms in the cost function which reflect prior knowledge on auxiliary parameter values (see below):

$$C(\theta) = \frac{1}{2} \sum_{i \in multiplets} \left(\frac{m_{i,sim} - m_{i,exp}}{\sigma_{i,exp}} \right)^2 + \sum_{j \in \theta} prior(\theta_j) \tag{5.1}$$

The $\sigma_{i,exp}$ represents the measurement error of the NMR intensity. This cost function is used for the optimisation procedures. It is also used as the argument of the normal probability distribution used for the MCMC procedure (see below). The cost function integrates data measured directly in the experiment with literature information incorporated in the priors on parameter values.

5.3.6 Priors on parameter values

The main objective of this study was to estimate J_{TCA} and P_{dil}, the two primary parameters which define aerobic and substrate metabolism and allow the calculation of oxygen consumption in the sample immediately before metabolic arrest. The three remaining parameters T_{trans}, P_{anap}, and J_{exch} are not our target parameters and cannot be determined with great precision. However, these auxiliary parameters are taken into account to evaluate their effect on the estimation of the primary parameters. To improve the estimation and to help avoid local minima in parameter space with physiologically implausible values of the auxiliary parameters, a priori information for such parameters (θ_i) can be directly entered into the cost function by adding a prior term to the cost function in Equation 5.1 for the deviation from a certain reference value θ_i^*

$$prior(\theta_i) = \frac{1}{2}\left(\frac{ln\theta_i - ln\theta_i^*}{\sigma_{ln\theta}}\right)^2 \tag{5.2}$$

where $\sigma_{ln\theta}$ is the standard deviation for the auxiliary parameter in log-space. The advantage of logarithmic parameters is that the parameter values with a Gaussian prior distribution are positive and dimensionless. Note that the prior probability in Equation 5.2 does not include the normalisation factor for the log-normal distribution of $\frac{1}{\sigma\sqrt{2\pi}}$. Normalisation was not necessary because our method applied the Metropolis-Hastings algorithm which uses the ratios of probabilities. In previous studies, the value of T_{trans} had been estimated to be 0.202 min which is compatible with the time constant of the enrichment of acetyl-CoA with radioactive label [van Beek et al., 1999, Binsl et al., 2010a]. We constrain T_{trans} around this prior value with $\sigma_{ln\theta}$ set to 0.336, a high value used in a previous study for unreported experimental errors [Hettling and van Beek, 2011]. This is slightly higher than the value for the standard deviation of these parameters determined in simulations by Binsl et al. [Binsl et al., 2010a]. The central 95% region of the prior for T_{trans} lies between 0.202/1.96 and 0.202 *1.96 min, since $\sigma_{ln\theta} = 0.336 = 1/4 * (ln(\theta_i * 1.96) - ln(\theta_i/1.96))$, (see [Hettling and van Beek, 2011]).

The accurate quantification of the exchange flux J_{exch} between α-amino and α-keto acids was found to be challenging [Chance et al., 1983, Jeffrey et al., 1999]. A previous analysis of the model used in this study revealed a low sensitivity of estimations of J_{TCA} to variations of J_{exch} in the physiological range from $\mu mol/(min * gdw)$ [van Beek et al., 1999]. Reported values of exchange flux in the literature vary substantially. Some report high values for the exchange flux (e.g. 13-fold the flux of J_{TCA} [Randle et al., 1970]). Several other studies report J_{exch} to be approximately equal to J_{TCA} [Nuutinen et al., 1981, Yu et al., 1995]. To address this issue, we set a prior on J_{exch} relative to the value of J_{TCA}. Instead of calculating the prior cost directly from J_{exch}, it is therefore determined by entering the ratio $\theta_i = J_{exch}/J_{TCA}$ into Equation 5.2. The reference value θ_i^* for the ratio is set to 1, based on values for J_{exch}/J_{TCA} reported by Nuutinen et al. [Nuutinen et al., 1981] and Yu et al. [Yu et al., 1995].

Because of the large spread of values found in the literature (see above), we assumed a high standard deviation for the ratio J_{exch}/J_{TCA} and set $\sigma_{ln\theta}$ to $1/4*(ln(\theta_i i* 15) - ln(\theta_i/15)) = 1.345$, with $\theta_i = 1$. It is thereby ensured that J_{exch} lies with 95% probability between $J_{TCA}/15$ and $J_{TCA} * 15$.

For the parameter P_{anap}, the anaplerotic flux relative to the TCA cycle flux, most of the values found in literature were smaller than 1 and the highest experimental value found was reported to be 1 ± 0.3 [Weiss et al., 1989, Lloyd et al., 2004]. Hence the prior cost for P_{anap} was set to be uniform for values of P_{anap} between 0 and 1 combined with a half-normal distribution which had a standard deviation of 0.3 taken from Lloyd et al. [Lloyd et al., 2004] for the values above 1:

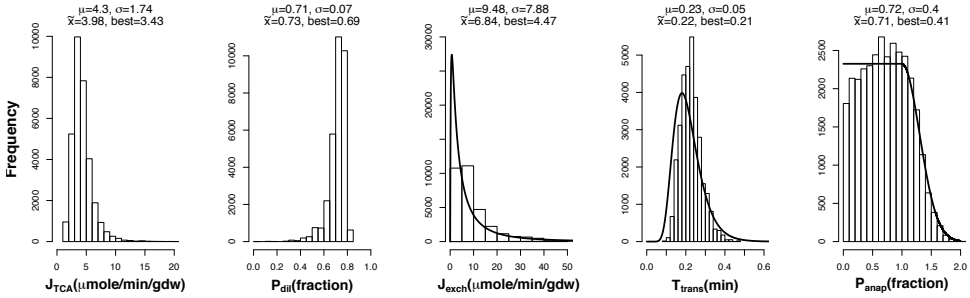

Figure 5.2: Posterior distributions for the parameter ensemble (corresponding to 35000 parameter sets) for one tissue sample of the control group. The probability density functions of the priors for the auxiliary parameters J_{exch}, T_{trans} and P_{anap} are plotted with solid lines. On top of each plot, ensemble mean, standard deviation, median $\tilde{x}$ and best fit value are reported. Note that the probability density functions are scaled to the observed frequencies on the histogram.

$$prior(P_{anap}) = \begin{cases} -ln(c1), & 0 \leq P_{anap} \leq 1 \\ -ln(\mathcal{N}(\mu = 1, \sigma = 0.3) * c2), & P_{anap} > 1 \end{cases} \tag{5.3}$$

with $c1 = 1 - \frac{0.5}{0.5 + \frac{1}{\sqrt{2\pi\sigma^2}}}$ and $c2 = \frac{1}{0.5 + \frac{1}{\sqrt{2\pi\sigma^2}}}$. The normalisation constants $c1$ and $c2$ ensure that the probability density function of the prior is continuous and that its integral is equal to one. $\mathcal{N}$ denotes the normal distribution. The probability density functions for prior(T_{trans}), prior(J_{exch}), and prior(P_{anap}) are shown in Figure 5.2 (solid lines).

5.3.7 Parameter estimation and sampling of parameter ensembles

In biological models, usually many different combinations of parameters can describe the experimental data [Gutenkunst et al., 2007a]. To address this, we decided to not merely rely on a single best-fit of the model parameters to the NMR data for fixed values of the auxiliary parameters, but instead, we systematically generated ensembles of model parameters that fit the data with reasonable precision. This approach clarifies how well the primary parameters are defined by the data despite uncertainty in the NMR intensities and auxiliary parameters. Through the use of an MCMC approach, confidence bounds can be set on the estimated parameter values. Sampling is based on Bayesian inference of a posterior parameter distribution

$$Pr(\theta|D) = Pr(D|\theta) * Pr(\theta) \tag{5.4}$$

where $Pr(D|\theta)$ is the probability of a parameter vector θ to describe the given data D and $Pr(\theta)$ is the prior probability of the parameters (see above). The right-hand

side of Equation 5.4 is equal to $e^{-C(\theta)}$ where the cost function of Equation 5.1 (which includes the priors of Equations 5.2 and 5.4) is used. The probability functions were not all normalised because this was not necessary for the MCMC procedure which relies on the ratios of probabilities rather than absolute values. Note that the cost function (Equations 5.1, 5.2, 5.3) forms the basis of a probability function (Equation 5.4) that defined the ensemble of estimated parameter values.

In order to estimate the model parameters and to quantify the uncertainty of the estimated values, we sampled an ensemble of parameter sets which could describe the available NMR data by performing a random walk through the parameter space through the application of the Metropolis-Hastings algorithm. The starting point of the random walk was an optimised set of parameters, which had been obtained by a grid optimisation strategy introduced by Binsl et al. [Binsl et al., 2010a]. The grid optimisation was designed to cope with a shallow basin shaped by the cost function in order to avoid optimisation towards local minima. The procedure covered the 5-dimensional parameter space by a grid so as to find the best starting point for optimisation. The second phase of optimisation starting at this grid point was then performed using the Nelder & Mead simplex algorithm, and in the third phase we used the Metropolis-Hastings algorithm to sample a parameter ensemble with its probability density proportional to a probability function based on the cost function $C(\theta)$ of Equation 5.1 entered in Equation 5.4.

5.3.8 Quality criteria for flux estimations in NMR samples

In many of the available *in vivo* samples, NMR peak intensities are low and often below the threshold of observability, i.e. often six or seven of the nine multiplets of glutamate are not discernible from noise and were assigned an intensity of zero. In some of these low intensity samples, Monte Carlo sampling leads to very large ensemble standard deviations of the estimated primary parameters. We excluded such samples which did not yield reliable estimates for J_{TCA}. The exclusion criterion was that the standard deviation of J_{TCA} in the posterior parameter ensemble exceeded 10 $\mu mol/(min * gdw)$.

5.3.9 Software package FluxEs

The analysis was performed using the R package FluxEs introduced by Binsl et al. [Binsl et al., 2010a]. In order to process parameter ensembles, a Monte Carlo module was added to the software. This module uses the AMCMC algorithm implemented within the package spBayes [Rosenthal, 2007, Finley et al., 2007]. The AMCMC algorithm is a Metropolis-Hastings variant which automatically adapts the proposal step size for the sampled parameters in the random walk. This leads to quicker convergence to a posterior distribution. For the primary parameters, the time constant of the autocorrelation function of the sampled ensemble was calculated in order to inspect whether the algorithm converged to a stationary distribution. For samples

with a high autocorrelation time in the primary parameters, we visually inspected the parameter trace.

A single model simulation run takes approximately 0.26 seconds on a computer with 2.26 GHz clock frequency. The grid optimisation for a single sample took on average 115 minutes, the subsequent sampling with the adaptive Metropolis-Hastings algorithm took on average 540 minutes per sample.

The calculations for all samples were performed in parallel on the Lisa computer cluster system at SARA Computing and Networking Services (`http://www.sara.nl`). All code required for the analysis and part of the experimental data can be found in the online supplemental material.

5.4 Results

5.4.1 Monte-Carlo sampling

We estimated the TCA cycle flux from the NMR peaks of glutamate for 347 tissue samples from 38 hearts. Applying the exclusion criterion described above we removed 85 low-quality samples - leaving 262 samples for further analysis. For each sample, an ensemble of 35,000 parameter sets was generated with the Metropolis-Hastings algorithm. Although convergence was not the first criterion for sample rejection, all ensemble estimates with a high autocorrelation time constant were rejected according to the quality criterion.

An example of a parameter ensemble for one single sample of the control group is given in Figure 5.2. For T_{trans}, J_{exch}, and P_{anap}, the probability density functions of the prior distributions are plotted together with the histograms of the posterior distributions. The posterior distributions for these auxiliary parameters are very broad and relatively close to their corresponding prior distributions. In this way the MCMC ensemble method allowed defining the uncertainty in the primary parameters taking into account the large spread in auxiliary parameters. Despite the broad distribution of the auxiliary parameters, the estimates for J_{TCA} and P_{dil} form relatively well-defined peaks and their standard deviations are relatively low.

For the primary parameters we can thus provide point estimates for each sample. To determine which measure best reflects the true value of a parameter, we conducted a simulation experiment in which multiple sets of artificial NMR multiplets were generated by model simulation and subsequent addition of Gaussian random measurement noise. The parameters were then re-estimated and we compared the estimates from the best fit after grid optimisation (i.e. the fit with the lowest cost function value and therefore the highest likelihood, see Equation 5.4) and the mean, median, and mode of the Monte Carlo ensemble with the 'true' parameter values from the initial simulation. Regarding the primary parameters, the best fit gave the most reliable point estimate. Below, we therefore report the best fit values for the primary

parameters.

5.4.2 Validation by estimation of myocardial oxygen consumption

In order to validate our flux estimation method we compared the LIPSSS estimated myocardial oxygen consumption (MVO_2, expressed in $\mu mol/(min * gdw)$) with independent 'gold standard' measurements. The 'gold standard' was determined by blood-gas oxygen and blood flow measurements and the LIPSSS estimated oxygen consumption was calculated from the parameter estimates of the model [Alders et al., 2004]. The MVO_2 for a single sample is determined from the primary LIPSSS flux parameters by stoichiometric biochemical relations and can be calculated as follows [Binsl et al., 2010a, Chatham et al., 1995]:

$$MVO_2^{sample} = (2 + P_{dil}) * J_{TCA} \tag{5.5}$$

The MVO_2 determined from blood-gas measurements reflects the oxygen consumption of the entire heart. When averaging the samples taken for LIPSSS measurements to estimate oxygen consumption for the entire heart (MVO_2^{heart}), individual sample sizes were taken into account. As in Binsl et al., the contributions of the individual samples were weighted by the dry weight w^{sample} for each sample [Binsl et al., 2010a]].

$$MVO_2^{heart} = \frac{\sum w^{sample} * MVO_2^{sample}}{\sum w^{sample}} \tag{5.6}$$

For all six experimental groups, the comparison of MVO_2 estimated with the LIPSSS method (from the model parameters P_{dil} and J_{TCA}) with the 'gold standard' oxygen measurements is shown in Figure 5.3. One heart from the stenosis + adenosine group was excluded from the analysis since none of its samples satisfied the quality criterion.

For all groups, LIPSSS MVO_2 correlated with blood-gas MVO_2 relatively well. For the control group oxygen consumption measured by the two methods corresponded, but for the ischaemic conditions (stenosis with and without adenosine), oxygen consumption tended to be lower for the LIPSSS method. We calculated Pearson correlation coefficients of 0.49 for control (n = 7, p = 0.26), 0.69 for mild stenosis (n = 7, p = 0.09), 0.66 for moderate stenosis (n = 6, p = 0.15), 0.99 for dobutamine (n = 6, p = 0.0003), 0.71 for adenosine (n = 4, p = 0.29), and 0.87 for the stenosis + adenosine group (n = 7, p = 0.01). The Pearson correlation for all groups combined was 0.85 (n = 37, $p < 10^{-10}$). The dobutamine group showed higher oxygen consumption than the other groups reflecting the increased cardiac work load. It is important to note that the small tissue biopsies used in the LIPSSS experiment only covered a relatively small cardiac region, in contrast to the physiological blood-gas

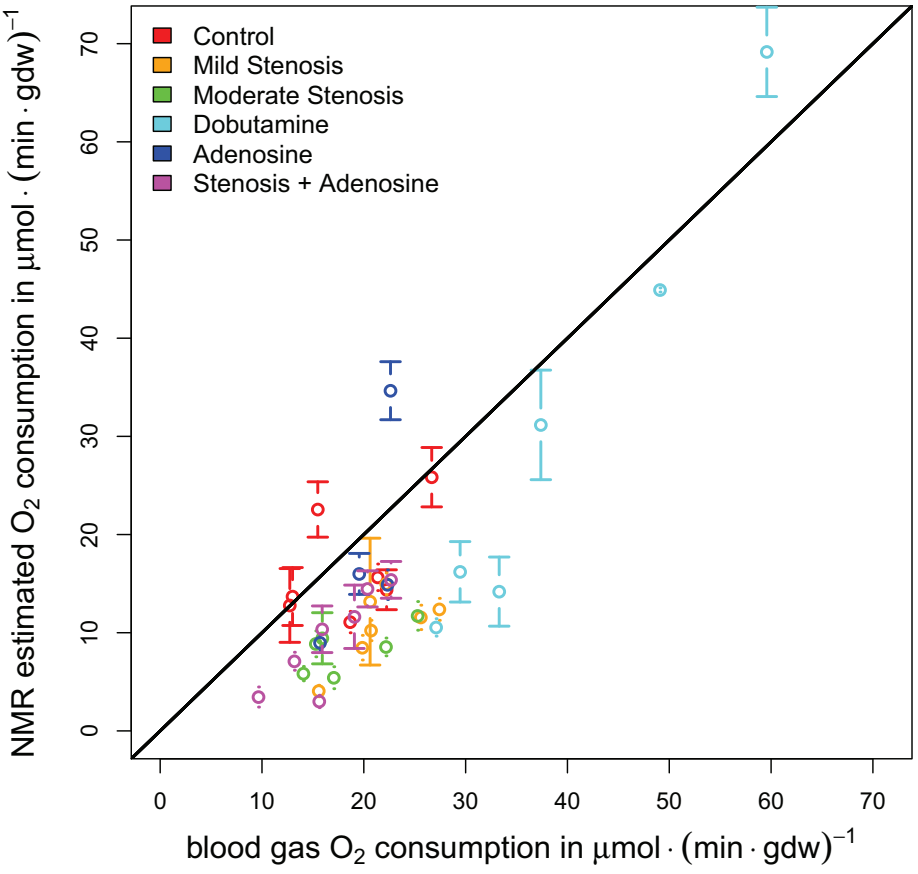

Figure 5.3: 'Gold standard' oxygen consumption (x-axis) calculated from blood gas and blood flow measurements versus oxygen consumption calculated from the parameter estimates derived with the LIPSSS method (y- axis). Each data point corresponds to one heart. The line of identity is plotted in black. Error bars correspond to the standard error of the mean of the oxygen consumption based on NMR measurements over all samples taken from one heart. Note that the error for blood-gas measurements using radioactive microspheres to measure local blood flow is estimated to be about 9% accounting for measurement error, spatial and temporal variation.

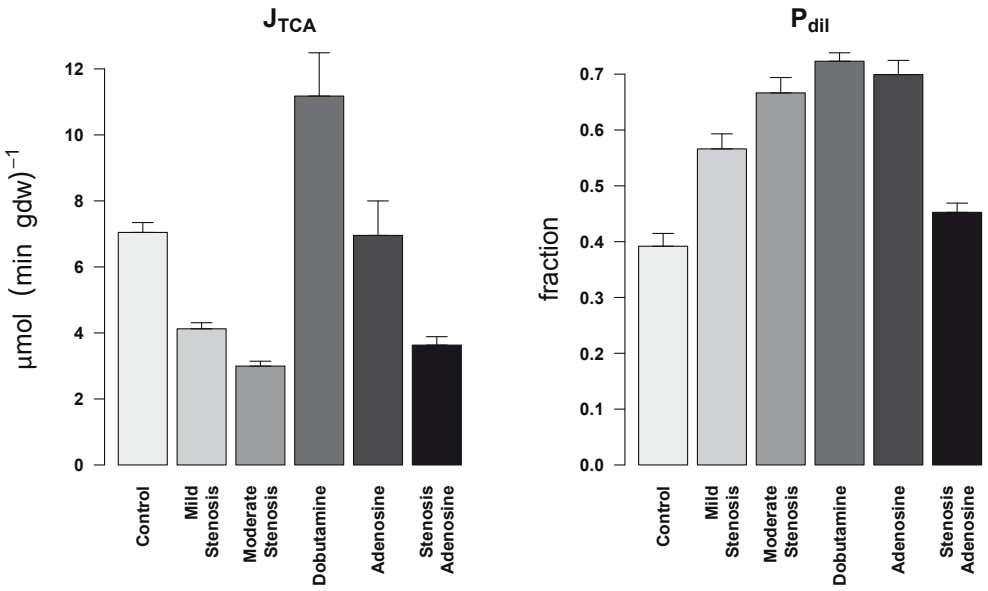

Figure 5.4: Estimates for the primary model parameters for all experimental groups. J_{TCA} and P_{dil} denote the overall TCA cycle flux and the dilution fraction of labelled acetate entering the TCA cycle due to unlabelled endogenous substrates such as glucose and fatty acids, respectively. Estimates were first averaged for each heart and then all hearts were averaged for each group. The error bars represent the standard error of the mean (SEM) of the estimates of all hearts in one experimental group.

measurements which covered the entire left ventricle. Furthermore, the estimation of MVO_2 from the parameters of the TCA cycle model only reflects myocardial oxygen consumption linked to the TCA cycle flux, disregarding other oxygen consuming reactions which were covered by the blood-gas measurements. The oxygen consumption measurements in a small ischaemic region dependent on a constricted coronary artery would be very difficult to obtain with classic blood-gas measurements.

5.4.3 Estimation of TCA cycle fluxes

LIPSSS-based estimates for the primary model parameters under all experimental conditions are shown in Figure 5.4. Estimates for J_{TCA} in the control group averaged 7.04 ± 0.79 (mean ± SEM) $\mu mol/(min * gdw)$. For mild and moderate constriction of the coronary vessels, we estimated J_{TCA} to be 4.12 ± 0.49 and 2.99 ± 0.36 $\mu mol/(min * gdw)$, respectively.

Dobutamine infusion, which stimulates cardiac contraction, leads to a high average J_{TCA} estimate of 11.18 ± 1.31 $\mu mol/(min * gdw)$ of tissue. Estimations for the adenosine group show no difference with the control condition. The TCA cycle flux in the stenosis + adenosine group is in between the mild and moderate stenosis condition.

The relative contribution to the TCA cycle flux of substrates other than labelled acetate, i.e. P_{dil} is higher in all experimental groups compared with the baseline condition (see Figure 5.4). Low fractional acetate usage, i.e. high dilution has been previously documented in experiments with dobutamine [Robitaille et al., 1993].

The estimations of T_{trans} for all the groups did not differ substantially from the prior value of 0.202 minutes (data not shown). Ensembles for the auxiliary parameters J_{exch} and P_{anap} show large standard deviations. This indicates that these parameters cannot be estimated properly from the NMR data. Indeed, the experimental protocol was optimised to estimate the primary parameters, disregarding the auxiliary parameters. Nevertheless the effect of the potential spread in these auxiliary parameters on the uncertainty limits of the primary parameters was taken into account. Estimations of the auxiliary model parameters are described in the supplemental text (section 5.7).

5.5 Discussion

The fluxes of biochemical reactions linked to cardiac energy metabolism are of significant interest. Here we investigated a computational method to quantify fluxes in the TCA cycle using NMR data from ^{13}C labelling experiments in porcine hearts. We took measurement error in the data and uncertainty of model parameters directly into account. To test the method, distinct ^{13}C labelling patterns (isotopomers) in glutamate were measured under six different cardiac stress and control conditions. The data were analysed with a detailed model of carbon transitions in the TCA cycle and two primary flux parameters of interest (reflecting total aerobic metabolism and uptake of the labelled substrate) were estimated. Possible variation in three auxiliary parameters, taken from experimental literature was included in the application of Bayesian priors. To define the uncertainty in estimated flux parameters from measurement error and uncertainty in prior knowledge, we used an MCMC method. As a result, we were able to derive estimates for the TCA cycle fluxes under various experimental conditions despite the high noise level in the available NMR data. For validation, we compared blood-gas measurements of myocardial oxygen consumption with oxygen consumption calculated from our own parameter estimates. The oxygen consumption estimated with our model correlated with the classic physiological measurements for the whole heart (Figure 5.3).

However, because the LIPSSS parameter estimates relied on small samples obtained from the heart while the blood gas measurements represented the oxygen consumption for the whole heart, the LIPSSS estimates are expected to deviate from the whole heart measurement. The deviation may have a random component because of the limited tissue sample size, and a systematic component because of functional differences between regions in the heart. The random component is expected because heterogeneity of blood flow and metabolism has been measured in heart muscle [Alders et al., 2004, Weiss and Sinha, 1978]. A systematic component is expected

especially in the stenosis groups, because the LIPSSS NMR measurements are taken from regions with lower oxygen consumption caused by low perfusion. However, it should be noted that this reasoning is incomplete because the blood gas estimation of oxygen consumption takes the local blood flow measured in the stenosed region into account. Nevertheless, systematic differences between the small region and the average for the whole heart may contribute to the deviation from the line of identity (see Figure 5.3) at low oxygen consumptions.

Additional physiological measurements of oxygen consumption and metabolic fluxes, independent from the stable isotope labelling experiments, are desirable for further validation of our method. Regional rates of oxygen consumption can be measured by measuring oxygen content in small veins [Weiss and Sinha, 1978] with a spectroscopical method in frozen tissue. The latter method is difficult and its validation has been criticised. A further method is the simultaneous determination of myocardial perfusion and oxygen content in small regions of the heart [Reeder et al., 1999]. Oxygen consumption can also be measured using positron emission tomography (PET) and TCA cycle fluxes using *in vivo* NMR (e.g. [Kofoed et al., 2000]). However, these methods mostly have very limited spatial resolution [Reeder et al., 1999] and were in turn subject to rather limited validation themselves. The difficulty in measuring local energy metabolic flux provided motivation to develop our present method in the first place. Despite the limited possibilities, further validation of the LIPSSS method in the future is desirable.

Part of the dataset used here, namely the control and dobutamine group, had been analysed in a previous study [Binsl et al., 2010a]. The estimates of Binsl et al. relied on prior information on the model parameters T_{trans} and P_{anap}. The latter parameter, describing anaplerosis relative to the TCA cycle flux was constrained to 6 $\pm$ 3% of J_{TCA}, based on information from literature studies on isolated hearts. The latter studies however, only accounted for anaplerosis from either propionate [Martini et al., 2003] or from pyruvate [Panchal et al., 2000, 2001]. It has been suggested that relative anaplerosis is often underestimated by conventional approaches, including isotopomer analysis or fractional enrichments of carbons in glutamate [Cohen and Bergman, 1997]. Tracer experiments also exist using ^{13}C labelled propionate that report the relative anaplerotic flux in rat hearts to be much higher than 6%, e.g. 16% [Kasumov et al., 2007] or 29% [Malloy et al., 1996]. Higher relative anaplerotic fluxes were reported during low flow ischaemia, reaching 100% [Lloyd et al., 2004] and 35% [Weiss et al., 1989]. Higher values have also been reported for hypertrophy [Sorokina et al., 2007]. Although our estimates for the parameter P_{anap} in the present study have a relatively low precision, they suggest the possibility that in porcine heart anaplerotic flux in vivo is relatively high in contrast to low values often found in isolated hearts (see section 5.7).

Since three different stenosis conditions were included in the present study, we chose a less constraining Bayesian prior on the parameter P_{anap} which covered a broad range. It is important to note that the Bayesian priors were the same for the analysis of NMR data from all experimental conditions. Despite the use of different choices of

priors on the parameters, and although a higher anaplerotic flux was estimated (see chapter 5.7), our present estimates for fluxes in the control and dobutamine groups did not differ much from the previous estimates of Binsl et al. [Binsl et al., 2010a]. Our estimates for cardiac ischaemia induced by coronary stenosis show that the TCA cycle flux decreases whilst the relative anaplerosis increases (see Figure 5.4 and 5.7) which is compatible with existing literature (see references cited above).

Due to the high velocity of the exchange reactions between α-amino and α-keto acids, the determination of J_{exch} using tracer experiments is expected to be practically infeasible [Jeffrey et al., 1999]. Because of the uncertainty on J_{exch}, we decided to evaluate the effect of variation in J_{exch}. Values for J_{exch}/J_{TCA} found in literature vary between 0.2 and 13 [Randle et al., 1970, Jeffrey et al., 1999], but are often around 1 in the heart [Yu et al., 1995], in contrast to the very high J_{exch}/J_{TCA} reported for the human brain [Mason et al., 1995]. Initial estimations of J_{exch} in our data showed that, particularly in samples with low NMR peak intensity, the simulated isotope enrichment was not very sensitive to J_{exch}. Rather than constraining J_{exch} around an absolute value, we chose to set a Bayesian prior relative to J_{TCA}. The standard deviation of the prior was set to a very high value, reflecting the high variability of J_{exch}/J_{TCA} measurements found in the literature. J_{exch}/J_{TCA} estimated with our method ranged from 0.74 (median dobutamine group) to 1.75 (median control group). Weiss et al. reported a decreased absolute exchange flux compared with control conditions during post-ischaemic reperfusion in rat hearts [Weiss et al., 1993]. A decrease in J_{exch} during stenosis was estimated in the present study (see chapter 5.7).

Literature information on parameter values was incorporated into the analysis as Bayesian priors because of the high noise level in the NMR data. Without using prior information, flux parameters sometimes reach physiologically infeasible regions in parameter space. We investigated the sensitivity of our estimates of the primary parameters to the priors for the auxiliary parameters by re-performing the analysis with doubled prior standard deviations in equations 5.2 and 5.3. The estimate for parameter P_{dil} is rather insensitive to changes in the prior standard deviation (absolute difference in the estimated value averaged over all groups is $4.4 \pm 4.0\%$) while estimates of J_{TCA} are more sensitive to alterations in the priors on auxiliary parameters (average absolute difference $20 \pm 21\%$). Especially in the moderate stenosis group, for which the NMR signals are on average very low, many estimates fail to meet the quality criteria if the standard deviation for all three priors simultaneously was made twice as large. This shows that the estimate of J_{TCA} is sensitive to the prior. However, Bayesian priors are necessary to constrain the estimates within reasonable ranges. It is therefore important to emphasise that the prior values and their standard deviations are not arbitrarily chosen. The prior distributions of P_{anap} and J_{exch} are based on experimental data [Randle et al., 1970, Nuutinen et al., 1981, Yu et al., 1995, Weiss et al., 1989, Lloyd et al., 2004] and were given large standard deviations. The prior on T_{trans} is based on previous estimates [van Beek et al., 1999, Binsl et al., 2010a] and its standard deviation allows for a broad range. We therefore argue

that although constraining parameters in this study was necessary due to the high noise in the data, our framework still allowed to define reasonable point estimates of flux parameters and additionally to define the variability in parameter estimates taking reasonable, sometimes deliberately high, values for the uncertainty of auxiliary parameters into account.

Parametric sensitivity analysis is commonly applied in systems biology [Perumal and Gunawan, 2011]. In this investigation, we chose an approach that explored the multidimensional space around a set of best-fit parameters using a random walk with the Metropolis-Hastings algorithm [Brown et al., 2004, Gutenkunst et al., 2007a, Hettling and van Beek, 2011]. The advantage of this method is that it takes into account possible correlations and nonlinear dependencies between the model parameters. Antoniewiecz et al. approached the problem of defining confidence regions for flux estimates by minimising a sum of squared residuals objective function as a function of the flux value [Antoniewicz et al., 2006]. In their approach, the confidence interval for a flux of interest is derived by setting the flux constant while optimising all remaining fluxes in the system. This step is repeated for a range of fixed flux values until the objective function value exceeds a predefined confidence limit. The advantage of the MCMC approach to determine confidence regions is that it takes all possible correlations between the fluxes into account, since no flux parameter is fixed during the MCMC sampling.

The challenge in analysing the data in this study was the high noise level. Up to seven of the nine measured multiplet intensities could sometimes not be detected. Ensemble modelling proved to be a feasible method to separate samples with flux parameters that could be estimated from samples with poor information on the fluxes in the system. This ensemble approach made it possible to identify 262 out of 347 samples that gave useful estimates for the primary parameters. The quality selection of the samples allowed us to use the best-fit parameters from each sample as a point estimate for the primary parameters. The MCMC approach allowed us to define confidence bounds on all estimated parameter values taking their correlations into account. This is a significant advantage compared with previous approaches, where linearised or analytical methods were used to calculate errors on estimated model parameters [van Beek et al., 1998, 1999, Binsl et al., 2010a].

Adding the Monte Carlo ensemble sampling to the LIPSSS framework enables us to estimate the confidence regions of flux parameters in a single sample. The small size of the tissue samples makes it feasible to identify the spatial variation of flux parameters expected because of the known heterogeneity in the tissue. The physiological meaning of our measurements of heterogeneity in metabolism in heart muscle will be addressed in future studies.

5.6 Conclusions

In this study we improved the LIPSSS method in order to quantify metabolic fluxes using stable isotope labelling integrated with mathematical models of carbon transitions: auxiliary information was taken into account in the form of Bayesian priors and emphasis was placed on the uncertainty analysis of the estimated flux parameters. The method was used to quantify TCA cycle fluxes from noisy NMR measurements in porcine hearts under different physiological conditions. Two important metabolic fluxes could be determined in single biopsies taken during animal experiments and confidence regions could be calculated for single samples.

5.7 Supplemental text: Estimation of auxiliary model parameters

For the auxiliary model parameters J_{exch} and P_{anap}, accurate point estimates were not expected (see above). Nevertheless, the uncertainty of the auxiliary parameters was explicitly taken into account in the ensemble sampling. To give an impression of the values and variation of these auxiliary parameters, Figure 5.5 shows box plots of the distributions as a result of combining all valid parameter ensembles for all samples per experimental condition.

Please note that these 'grand ensembles' aggregate the variability in the single NMR measurements, the spatial variability between the different samples and the variation amongst different hearts. The 'grand median' of the combined distributions for P_{anap} and J_{exch} and also the best-fit values averaged over all samples in all hearts are given in Figure 5.5. Here we report the median of the 'grand ensemble' rather than the mean, because it is more robust when dealing with skewed distributions. In the control hearts, anaplerosis has a median of the combined ensemble at 31% of J_{TCA}. This is higher than reported by Binsl et al. [Binsl et al., 2010a] (see 5.5). Relative anaplerosis tends to be increased under ischaemic conditions: for mild stenosis, moderate stenosis, and the stenosis + adenosine groups, P_{anap} becomes 39%, 71%, and 63% (grand median) of J_{TCA}, respectively. These findings agree with studies on perfused rat hearts which report a higher relative anaplerotic flux during ischaemia [Lloyd et al., 2004]. The determination of the parameter J_{exch} using ^{13}C MFA is a challenging task [Chance et al., 1983, Jeffrey et al., 1999]. As shown previously, the NMR intensities are insensitive to high J_{exch} values and consequently J_{exch} values in the high range cannot be defined accurately. Figure 5.5 presents a large number of outliers for the estimates of J_{exch} with our ensemble method. For all stress conditions, J_{exch} is found to be lower compared with the control condition. This is corroborated by a lower exchange flux from α-ketoglutarate to glutamate during dobutamine stress previously reported for dog hearts [O'Donnell et al., 2004]. The 'grand median' J_{exch} for control and dobutamine group was 12.5 and 7.6 $\mu mol/(min * gdw)$ which is in agreement with previous estimates by Binsl et al. [Binsl et al., 2010a].

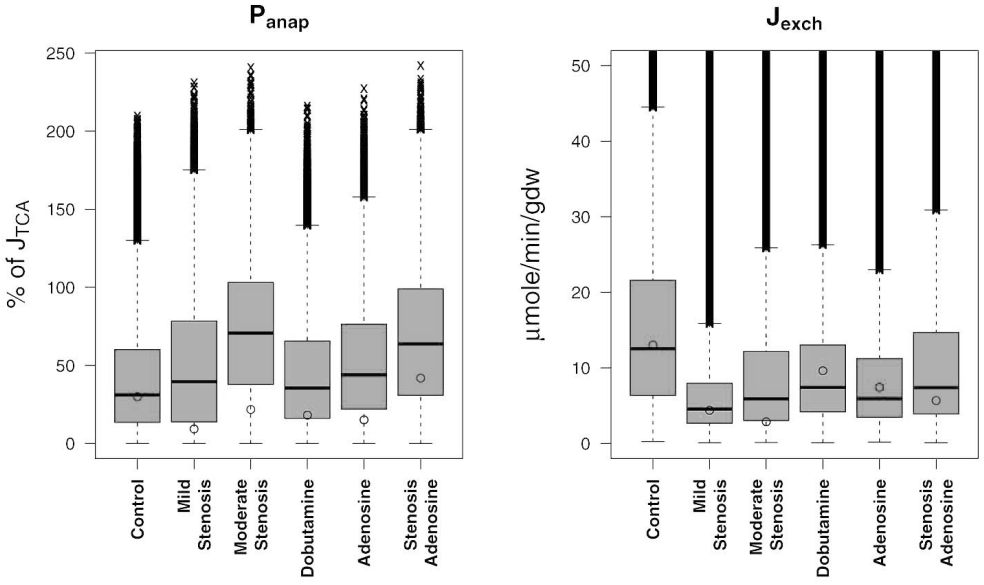

Figure 5.5: Composite ensemble distributions for the auxiliary parameters P_{anap} and J_{exch} representing relative anaplerosis and the exchange flux between TCA cycle intermediates and amino acids, respectively. Colored boxes show the interquartile range of the distributions and their medians. The ends of the whiskers stand for the lowest and highest value still within 1.5 times the interquartile range of the lower and upper quartile, respectively. Best-fit parameter values averaged over all samples and all hearts per group are plotted as open circles. Outliers are plotted with the symbol 'x'. Note that in the right hand plot outliers above 50 $\mu mole/(min * gdw)$ have been omitted.

5.8 Acknowledgements

We thank SARA Computing and Networking Services (`http://www.sara.nl`) for their support in using the Lisa Computer Cluster. We thank Frans de Kanter for assistance with the NMR measurements. We also thank Erik van Dijk for scientific advice and Irisa Ono for editorial assistance.

5.9 Author Contributions

DJA, ABJG and JHGMvB designed and conducted the animal experiments. HH, TWB and JHGMvB designed and conceived the in silico experiments, HH and TWB performed the in silico experiments. HH, DJA, JH and JHGMvB analysed the data. HH and JHGMvB wrote the manuscript. All authors read and approved the final manuscript.

Chapter 6

BiGGR: an open source R package for constraint-based modelling of metabolism using reconstruction databases

Hettling, H.[*]**, Gavai, A.K.**[*]**, Supandi, F.**[*] **and van Beek, J.H.G.M.**(2013) BiGGR: an open source R package for constraint-based modelling of metabolism using reconstruction databases. *Revised version submitted.*

[*]These authors contributed equally.

6.1 Abstract

Predicting the distribution of metabolic fluxes in biochemical reaction networks using Flux Balance Analysis (FBA) is of major interest in systems biology. Several databases provide metabolic reconstructions for different organisms. These reconstructions are curated and include genetic and biochemical literature information. Software to analyse flux distributions exists, among others for the proprietary Matlab environment. Given the large user base of the R computing environment and the wealth of analysis tools R provides, an implementation of flux analysis in R was desirable.

We developed the R software package BiGGR which incorporates public metabolic reconstruction databases, providing direct access to the BiGG database and the reconstruction of human metabolism Recon2. Models can be assembled by querying the databases for pathways, genes or reactions of interest. By linking to linear inverse modelling algorithms, fluxes can be estimated, among others by optimising a cost function. BiGGR allows to estimate flux distributions by minimising differences between experimental measurements and model predictions and can generate ensembles of possible flux distributions. R analysis tools and newly written R functions are easily mixed with the functionality provided by BiGGR. BiGGR automatically visualises selected parts of metabolic networks using hypergraphs, with hyperedge widths proportional to estimated flux intensity. BiGGR supports import and export of models encoded in the Systems Biology Markup language (SBML) and is therefore interoperable with different modelling and analysis tools. Note that the BiGGR version presented here is based upon a preliminary version introduced in section 4.8.1.

Availability

BiGGR is freely available via the Bioconductor repository at `http://www.bioconductor.org/packages/release/bioc/html/BiGGR.html`.

6.2 Introduction

Metabolism directly reflects cellular functioning. If the biochemical reactions that operate in a cell type are known along with uptake or release of some metabolites, the distribution of metabolic flux in the metabolic system can often be predicted. The reconstruction of the metabolic network of an organism is the starting point. In recent years, genome-scale metabolic networks have been reconstructed for various organisms, such as microorganisms, animals and humans [Radrich et al., 2010]. BiGG ([Schellenberger et al., 2010]) is a knowledgebase of reconstructions of metabolism consisting of large lists of metabolites and reactions for *M. barkeri*, *S. cerevisiae*, *H. pylori*, *E. coli*, *P. putida*, *M. tuberculosis* and *S. aureus*. Recon 2 [Thiele et al., 2013]

is presently the most comprehensive reconstruction of human metabolism. The reconstructions recorded in these databases consist of genes, proteins, metabolites and reactions that are connected with each other forming metabolic networks. Several FBA software tools support import from databases [Lakshmanan et al., 2012]. The free COBRA 2.0 toolbox allows to quantitatively predict cell metabolism from the reconstructions in the BiGG database [Schellenberger et al., 2011], but requires the commercial software package MATLAB. It is useful to have similar analysis tools available in the open source R environment because this has a wide user base in statistics and genomics. Recently two other R packages appeared for constrained based modelling [Gelius-Dietrich et al., 2013, Gangadharan and Rohatgi, 2012]. However, these packages have a different focus than BiGGR which is specifically designed for automated model generation by querying metabolic reconstruction databases and the visualisation of FBA results.

6.3 Software features

The R package BiGGR comprises the databases from BiGG [Schellenberger et al., 2011] and the recent reconstruction of human metabolism Recon 2 [Thiele et al., 2013] as SBML objects. Other metabolic reconstructions can be imported from e.g. the BioModels database [Le Novere et al., 2006]. BiGGR provides functionality to query the databases for specific pathways, reactions or genes and select subnetworks to which FBA can be applied. FBA is conducted by converting a network into a linear inverse model, which is then solved using linear programming algorithms initially implemented for analysis of ecological models [Oevelen et al., 2009]. In addition to model assembly and analysis, automatic visualisation of selected parts of metabolic reconstructions and estimated reaction rates is implemented using hypergraphs which provide graphically intuitive plots of biochemical pathways [Klamt et al., 2009]. A detailed tutorial is available as a vignette within the package documentation. BiGGR provides the following functionalities (see Figure 6.1 for a graphical summary):

- **Model creation:** Models can be assembled by querying the metabolic reconstruction databases for specific pathways (e.g. glycolysis), lists of reactions or metabolites or gene identifiers.
- **Model import/export:** Models in SBML format can be easily imported into BiGGR for further analysis. Model files exported from the web interface of the BiGG database can also be imported. Each model created or modified within BiGGR can easily be exported in SBML format.
- **Flux estimation:** BiGGR uses linear inverse model (LIM) approaches for flux estimation. The fluxes in an underdetermined system can be calculated based on a linear function (i.e. a weighted sum of the unknown variables) which is either minimised (a 'cost' function) or maximised (a 'profit' function). The function to

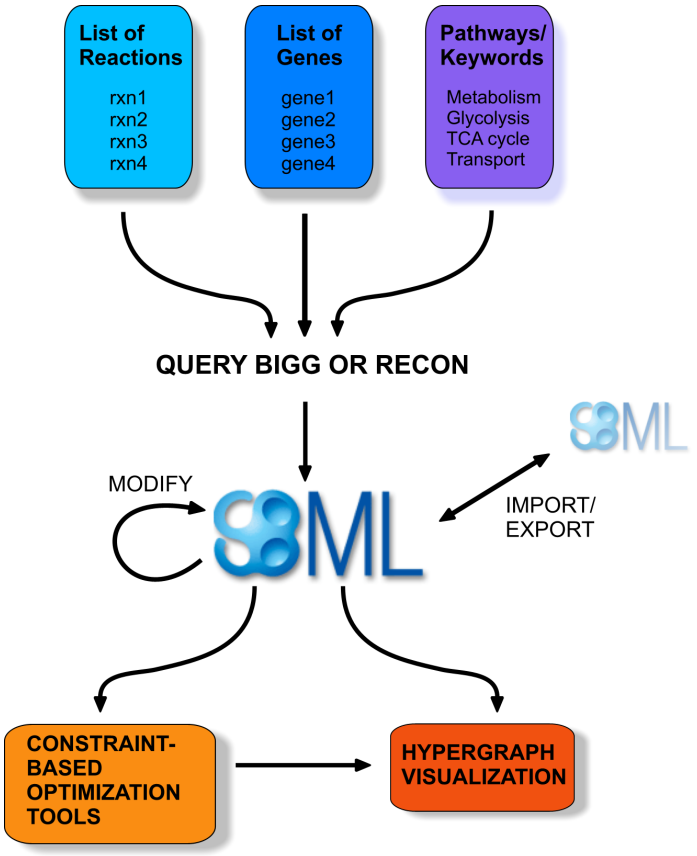

Figure 6.1: Overview of BiGGR functionality.

be minimised or maximised can be subject to inequality and equality constraints, e.g. due to irreversibility of fluxes and the metabolic steady state assumption, respectively. BiGGR can also generate ensembles of parameter combinations according to the likelihood of explaining measured data.

- **Visualisation:** BiGGR provides automatic visualisation of a network with hypergraphs using the hyperdraw package [Murrell, 2013]. The graph displays metabolites connected with each other using hyperedges which represent reactions. Edge widths represent the intensity of estimated fluxes. It usually works best to plot a selected subset of metabolites or reactions.

6.4 Discussion and conclusion

The BiGGR open source package is built in R and Bioconductor [Gentleman et al., 2004], and derives input from metabolic reconstruction databases. BiGGR can be used to automatically construct mathematical and graphical representations of metabolic networks on the fly. BiGGR provides easy access to metabolic flux analysis for the large user base of the R environment. The vast number of open source analysis tools available in R are easily combined with the functionality in BiGGR. By supporting the modelling standard SBML, BiGGR can be used in combination with other modelling tools, as e.g. sybil [Gelius-Dietrich et al., 2013].

6.5 Acknowledgements

The authors would like to thank Karline Soetaert for excellent advice on linear inverse modelling routines.

6.6 BiGGR package vignette

The BiGGR package vignette provides information on how to install and use the package and gives an example of a model analysis with BiGGR. The vignette can be found within the package source at `http://www.bioconductor.org/packages/devel/bioc/html/BiGGR.html`.

6.6.1 Introduction

The main purpose of this package is to analyse metabolic systems and estimate the biochemical reaction rates in metabolic networks. BiGGR works with the BiGG [Schellenberger et al., 2010] database and with files encoded in the Systems Biology Markup Language (SBML) from other sources. The BiGG database stores reconstructions of metabolic networks and is freely accessible. BiGGR is an entirely open source alternative for a more extensive software package, COBRA 2.0, which is available for Matlab [Schellenberger et al., 2011]. BiGGR makes it easy to apply a big variety of open source R packages to the analysis of metabolic systems. Although it contains less functionality than COBRA, BiGGR may be convenient for R users. The BiGG system provides metabolic reconstructions on humans, *M. barkeri*, *S. cerevisiae*, *H. pylori*, *E. coli* and *S. aureus*. BiGGR also works with the new reconstruction of human metabolism Recon 2 [Thiele et al., 2013]. These reconstructions consist of genes, metabolites, reactions and proteins that are identified and connected with each other to form a network structure. The BiGGR package provides various functions to interface to the BiGG database, and to perform flux balance analysis (FBA) after importing selected reactions or pathways from the database. Other functions included

in this package allow users to create metabolic models for computation, linear optimisation routines, and likelihood based ensembles of possible flux distributions fitting measurement data. To this end, BiGGR interfaces with the LIM package [Oevelen et al., 2009]. BiGGR provides models in standard SBML R object format for each organism within the BiGG database as well as the new reconstruction of human metabolism from the Biomodels database [Thiele et al., 2013] (see 'data' directory in the package). This format allows easy construction of the stoichiometric matrix of the entire system which may serve as the core of further computational analysis. Finally, the package allows automatic visualisation of reaction networks based on a hypergraph framework using the hyperdraw [Murrell, 2013] package.

6.6.2 Installation

BiGGR is installed as follows from the R console:

```
> source("http://bioconductor.org/biocLite.R")
> biocLite("BiGGR")
```

BiGGR depends on the Bioconductor packages rsbml [Lawrence, 2013], hyperdraw [Murrell, 2013] (which in turn requires the package hypergraph) and the CRAN package LIM [Oevelen et al., 2009]. For detailed installation instructions of the dependencies we refer to the package documentations at `http://www.bioconductor.org/` and `http://www.cran.r-project.org/`.

6.6.3 Example: A flux balance analysis with BiGGR

The package is imported as follows:

```
> library("BiGGR")
```

To get an overview about the functions and databases available in the package, you can use:

```
> library(help="BiGGR")
```

The reference manual which describes all functions of BiGGR in detail can be found in the documentation directory ('doc') of the package. In the following we will provide a step-by-step guide demonstrating a flux estimation procedure in a model of human brain metabolism. The general work flow using this package consists of the following steps:

- Retrieve a model in SBML object format as provided in the package (alternatively an R object containing the model can be generated from an SBML file)

- Specify optimisation objective and model constraints and create a LIM model file as input for the linear programming package LIM
- Estimate the reaction fluxes with linear programming
- Visualise the results using the hypergraph framework

Generate Model

There are several ways to create a model within BiGGR:

- Query one of the databases contained in the BiGGR package (use the command `data()` to see all available databases). You can query with a list of genes (function `buildSBMLFromGenes`), a list of reaction identifiers (`buildSBMLFromReactionIDs`) or for specific pathways (`buildSBMLFromPathways`).
- Alternatively: Retrieve a text file with metabolic reactions from the web interface of the BiGG database (`http://bigg.ucsd.edu/bigg/main.pl`). The user can query and select reactions from BiGG which can then be exported in SBML or text format. BiGG reactions saved in text format can be converted to an internal SBML object by the function `buildSBMLFromBiGG`. An SBML file can be imported using the `rsbml_read` function from the rsbml package.

Below we will demonstrate how to build an SBML model from a set of reaction identifiers using the Recon 1 database. The list of reaction IDs can be found in the 'extdata' subdirectory in the package. The model comprises the reactions of glycolysis, pentose-phosphate pathway and TCA cycle. Note that the model was already introduced in section 4.8.2.

```
> ##load reaction identifiers from package examples
> file.name <- system.file("extdata",
+                          "brainmodel_reactions.txt",
+                          package="BiGGR")
> reaction.ids <- scan(file.name, what=" ")
> ##load database
> data("H.sapiens_Recon_1")
> ##build SBML model
> sbml.model <- buildSBMLFromReactionIDs(reaction.ids, H.sapiens_Recon_1)
```

The model object `sbml.model` is an rsbml object of class `Model`. It has 92 metabolites participating in 73 reactions in 3 compartments.

Specify constraints, optimisation objective and estimate fluxes

After building the model, we specify additional parameters necessary to run the flux estimation. In the present model, several metabolites are unbalanced because not all the biochemical reactions involving them are represented inside the model. Another unbalanced situation is when metabolites accumulate inside or outside the cell. These metabolites must therefore not be subjected to the equality constraints (i.e. the steady state constraint) of the linear programming routine for flux estimation. These metabolites are termed external metabolites or, in short, externals. The objective of this flux balance analysis is to maximise the net ATP production in the reaction network given the constraints in the model. Note that, of course, also minimising a linear function of fluxes in the model is possible in BiGGR ('loss' function as opposed to 'profit' function). Below we specify the objective function and the external metabolites.

```
> ##following term is to be maximised
> maximize <- "R_ATPS4m - R_NDPK1m - R_HEX1 - R_PFK - R_PGK + R_PYK"
> ##specify the external metabolites of the system
> externals <- c("M_glc_DASH_D_e", "M_lac_DASH_L_e", "M_ala_DASH_L_e",
+                "M_gln_DASH_L_e", "M_h2o_e", "M_co2_e",
+                "M_o2_e", "M_h_e", "M_o2s_m",
+                "M_adp_c", "M_atp_c", "M_pi_c",
+                "M_h_c", "M_nadp_c", "M_nadph_c",
+                "M_na1_c", "M_na1_e", "M_gln_DASH_L_c",
+                "M_nh4_c", "M_pyr_e")
```

Additional equality and inequality constraints can be given for fluxes for which the values are known beforehand, e.g. if they rely on experimental measurements. Below we use measurements of cerebral metabolic substrate uptake and release rates of glucose, lactate, glutamine and pyruvate in human brain [Lying-Tunell et al., 1980]. Assuming a brain mass of 1.4 kg, the rates for the fluxes given in `equation.vars` are constrained to the values given in `equation.values`. Based on measurements of the pentose phosphate pathway flux in the brain [Dusick et al., 2007], also the flux of the glucose-6-phosphate dehydrogenase reaction (R_G6PDH2R) is constrained. The total oxygen uptake rate is constrained to be not higher than six times glucose uptake rate according to the balance equation of aerobic respiration. Equality and inequality constraints are given as lists in the variables `eqns` and `ineq`. Finally a LIM model file is created using the function `createLIMFromSBML`.

```
> ##equality constraints
> equation.vars <- c("R_GLCt1r", "R_L_LACt2r", "R_GLNtN1",
+                    "R_PYRt2r", "R_GLUDC", "R_G6PDH2r")
> equation.values <- c(0.2842,-0.01288,-0.0154,
+                      -0.00336,0.11368,0.0196098)
```

```
> eqns <- list(equation.vars, equation.values)
> ##inequality constraint: R_O2t is smaller or equal to 6 * R_GLCt1r
> inequalities <- list("R_O2t", "6 * R_GLCt1r", "<=")
> ##write LIM file to system's temporary directory
> limfile.path <- tempfile()
> createLIMFromSBML(sbml.model, maximize, equations=eqns,
+                   inequalities=inequalities,
+                   externals=externals, file.name=limfile.path)
```

Running simulations to estimate fluxes

BiGGR uses Linear Inverse Models for estimating the fluxes as provided by LIM. All the functionality of this package can be used in this framework. The function `lsei` in LIM provides least squares estimation with equalities and inequalities which is useful to fit the model to biochemical measurements of metabolite exchange. The interface to LIM's `lsei` in BiGGR is `getRates` which takes the model file (or a LIM object) as an input parameter to estimate the fluxes with respect to the objective function.

```
> rates <- getRates(limfile.path)
```

Sampling of feasible flux distributions

Experimentally quantified fluxes are always subject to measurement error. In the above example, the rates for, among others glucose and glutamine uptake (R_GLCt1r and R_GLNtN1, respectively) and uptake of lactate and pyruvate (R_L_LACt2r and R_PYRt2r) were fixed. However, it is of interest how the estimated fluxes vary if measurement error on the known fluxes is taken into account. BiGGR provides the functionality to calculate the uncertainty of all estimated fluxes by performing a random walk in the feasible flux space with a Markov chain Monte Carlo (MCMC) method. To this end, BiGGR provides an interface to the `xsample` function from the package limSolve [den Meersche et al., 2009]. Ensembles of feasible flux vectors within the precision limits of the known fluxes can be sampled with the function `sampleFluxEnsemble`. As an example, we will sample an ensemble of 100000 flux vectors within the precision limits of the data [Lying-Tunell et al., 1980] given as the standard deviation. Note that standard deviations were estimated from the reported median and range using an approach of Hozo et al. [Hozo et al., 2005]. Starting point for the random walk is the previously optimised flux vector. For quicker convergence of the MCMC procedure, we set the jump length manually (see `?sampleFluxEnsemble` for details).

```
> ##specify the fluxes with uncertainty given as SD in a data frame
> uncertain.vars <- data.frame(var=equation.vars[c(1,2,3,4,6)],
+                              value=equation.values[c(1,2,3,4,6)],
```

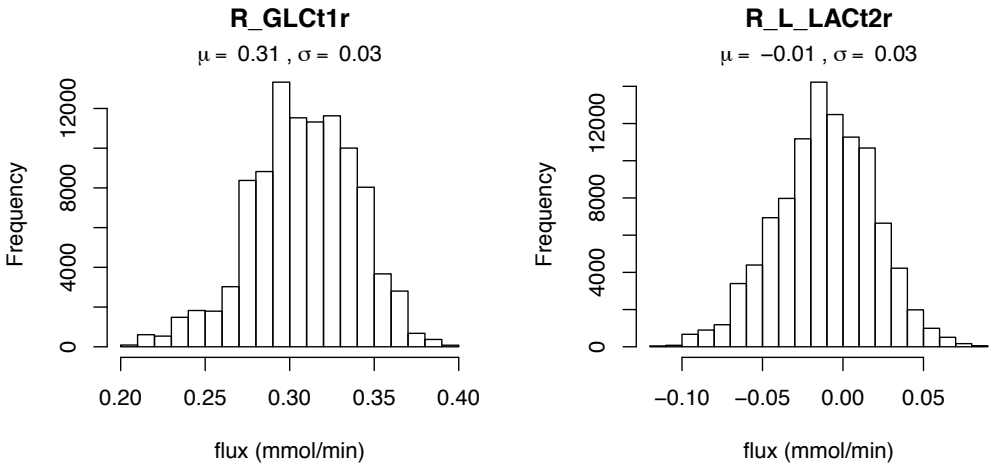

Figure 6.2: Posterior distributions of the fluxes R_GLCt1r and R_L_LACt2r after the sampling with Markov Chain Monte Carlo.

```
+                                sd=c(0.058,0.032,0.034,0.004, 0.007))
> ##sample feasible flux distributions with MCMC
> ensemble <- sampleFluxEnsemble(limfile.path, uncertain.vars,
+                                 x0=rates, iter=100000, jmp=0.1)
>
```

The sampled posterior distributions can then simply be plotted as histograms as shown in figure 6.2 for glucose uptake and lactate release.

```
> par(mfrow=c(1,2))
> hist(ensemble[,"R_GLCt1r"],
+      xlab="flux (mmol/min)", main="R_GLCt1r")
> mtext(bquote(mu== ~.(round(mean(ensemble[,"R_GLCt1r"]),2))
+              ~ "," ~ sigma== ~.(round(sd(ensemble[,"R_GLCt1r"]),2))))
> hist(ensemble[,"R_L_LACt2r"],
+      xlab="flux (mmol/min)", main="R_L_LACt2r")
> mtext(bquote(mu== ~.(round(mean(ensemble[,"R_L_LACt2r"]),2))
+              ~ "," ~ sigma== ~.(round(sd(ensemble[,"R_L_LACt2r"]),2))))
```

Furthermore, it is now possible to assess the effect of possible measurement error in R_GLCt1r and R_O2t on other fluxes present in the system. As an example, we calculate the net rate of ATP production for the whole ensemble from the linear flux combination R_ATPS4m - R_NDPK1m - R_HEX1 - R_PFK - R_PGK + R_PYK. Note that the net ATP production was the profit function of the flux balance analysis presented above.

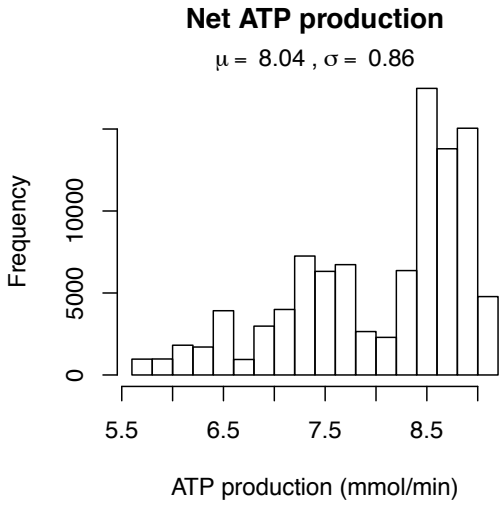

Figure 6.3: Posterior distribution of net ATP production rate in the model.

```
> ##calculate net ATP production for each flux vector in the ensemble
> atp.prod.ens <- eval(parse(text=maximize), envir=data.frame(ensemble))
> hist(atp.prod.ens,
+      xlab="ATP production (mmol/min)", main="Net ATP production")
> mtext(bquote(mu== ~.(round(mean(atp.prod.ens),2))
+              ~ "," ~ sigma== ~.(round(sd(atp.prod.ens),2))))
```

Figure 6.3 shows the spread of the rates of net ATP production given the variability in the measured input fluxes. In this way, the uncertainty of the objective function value can be investigated with respect to possible measurement noise of the fluxes in the model.

Visualisation of networks and fluxes

BiGGR provides visualisation using hypergraphs. To this end, BiGGR uses the package hyperdraw which in turn uses the Graphviz engine. Hypergraphs are graphs which can connect multiple nodes by one edge. Metabolites are represented by nodes and reactions are represented by edges connecting the nodes. The fluxes of the biochemical reactions can be represented by the width of the edges (a wider edge corresponds to a higher flux value). An SBML model can be converted into a hyperdraw object using the function `sbml2hyperdraw`. Since many models contain numerous metabolites and reactions, a 'human readable' automatic graphical representation of the system in one single plot is often infeasible. Therefore, specific subsets of metabolites and/or reactions can be passed as an argument to the `sbml2hyperdraw` function and only

metabolites or reactions belonging to the specified sets are visualised. Below we will visualise selected metabolites and reactions in the glycolytic pathway and the pentose phosphate pathway, which is a subset of our example model. As a second argument we pass the reaction rates calculated in 6.6.3 in order to represent the reaction rates by the width of the edges.

```
> relevant.species <- c("M_glc_DASH_D_c", "M_g6p_c", "M_f6p_c",
+                       "M_fdp_c", "M_dhap_c", "M_g3p_c",
+                       "M_13dpg_c", "M_3pg_c", "M_2pg_c",
+                       "M_pep_c", "M_pyr_c", "M_6pgl_c", "M_6pgc_c",
+                       "M_ru5p_DASH_D_c", "M_xu5p_DASH_D_c",
+                       "M_r5p_c", "M_g3p_c", "M_s7p_c")
> relevant.reactions <- c("R_HEX1", "R_PGI", "R_PFK", "R_FBA",
+                         "R_TPI", "R_GAPD", "R_PGK", "R_PGM",
+                         "R_ENO", "R_PYK",  "R_G6PDH2r", "R_PGL",
+                         "R_GND", "R_RPE", "R_RPI", "R_TKT1")
> hd <- sbml2hyperdraw(sbml.model, rates=rates,
+                      relevant.species=relevant.species,
+                      relevant.reactions=relevant.reactions,
+                      layoutType="dot", plt.margins=c(20, 0, 20, 100))
```

The hypergraph object can then simply be plotted using the plot function:

```
> plot(hd)
```

The resulting plot is shown in Figure 6.4. Flux values are displayed following each reaction identifier. The forward direction is defined in the BiGG database according to biochemical conventions, but if the actual calculated flux is backwards according to the definition the arrow is dashed [1]. Additional graphical arguments are documented in the help file (see `?sbml2hyperdraw`).

Below, we give various reactions and metabolites in the TCA cycle which are present in our example model and plot all components using a circular layout (see Figure 6.5):

```
> relevant.species <- c("M_cit_m", "M_icit_m" , "M_akg_m",
+                       "M_succoa_m", "M_succ_m", "M_fum_m",
+                       "M_mal_DASH_L_m", "M_oaa_m")
> relevant.reactions <- c("R_CSm", "R_ACONTm", "R_ICDHxm",
+                         "R_AKGDm", "R_SUCOAS1m", "R_SUCD1m",
+                         "R_FUMm", "R_MDHm", "R_ICDHyrm", "R_ME1m",
+                         "R_ME2m", "R_ASPTAm","R_AKGMALtm", "R_GLUDym",
```

[1] Note that in the BiGGR program available online, arrows depicting reverse fluxes are coloured red.

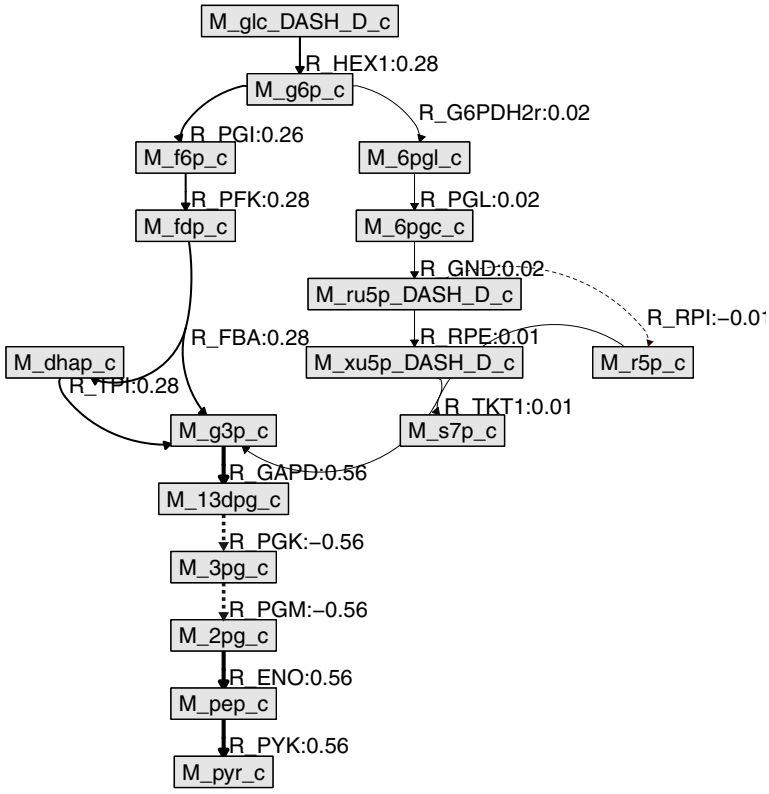

Figure 6.4: Estimated fluxes in the glycolytic pathway and parts of the pentose phosphate pathway. For each reaction, the arrow points in the direction of the calculated flux. If that is backward relative to the direction defined as forward in the metabolic reconstruction, the arrow is dashed. Note that only a subset of all metabolites and reactions is plotted.

```
+                          "R_ABTArm", "R_SSALxm","R_CITtam")
> hd <- sbml2hyperdraw(sbml.model, rates=rates,
+                      relevant.reactions=relevant.reactions,
+                      relevant.species=relevant.species,
+                      layoutType="circo", plt.margins=c(100, 235, 100, 230))
> dev.new() ##Open a new plotting device
> plot(hd)
```

In this example, reactions with a flux equal to zero are displayed in grey. Note that metabolites which are not specified are not plotted, even if reactions in which they participate are drawn. This is for instance the case for the exchange reaction below:

```
M_akg_m + M_mal_DASH_L_c -> M_akg_c + M_mal_DASH_L_m
```

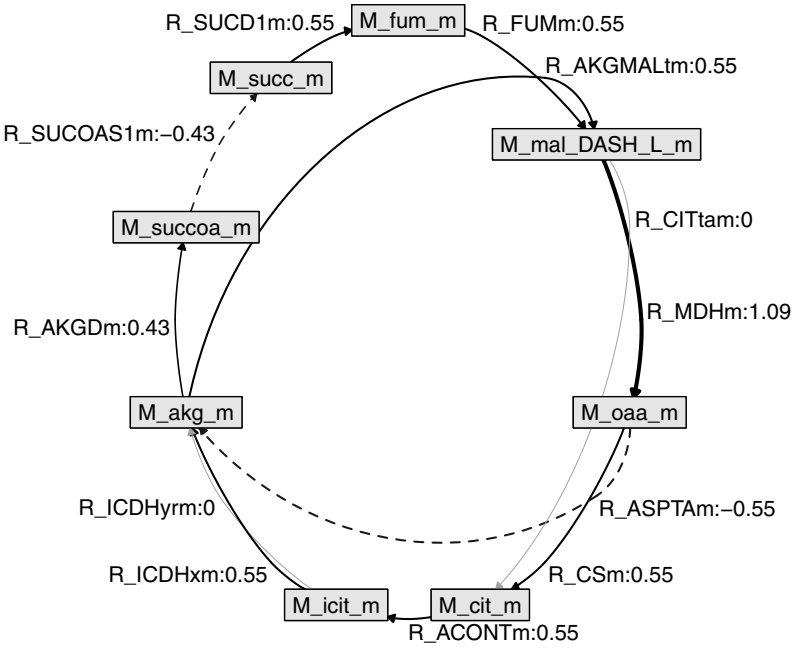

Figure 6.5: Estimated fluxes in the citric acid cycle in the mitochondrion.

The visualisation function `sbml2hyperdraw` is not restricted to FBA models, but `sbml2hyperdraw` can be used as a generic plotting function for SBML models. To this end, in case that no reaction rates are given as argument, all edges are plotted with the same width and in the same colour.

6.6.4 Troubleshooting BiGGR

Model building is an iterative process and requires careful selection of parameters and arguments. Some of the most common problems and solutions are described below:

- **Infeasible solution:** This problem can be encountered when using the `linp` method form the LIM package. This problem occurs when the constraints provided by the user for the model are conflicting. (A trivial example is that a constraint says that a specific flux is greater than 5 units and another constraint says the same flux is smaller than 4. Such conflicts can be much more subtle). The reactions in the model file may sometimes be defined incorrectly, for instance with regard to their reversibility.
- **Visualising too many metabolites and reactions:** If the plotting area is too small to fit all boxes for metabolites, the following error is produced by the hyperdraw package:

```
Error in `[.unit`(pts$x, ref + step) :
Index out of bounds (unit subletting)
```

In case you encounter this error when plotting your model, you can consider several possibilities:

- Increase the size of the plotting area: When plotting to the screen, width and height of the plotting window can be set with the `x11()` command. Type `?x11` for more information. Similarly, figure dimensions can be set when plotting to a jpeg, png, pdf, eps etc. device. Type for instance `?pdf` for the documentation.
- Consider plotting only a subset of the metabolites and reactions in the model. It is possible to pass a list or vector of relevant species and/or relevant reactions to the function `sbml2hyperdraw`. See `?sbml2hyperdraw` for more information.

✧ **Resizing the plotting window:** Resizing the plotting window after plotting a model can cause the edges to get distorted. We advice not to manually resize the plotting window. Instead, if a larger plotting area is desired, the dimensions of the plotting area can be set as described above.

6.7 Author Contributions

Implemented the software: HH and AKG. Designed the model: FS and JHGMvB. Conducted the model simulations: HH, FS, JHGMvB.

Chapter 7

Discussion

Several years of my research in the Systems Biology of energy metabolism have been summarised in this thesis. During this research, approaches in diverse fields were applied, including bioinformatics, biophysics, mathematics and software development. This chapter aims to put the results obtained during this PhD project into perspective and elaborates on the achievements and limitations of the work.

7.1 The importance of mathematical modelling of energy metabolism

Energy is necessary for all living entities to cater for growth, reproduction and function. Biochemical mechanisms for the conversion of energy are therefore essential components of any living organism and the investigation of these mechanisms is crucial for understanding the molecular basis of life. All computational approaches developed and applied in this thesis were used to study energy metabolism. The importance of computational methods in the research field of bioenergetics should not be disregarded since these methods play an important role especially in the quantitative investigation of cellular processes involved in energy metabolism [Cloutier and Wellstead, 2010, Sangar et al., 2012, Schmitz et al., 2011]. Computational models can be used to analyse experimental data and to test hypotheses about the modelled system, thereby facilitating new experimental strategies and designs. To reproduce *in vitro* and *in vivo* biological processes *in silico*, the majority of models rely on experimental data obtained in laboratory experiments. However, the precision of experimental data that can be obtained is subject to noise in a number of ways. Any apparatus or technical device used to extract biological data has limited precision or resolution, and therefore a level of uncertainty must be assigned to all measurements. Variation within populations also accounts for noise in experimental data. For instance, if a certain metabolite or enzyme concentration is measured in two different animals within the same species, the concentrations will most likely not be exactly equal. Furthermore, biological systems tend to be subject to intrinsic noise within the system. Model parameters are components of a computational model which are often not determined precisely and therefore are subject to noise. In order to make model simulations as realistic as possible, noise in model parameters must be taken into account. In order to obtain useful predictions from computational models, their uncertainty with respect to changes in parameter values must be considered. Within this thesis, a major bottleneck in the computational modelling of energy metabolism was approached: the consequences of uncertainty in model parameters describing physical and biological properties of the system.

7.2 Making use of literature information on model parameters - towards viable models of energy metabolism

Often, mathematical models of metabolic processes are subject to uncertainty in their predictions. While model variables, such as a concentration of a certain metabolite, represent biological entities that can change over time, model parameters, for instance enzyme velocities or binding constants, represent biological properties that do not change during a model simulation. The dynamics of model variables are defined by the values of the model parameters. The uncertainty in the model parameters therefore directly affects the uncertainty in model predictions. Small changes in single parameter values can therefore have significant effect on the dynamics of the model variables. In other cases, the model might be over-parametrised and even strong variability in model parameters do not significantly alter the dynamics of the model variables.

A novel aspect in the research of this thesis is the quantitative incorporation of noise on parameter values in the analysis of metabolic models. In this thesis it is argued that a set of model parameters should never be treated as a 'true' set of parameters that correctly reflects biological reality. Models with only one valid and unshiftable set of parameters might introduce a bias in the model predictions by setting *a priori* information represented by the parameters into stone. A spread in model predictions is therefore not taken into account which in some cases can lead to false assumptions on biological processes which are investigated with the model, because possible error is disregarded. In the worst case, a model parameter could be ill-determined and therefore the model would yield unreasonable predictions, which would render the model useless. Also, we believe that if parameters are optimized to experimental data, it is dangerous to merely rely on single estimates. Useful, practical and therefore *viable* models must take any source of uncertainty explicitly into account.

To make models of energy metabolism more viable, we propose that each parameter (e.g. the maximum velocity of the creatine kinase reaction) is treated as a probability distribution which reflects the spread in its probable values. The adjustment of model parameters to experimental data is conducted in a Bayesian fashion. This means that for each model parameter, a posterior distribution of values is sampled that could describe the experimental data [Liebermeister and Klipp, 2005, Brown et al., 2004, Gutenkunst et al., 2007b]. In some cases, values of model parameters are relatively well described in the scientific literature. In Bayesian parameter estimation, it is then feasible to incorporate this prior information on the parameter value and its possible spread. In this thesis, the incorporation of literature information as prior information on parameter values has been done in a quantitative manner. As a result, the uncertainty of predictions made with models of energy metabolism could be narrowed based on this information. In this thesis, we applied this strategy to models

of the creatine kinase system (chapters 3 and 4) and of the citric acid cycle (chapter 5). The approach is thus shown to be versatile and can be applied to very different types of models: models of differential equations with detailed enzyme kinetics, in which parameters represent kinetic properties (chapter 3 and 4) and models for ^{13}C flux estimation (chapter 5) where each parameter represents a reaction flux.

7.3 Computational investigation of creatine kinase function

Creatine kinase catalyses the reversible transfer of high-energy phosphate from ATP to creatine to form phosphocreatine. In this thesis, a simple model of the creatine kinase system was used (chapters 2 to 4). First a naïve parameter sensitivity analysis was conducted on the model (chapter 2). Previous model analysis suggested that energy transport in cardiac muscle is carried out predominantly by ATP and not by phosphocreatine [van Beek, 2007], contradicting what is known as the 'phosphocreatine shuttle hypothesis' [Bessman and Geiger, 1981] which states that phosphocreatine is the chief energy transport molecule in muscle.

By adding moderate random noise on all parameter values, model predictions about the phosphocreatine shuttle hypothesis became very uncertain. However, it was suggested by Gutenkunst et al. that even with a high degree of uncertainty in model parameters, models still can yield useful predictions, if the multidimensional parameter space is explored extensively and the uncertainties in model parameters are known [Gutenkunst et al., 2007b]. The parameter space can be explored by sampling a posterior distribution of parameter sets with the Metropolis-Hastings algorithm [Brown et al., 2004].

Based on the theoretical framework proposed by Brown et al. and Gutenkunst et al., Bayesian parameter inference on the simple model of the creatine kinase system was conducted (chapter 3). The model was calibrated with data from experiments on isolated perfused rabbit hearts, in which the response time of ATP production to a change in energy demand was measured with and without the presence of the functional creatine kinase enzyme. Also, measurements on molecular and organellar biochemical parameters were included into the analysis. Parameter values and their standard errors were extracted from the scientific literature and used as priors in the Bayesian parameter estimation. Parameter ensembles were generated that could describe the experimental response time data relatively well. Using the parameter ensembles, the predicted fraction of energy transported via the phosphocreatine shuttle was relatively low at 15±8%.

When these results were published, the study was harshly criticised by a group of researchers advocating the phosphocreatine shuttle hypothesis [Aliev et al., 2011]. A major assumption for the phosphocreatine shuttle hypothesis is a diffusion restriction for ATP and ADP between mitochondria and myofibrils (see e.g. [Saks et al., 2008]).

In chapter 3, the parameter describing the diffusion of ATP is estimated to be much higher than assumed in other studies [Aliev and Saks, 1997, Vendelin et al., 2000]. In [Aliev et al., 2011], the authors argue that their model also can accurately describe the response time of ATP production in response to increased cardiac workload, even when assuming limited ATP diffusion. However, we found that the models by Saks and colleagues fail to describe a shorter response time in the case that creatine kinase is inhibited, as already shown in our original publication that was criticised ([Hettling and van Beek, 2011] also see Chapter 3.8). A decrease of the response time of cardiac ATP production when creatine kinase is inhibited or absent was measured in different laboratories and different mammalian species or isolated myocytes [Gustafson and Van Beek, 2002, Harrison et al., 1999, Kindig et al., 2005]. In order to refer to the research from chapter 3 as a 'failure' [Aliev et al., 2011], it should be shown by Saks et al., that the response time with impaired creatine kinase activity can be described under the assumption of an obligatory phosphocreatine shuttle. In opposition to the claims of Saks et al., I think that the shuttle hypothesis is on no account a definite fact that cannot not be challenged [Beard and Kushmerick, 2009, Vendelin et al., 2010, Meyer et al., 1984]. Unlike in cardiac tissue, the function of creatine kinase in skeletal muscle tissue is not as hotly disputed. In chapter 4, the model used in chapters 2 and 3 is modified and adapted to match experimental data from human skeletal muscle. The model was calibrated with measurements of metabolite concentrations and muscle acidity in human subjects after a cycling exercise bout. Model simulations showed that creatine kinase is essentially responsible for the damping of high peaks of energy demand in the muscle tissue. This is an important aspect of energy housekeeping, but does not make the phosphocreatine shuttle the predominant vehicle of energy transport.

7.4 ^{13}C metabolic flux analysis and parameter uncertainty

To understand the dynamics of energy metabolism *in vivo*, quantitative knowledge on the velocity of biochemical reactions is of major importance. In chapter 5, isotope labelling data were analysed using a computational model in order to quantify the fluxes in the citric acid cycle in porcine cardiac muscle. The data consisted of NMR measurements in glutamate in 347 tissue samples. The computational model contains detailed descriptions of the carbon transitions within reactions of the pathway. These relationships between the metabolite pools in the model enabled the calculation of the network fluxes from the measurement of a single metabolite. Two out of five flux parameters of interest could be reasonably well estimated. For the remaining three parameters, no reliable point estimates could be made (i) because of the high measurement noise in the NMR data, (ii) because measurements were performed at exactly one point in time to enable the study of heterogeneity between different biopsies [van Beek et al., 1999] and (iii) because only one intermediate metabolite,

glutamate, could be measured with NMR. Because of the high noise level in the data, it would be dangerous to rely on mere point estimates obtained from fitting the flux parameters. The uncertainty of the estimated fluxes was therefore investigated by performing a random walk through parameter space (similar to the methodology in chapter 3). Samples for which the spread in the estimated parameters was too high were rejected from the analysis. Although sophisticated bioinformatics methods were used to quantify the flux parameters, the method is limited by the quality of the data used.

A possible improvement of the procedure described in chapter 5 would be measurements of multiple intermediate metabolites in addition to glutamate. Also NMR spectroscopy is a popular analytical method in metabolomics but has limited sensitivity when compared to mass spectrometry (MS) [Pan and Raftery, 2007] and it would therefore be desirable to obtain MS measurements for the tissue samples. Maier et al., for instance, measured ^{13}C citric acid cycle intermediates such as citrate and alpha-ketoglutarate with gas-chromatography-mass spectrometry (GC-MS) [Maier et al., 2008]. It would therefore be feasible to obtain more accurate data for the tissue biopsies. In order to test if this data would improve the performance of the method, the following simulation study was implemented: One sample of artificial MS data was generated by simulating the model used in chapter 5 with the best fit parameters from Figure 5.2. All mass isotopomers of the metabolites citrate, α-ketoglutarate, oxaloacetate and glutamate were considered, resulting in 18 data points for the sample (in comparison: 9 data points could be obtained per sample with NMR). Gaussian random noise was added to the simulated mass isotopomers according to the standard deviations found in [Maier et al., 2008], which ranged from 1 to 71% but was assumed to be at least 10%.

Figure 7.1 displays the results of the parameter estimation procedure with Markov-chain Monte Carlo described in chapter 5 for the NMR data (top row) and the artificial MS data (bottom row). Especially the parameter J_{exch} becomes identifiable in the analysis with artificial MS data, since both α-ketoglutarate and glutamate pools are part of the data. It is therefore concluded that the accuracy of the methodology of ^{13}C flux analysis with the LIPSSS method (chapter 5) could be significantly improved if MS measurements for more metabolites than merely glutamate were available.

7.5 A note on doping abuse and bicycle racing

In the first part of chapter 4, a computational model is presented to simulate heat transfer in a bicyclist climbing the Alpe d'Huez during a time trial stage in the Tour de France 2004. As an input for the model, physiological data from Lance Armstrong was used which was measured over a period of time by Coyle [Coyle, 2005]. An exceptionally high oxygen uptake of the cyclist during the time trial was estimated. Additional to the calculations with the whole body model, a different model was used to describe the dynamics of mitochondrial ATP production during bicycle racing on

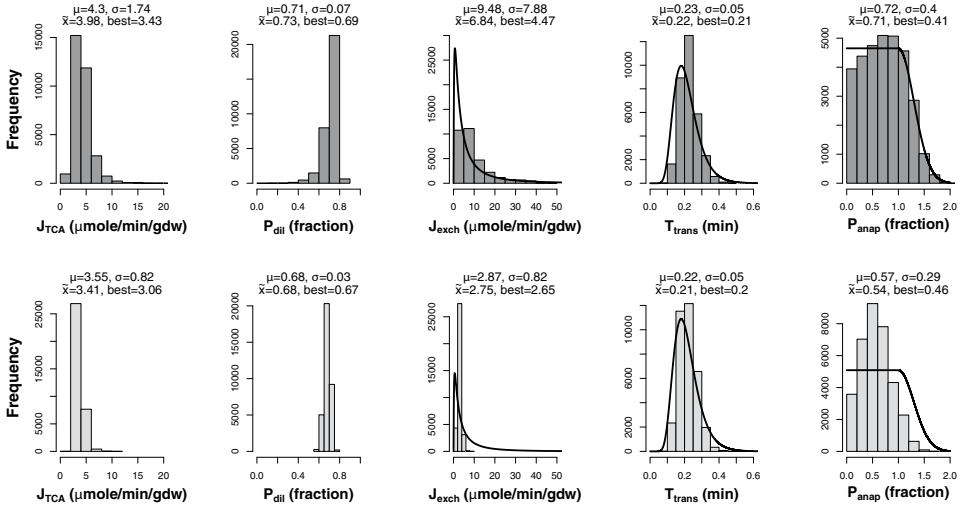

Figure 7.1: ^{13}C *metabolic flux estimation for one tissue sample analysed with the method described in chapter 5. The top row shows the analysis with real NMR data and the incorporation of prior information. In the bottom row, the same analysis was conducted with artificial MS data and the same prior information. The artificial data has been generated by simulating the model with the set of best fitting parameters from the analysis in the top row. The probability density functions of the priors for the auxiliary parameters* J_{exch}, T_{trans} *and* P_{anap} *are plotted with solid lines. On top of each plot, ensemble mean, standard deviation, median* $\tilde{x}$ *and best fit value are reported. Note that the top row is the same as Figure 5.2.*

a smaller scale.

On August 24th 2012, the United States Anti-Doping Agency (USADA) announced a life-time sports competition ban for Lance Armstrong and the revoke of all his seven Tour de France titles for the illegal usage of performance enhancing drugs. Already two years before that, at the time when the article which constitutes the research in chapter 4 was being reviewed, the question was raised whether the models used in this study could be used to investigate possible substance abuse in bicycle racing. Chief evidence against Lance Armstrong, besides the testimony of his teammates, was the detection of erythropoetin (EPO) in blood and urine samples. EPO is a peptide hormone which stimulates the production of erythrocytes (red blood cells), increasing the body's capacity of oxygen uptake to boost oxidative energy production by the mitochondria. The published data for Lance Armstrong [Coyle, 2005] used for the whole body model and the data used to calibrate the model on the subcellular level [Sahlin et al., 1987] are mainly measurements of aerobic capacity, which, among others, depends on oxygen transport in the blood. In order to investigate the effect of EPO or blood doping, the model would have to be extended to take these other processes into account. However, even if the model was extended it would be very difficult to discern the effects of erythrocyte enhancing doping methods to for instance

the effect of training in high altitudes. Besides EPO and blood doping, prohibited substances that regulate the cyclists' temperature were also mentioned by his former soigneur Emma O'Reilly during the testimony[1]. The use of such substances may have affected the regulatory processes modelled in chapter 5 and invalidated the computational predictions. An interesting future application of the whole body model would be the testing of the effect of temperature regulating drugs *in silico*.

7.6 The metabolic modelling software tools BiGGR and FluxEs

Computational analysis of metabolic processes requires sophisticated software to simulate, optimise and visualise computational models. For some calculations conducted during the research for this thesis, existing software tools and libraries could be used (e.g. [Gutenkunst et al., 2007a, Binsl et al., 2010a]). However, the application and extension of different methods conducted in this thesis often required functional extension of the existing tools. In the optimal case, this results in the development of new software which can be provided to the scientific community. For the compilation and the simulation of the model of brain metabolism (chapter 4) and similar models, the software package BiGGR was developed (chapter 6). BiGGR is a software application which allows for model generation, model visualisation and flux estimation via Flux Balance Analysis. Furthermore, BiGGR provides an interface to the metabolic reconstruction databases BiGG and Recon [Schellenberger et al., 2010, Thiele et al., 2013]. An example of a flux balance analysis using BiGGR is given in chapter 6.6. To this end, a computational model of brain energy metabolism [van Beek et al., 2011] was used to estimate fluxes in the glycolysis pathway and to visualise the results. The work in this thesis also comprised the development of software for flux analysis with stable isotopes (chapter 5). To this end, the FluxEs software developed by Binsl et al. [Binsl et al., 2010a] was rewritten and extended to facilitate the sampling of parameter ensembles with Monte-Carlo Methods. FluxEs supports an easy to use file format for encoding carbon transition networks. The software automatically generates the differential equations of a model and facilitates model simulation and flux parameter estimation. Although FluxEs was developed to analyse LIPSSS data taken at one 'snapshot' point in time (see the analysis of in vivo tissue biopsies in chapter 5), it is easily possible to include time-series data into the analysis. Theoretically, other tracer experiments which rely for instance on the isotope ^{18}O, could be simulated using FluxEs. Both FluxEs and BiGGR are written in the R programming language and are therefore open source, free of charge and platform independent. By supporting the standard model exchange format SBML, interoperability with different software packages is possible. BiGGR is publicly available in the BioConductor software repository [Gentleman et al., 2004]. In the future, a new improved version of FluxEs will be made available for the public.

[1] The testimony is available at `http://cyclinginvestigation.usada.org`.

7.7 Conclusions

The outcome of the project of this thesis is manifold. One part of the work was the development and modification of theoretical methods for modelling in Systems Biology. Main emphasis was put on developing strategies to take into account literature information on model parameters. Another part was the application of these methods to biological processes using existing models and biological data. To this end, (i) the biological functions of the creatine kinase enzyme in muscle (chapter 3 and 4) and (ii) the effect of various interventions in heart muscle in vivo on reactions in the citric acid cycle (chapter 5) was investigated. Validations of the methods used showed that the findings obtained have biological significance. Also, basic modelling was part of the work in this thesis. The creatine kinase model was extended in order to describe experimental data on human subjects in endurance exercise (chapter 4). Finally, in the process of this work, two software tools were developed that facilitate the analysis of computational models. The results in this thesis thus contribute to new improvements in the research of energy metabolism using a Systems Biology approach.

Bibliography

[Alders et al., 2004] D. J. C. Alders, A. B. J. Groeneveld, F. J. J. de Kanter, and J. H. G. M. van Beek. Myocardial O2 consumption in porcine left ventricle is heterogeneously distributed in parallel to heterogeneous O2 delivery. *American journal of Physiology - Heart and Circulatory Physiology*, 287(3):H1353–61, Sept. 2004.

[Alders et al., 2007] D. J. C. Alders, R. N. Cornelussen, F. W. Prinzen, P. A. C. Specht, M. I. M. Noble, A. J. Drake-Holland, F. J. J. de Kanter, and J. H. G. M. van Beek. Regional sympathetic denervation affects the relation between canine local myocardial blood flow and oxygen consumption. *Experimental physiology*, 92(3): 541–8, May 2007.

[Alders et al., 2011] D. J. C. Alders, A. B. J. Groeneveld, T. W. Binsl, F. J. de Kanter, and J. H. G. M. van Beek. Endotoxemia decreases matching of regional blood flow and O2 delivery to O2 uptake in the porcine left ventricle. *American Journal of Physiology - Heart and Circulatory Physiology*, 300(4):H1459—-66, Apr. 2011.

[Aliev et al., 2011] M. Aliev, R. Guzun, M. Karu-Varikmaa, T. Kaambre, T. Wallimann, and V. Saks. Molecular system bioenergics of the heart: experimental studies of metabolic compartmentation and energy fluxes versus computer modeling. *International journal of molecular sciences*, 12(12):9296–331, Jan. 2011.

[Aliev and Saks, 1997] M. K. Aliev and V. A. Saks. Compartmentalized Energy Transfer in Cardiomyocytes : Use of Mathematical Modeling for Analysis of In Vivo Regulation of Respiration Mitochondrion Cytolas. *Biophysical Journal*, 73(July):428–445, 1997.

[Antoniewicz et al., 2006] M. R. Antoniewicz, J. K. Kelleher, and G. Stephanopoulos. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metabolic engineering*, 8(4):324–337, July 2006.

[Antoniewicz et al., 2007] M. R. Antoniewicz, J. K. Kelleher, and G. Stephanopoulos. Elementary metabolite units (emu): a novel framework for modeling isotopic distributions. *Metabolic Engineering*, 9(1):68–86, Jan. 2007.

[Beard, 2005] D. A. Beard. A biophysical model of the mitochondrial respiratory system and oxidative phosphorylation. *PLoS computational biology*, 1(4):e36, 2005.

[Beard, 2006] D. A. Beard. Modeling of oxygen transport and cellular energetics explains observations on in vivo cardiac energy metabolism. *PLoS computational biology*, 2 (9):e107, Sept. 2006.

[Beard and Kushmerick, 2009] D. A. Beard and M. J. Kushmerick. Strong inference for systems biology. *PLoS computational biology*, 5(8):e1000459, Aug. 2009.

[Bergström, 1967] J. Bergström. Local changes of ATP and phosphorylcreatine in human muscle tissue in connection with exercise. *Circulation Research*, 21(Supp.1):91–98, 1967.

[Bessman and Geiger, 1981] S. P. Bessman and P. J. Geiger. Transport of energy in muscle: the phosphorylcreatine shuttle. *Science*, 211(4481):448–52, Jan. 1981.

[Binsl et al., 2010a] T. W. Binsl, A. A. De Graaf, K. Venema, J. Heringa, A. Maathuis, P. De Waard, and J. H. Van Beek. Measuring non-steady-state metabolic fluxes in starch-converting faecal microbiota in vitro. *Beneficial Microbes*, 1(4):391–405, Nov. 2010a.

[Binsl et al., 2010b] T. W. Binsl, M. K.M, van Stokkum I.H.M, J. Heringa, and J. van Beek. FLUXSIMULATOR: an R package to simulate isotopomer distributions in metabolic networks. *Journal of Statistical Software*, 18:1–17, 2010b.

[Brewer and Wallimann, 2000] G. J. Brewer and T. W. Wallimann. Protective effect of the energy precursor creatine against toxicity of glutamate and beta-amyloid in rat hippocampal neurons. *Journal of Neurochemistry*, 74(5):1968–78, May 2000.

[Brown and Sethna, 2003] K. S. Brown and J. P. Sethna. Statistical mechanical approaches to models with many poorly known parameters. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 68(2 Pt 1):21904, 2003.

[Brown et al., 2004] K. S. Brown, C. C. Hill, G. A. Calero, C. R. Myers, K. H. Lee, J. P. Sethna, and R. A. Cerione. The statistical mechanics of complex signaling networks: nerve growth factor signaling. *Physical Biology*, 1(3-4):184–195, 2004.

[Casey et al., 2007] F. P. Casey, D. Baird, Q. Feng, R. N. Gutenkunst, J. J. Waterfall, C. R. Myers, K. S. Brown, R. A. Cerione, and J. P. Sethna. Optimal experimental design in an epidermal growth factor receptor signalling and down-regulation model. *IET Systems Biology*, 1(3):190–202, 2007.

[Cerami et al., 2011] E. G. Cerami, B. E. Gross, E. Demir, I. Rodchenkov, O. Babur, N. Anwar, N. Schultz, G. D. Bader, and C. Sander. Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Research*, 39(Database issue):D685–90, Jan. 2011.

[Chance, 1965] B. Chance. The energy-linked reaction of calcium with mitochondria. *The Journal of Biological Chemistry*, 240:2729–48, June 1965.

[Chance et al., 1983] E. M. Chance, S. H. Seeholzer, K. Kobayashi, and J. R. Williamson. Mathematical analysis of isotope labeling in the citric acid cycle with applications to 13C NMR studies in perfused rat hearts. *The Journal of Biological Chemistry*, 258(22):13785–94, Nov. 1983.

[Chatham et al., 1995] J. C. Chatham, J. R. Forder, J. D. Glickson, and E. M. Chance. Calculation of absolute metabolic flux and the elucidation of the pathways of glutamate labeling in perfused rat heart by 13C NMR spectroscopy and nonlinear least squares analysis. *The Journal of Biological Chemistry*, 270(14):7999–8008, Apr. 1995.

[Chouikha and Schnieder, 1998] M. Chouikha and E. Schnieder. Modelling of Continuous-discrete Systems with Hybrid Petri Nets. In *In IEEE-MC Multiconference on Computational Engeneering in Systems Applications*, pages 606–612, 1998.

[Cloutier and Wellstead, 2010] M. Cloutier and P. Wellstead. The control systems structures of energy metabolism. *Journal of the Royal Society, Interface / the Royal Society*, 7(45):651–65, Apr. 2010.

[Cohen and Bergman, 1997] D. M. Cohen and R. N. Bergman. Improved estimation of anaplerosis in heart using 13C NMR. *The American Journal of Physiology - Endocrinology and Metabolism*, 273(6):E1228—-1242, Dec. 1997.

[Coyle, 2005] E. F. Coyle. Improved muscular efficiency displayed as Tour de France champion matures. *Journal of applied physiology (Bethesda, Md. : 1985)*, 98(6):2191–6, June 2005.

[Croft et al., 2011] D. Croft, G. O'Kelly, G. Wu, R. Haw, M. Gillespie, L. Matthews, M. Caudy, P. Garapati, G. Gopinath, B. Jassal, S. Jupe, I. Kalatskaya, S. Mahajan, B. May, N. Ndegwa, E. Schmidt, V. Shamovsky, C. Yung, E. Birney, H. Hermjakob, P. D'Eustachio, and L. Stein. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Research*, 39(Database issue):D691–7, Jan. 2011.

[Dalsgaard, 2006] M. K. Dalsgaard. Fuelling cerebral activity in exercising man. *Journal of cerebral blood flow and metabolism*, 26(6):731–50, June 2006.

[Dalsgaard et al., 2002] M. K. Dalsgaard, K. Ide, Y. Cai, B. r. Quistorff, and N. H. Secher. The intent to exercise influences the cerebral O(2)/carbohydrate uptake ratio in humans. *The Journal of physiology*, 540(Pt 2):681–9, Apr. 2002.

[Dalsgaard et al., 2004] M. K. Dalsgaard, B. r. Quistorff, E. R. Danielsen, C. Selmer, T. Vogelsang, and N. H. Secher. A reduced cerebral metabolic ratio in exercise reflects metabolism and not accumulation of lactate within the human brain. *The Journal of physiology*, 554(Pt 2):571–8, Jan. 2004.

[de Groot, 1999] B. de Groot. The role of mitochondria and intracellular energy transfer in the pathogenesis of heart failure. *Ph.D thesis, VU University Amsterdam*, 1999.

[den Meersche et al., 2009] K. V. den Meersche, K. Soetaert, and D. V. Oevelen. `xsample()`: An R function for sampling linear inverse problems. *Journal of Statistical Software*, 30(1):1–15, 4 2009.

[Duarte et al., 2007] N. C. Duarte, S. A. Becker, N. Jamshidi, I. Thiele, M. L. Mo, T. D. Vo, R. Srivas, and B. O. Palsson. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences of the United States of America*, 104(6):1777–82, Feb. 2007.

[Dusick et al., 2007] J. R. Dusick, T. C. Glenn, W. N. Lee, P. M. Vespa, D. F. Kelly, S. M. Lee, D. A. Hovda, and N. A. Martin. Increased pentose phosphate pathway flux after clinical traumatic brain injury: a [1,2-13C2] glucose labeling study in humans. *Journal of cerebral blood flow and metabolism*, 27(9):1593–1602, Sep 2007.

[Erickson-Viitanen et al., 1982] S. Erickson-Viitanen, P. Viitanen, P. J. Geiger, W. C. Yang, and S. P. Bessman. Compartmentation of mitochondrial creatine phosphokinase. I. Direct demonstration of compartmentation with the use of labeled precursors. *The Journal of Biological Chemistry*, 257(23):14395–404, Dec. 1982.

[Ferretti et al., 1997] G. Ferretti, C. Moia, J. M. Thomet, and B. Kayser. The decrease of maximal oxygen consumption during hypoxia in man: a mirror image of the oxygen equilibrium curve. *The Journal of physiology*, 498:231–7, Jan. 1997.

[Finley et al., 2007] A. O. Finley, S. Banerjee, and B. P. Carlin. spBayes: An R Package for Univariate and Multivariate Hierarchical Point-referenced Spatial Models. *Journal of Statistical Software*, 19(4):1–24, 2007.

[Funahashi et al., 2003] A. Funahashi, M. Morohashi, H. Kitano, and N. Tanimura. CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *Biosilico*, 1(5):159–162, 2003.

[Gangadharan and Rohatgi, 2012] A. Gangadharan and N. Rohatgi. abcdefba: A-biologist-can-do-everything of flux balance analysis with this package. *R package version 0.4. Available at* `http://www.cran.r-project.org/`, 2012.

[Gelius-Dietrich et al., 2013] G. Gelius-Dietrich, A. Amer Desouki, C. J. Fritzemeier, and M. J. Lercher. sybil : Efficient constraint-based modelling in R. *BMC Systems Biology*, 7(1):125, Nov 2013.

[Gentleman et al., 2004] R. C. Gentleman, V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dumydoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. H. Yang, and J. Zhang. Bioconductor: open software development for computational biology and bioinformatics. *Genome biology*, 5(10):R80, Jan. 2004.

[Gnaiger, 2009] E. Gnaiger. Capacity of oxidative phosphorylation in human skeletal muscle: new perspectives of mitochondrial physiology. *The international journal of biochemistry & cell biology*, 41(10):1837–45, Oct. 2009.

[Gokulakrishnan et al., 2006] P. Gokulakrishnan, A. D. Lawrence, P. J. McLellan, and E. W. Grandmaison. A functional-PCA approach for analyzing and reducing complex chemical mechanisms. *Computers & Chemical Engineering*, 30(6-7):1093–1101, 2006.

[Greenhaff, 2001] P. L. Greenhaff. The creatine-phosphocreatine system: there's more than one song in its repertoire. *The Journal of Physiology*, 537(Pt 3):657, 2001.

[Gustafson and Van Beek, 2002] L. A. Gustafson and J. H. Van Beek. Activation time of myocardial oxidative phosphorylation in creatine kinase and adenylate kinase knockout mice. *American Journal of Physiology - Heart and Circulatory Physiology*, 282(6):H2259–64, 2002.

[Gutenkunst et al., 2007a] R. N. Gutenkunst, J. C. Atlas, F. P. Casey, R. S. Kuczenski, J. J. Waterfall, C. R. Meyers, and J. P. Sethna. SloppyCell. See `http://sloppycell.sourceforge.net/`, 2007a.

[Gutenkunst et al., 2007b] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS computational biology*, 3(10):1871–1878, 2007b.

[Harrison et al., 1999] G. J. Harrison, M. H. van Wijhe, B. de Groot, F. J. Dijk, and J. H. van Beek. CK inhibition accelerates transcytosolic energy signaling during rapid workload steps in isolated rabbit hearts. *American Journal of Physiology*, 276(1 Pt 2):H134–H140, 1999.

[Harrison et al., 2003] G. J. Harrison, M. H. van Wijhe, B. de Groot, F. J. Dijk, L. A. Gustafson, and J. H. G. M. van Beek. Glycolytic buffering affects cardiac bioenergetic signaling and contractile reserve similar to creatine kinase. *American Journal of Physiology - Heart and Circulatory Physiology*, 285(2):H883–H890, 2003.

[Heineman and Balaban, 1990] F. W. Heineman and R. S. Balaban. Phosphorus-31 nuclear magnetic resonance analysis of transient changes of canine myocardial metabolism in vivo. *The Journal of Clinical Investigation*, 85(3):843–852, 1990.

[Hersch et al., 2006] S. M. Hersch, S. Gevorkian, K. Marder, C. Moskowitz, A. Feigin, M. Cox, P. Como, C. Zimmerman, M. Lin, L. Zhang, A. M. Ulug, M. F. Beal, W. Matson, M. Bogdanov, E. Ebbel, A. Zaleta, Y. Kaneko, B. Jenkins, N. Hevelone, H. Zhang, H. Yu, D. Schoenfeld, R. Ferrante, and H. D. Rosas. Creatine in Huntington disease is safe, tolerable, bioavailable in brain and reduces serum 8OH2'dG. *Neurology*, 66(2):250–2, Jan. 2006.

[Hettling and van Beek, 2011] H. Hettling and J. H. G. M. van Beek. Analyzing the Functional Properties of the Creatine Kinase System with Multiscale 'Sloppy' Modeling. *PLoS Computational Biology*, 7(8):e1002130, 2011.

[Hindmarsh, 1983] A. Hindmarsh. ODEPACK, A Systematized Collection of ODE Solvers , R. S. Stepleman et al. (eds.), North-Holland, Amsterdam, (vol. 1 of), pp. 55-64. *IMACS Transactions on Scientific Computation*, 1:55 – 64, 1983.

[Hozo et al., 2005] S. P. Hozo, B. Djulbegovic, and I. Hozo. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology*, 5:13, 2005.

[Ivy et al., 2002] J. L. Ivy, H. W. Goforth, B. M. Damon, T. R. McCauley, E. C. Parsons, and T. B. Price. Early postexercise muscle glycogen recovery is enhanced with a carbohydrate-protein supplement. *Journal of applied physiology (Bethesda, Md. : 1985)*, 93(4):1337–44, Oct. 2002.

[Jacobus and Saks, 1982] W. E. Jacobus and V. A. Saks. Creatine kinase of heart mitochondria: changes in its kinetic properties induced by coupling to oxidative phosphorylation. *Archives of Biochemistry and Biophysics*, 219(1):167–178, 1982.

[Jacobus et al., 1982] W. E. Jacobus, R. W. Moreadith, and K. M. Vandegaer. Mitochondrial respiratory control. Evidence against the regulation of respiration by extramitochondrial phosphorylation potentials or by [ATP]/[ADP] ratios. *The Journal of Biological Chemistry*, 257(5):2397–2402, 1982.

[Jeffrey et al., 1999] F. M. H. Jeffrey, A. Reshetov, C. J. Storey, R. A. Carvalho, A. D. Sherry, and C. R. Malloy. Use of a single 13C NMR resonance of glutamate for measuring oxygen consumption in tissue. *American journal of physiology - Endocrinology and metabolism*, 277(6):E1111–1121, Dec. 1999.

[Jensen, 1997] K. Jensen. A brief introduction to coloured Petri nets. In *Lecture Notes in Computer Science*, pages 55–58. Springer-Verlag, 1997.

[Jeukendrup et al., 2000] A. E. Jeukendrup, N. P. Craig, and J. A. Hawley. The bioenergetics of World Class Cycling. *Journal of science and medicine in sport / Sports Medicine Australia*, 3(4):414–33, Dec. 2000.

[Joubert et al., 2002] F. Joubert, J.-L. Mazet, P. Mateo, and J. A. Hoerter. 31P NMR detection of subcellular creatine kinase fluxes in the perfused rat heart: contractility

modifies energy transfer pathways. *The Journal of Biological Chemistry*, 277(21): 18469–18476, 2002.

[Joubert et al., 2004] F. Joubert, P. Mateo, B. Gillet, J. C. Beloeil, J. L. Mazet, and J. A. Hoerter. CK flux or direct ATP transfer: versatility of energy transfer pathways evidenced by NMR in the perfused heart. *Molecular and Cellular Biochemistry*, 256-257(1-2):43–58, 2004.

[Joyner and Coyle, 2008] M. J. Joyner and E. F. Coyle. Endurance exercise performance: the physiology of champions. *The Journal of physiology*, 586(1):35–44, Jan. 2008.

[Kadirkamanathan et al., 2006] V. Kadirkamanathan, J. Yang, S. A. Billings, and P. C. Wright. Markov Chain Monte Carlo Algorithm based metabolic flux distribution analysis on Corynebacterium glutamicum. *Bioinformatics (Oxford, England)*, 22 (21):2681–7, Nov. 2006.

[Kanehisa et al., 2012] M. Kanehisa, S. Goto, Y. Sato, M. Furumichi, and M. Tanabe. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research*, 40(Database issue):D109–14, Jan. 2012.

[Kasumov et al., 2007] T. Kasumov, A. V. Cendrowski, F. David, K. A. Jobbins, V. E. Anderson, and H. Brunengraber. Mass isotopomer study of anaplerosis from propionate in the perfused rat heart. *Archives of biochemistry and biophysics*, 463(1): 110–117, July 2007.

[Kelleher, 2001] J. K. Kelleher. Flux estimation using isotopic tracers: common ground for metabolic physiology and metabolic engineering. *Metabolic Engineering*, 3(2): 100–110, Apr. 2001.

[Kindig et al., 2005] C. A. Kindig, R. A. Howlett, C. M. Stary, B. Walsh, and M. C. Hogan. Effects of acute creatine kinase inhibition on metabolism and tension development in isolated single myocytes. *Journal of applied physiology (Bethesda, Md. : 1985)*, 98(2):541–9, Feb. 2005.

[Kitano, 2002a] H. Kitano. Systems biology: a brief overview. *Science*, 295(5560):1662–4, Mar. 2002a.

[Kitano, 2002b] H. Kitano. Computational systems biology. *Nature*, 420(6912):206–10, Nov. 2002b.

[Klamt et al., 2009] S. Klamt, U.-U. Haus, and F. Theis. Hypergraphs and cellular networks. *PLoS computational biology*, 5(5):e1000385, May 2009.

[Klivenyi et al., 1999] P. Klivenyi, R. J. Ferrante, R. T. Matthews, M. B. Bogdanov, A. M. Klein, O. A. Andreassen, G. Mueller, M. Wermer, R. Kaddurah-Daouk, and M. F. Beal. Neuroprotective effects of creatine in a transgenic animal model of amyotrophic lateral sclerosis. *Nature Medicine*, 5(3):347–50, Mar. 1999.

[Kofoed et al., 2000] K. Kofoed, P. Hansen, S. Holm, J. Hove, K. Chen, W. Jin, M. Jensen, H. Iida, B. Hesse, and J. Svendsen. Regional myocardial oxygen consumption estimated by carbon-11 acetate and positron emission tomography before and after repetitive ischemia. *Journal of Nuclear Cardiology*, 7(3):228–234, May 2000.

[Kohl et al., 2010] P. Kohl, E. J. Crampin, T. A. Quinn, and D. Noble. Systems biology: an approach. *Clinical pharmacology and therapeutics*, 88(1):25–33, July 2010.

[Kongas and van Beek, 2007] O. Kongas and J. H. van Beek. Creatine kinase in energy metabolic signaling in muscle. *Nature Precedings*, Nov. 2007.

[Korshunov et al., 1997] S. S. Korshunov, V. P. Skulachev, and A. A. Starkov. High protonic potential actuates a mechanism of production of reactive oxygen species in mitochondria. *FEBS Letters*, 416(1):15–8, Oct. 1997.

[Lakshmanan et al., 2012] M. Lakshmanan, G. Koh, B. K. S. Chung, and D.-Y. Lee. Software applications for flux balance analysis. *Briefings in bioinformatics*, Nov. 2012.

[Lawrence, 2013] M. Lawrence. rsbml: R support for sbml, using libsbml. *R package version 2.20.1. Available at* `http://www.bioconductor.org/`, 2013.

[Lawson and Veech, 1979] J. W. Lawson and R. L. Veech. Effects of pH and free Mg2+ on the Keq of the creatine kinase reaction and other phosphate hydrolyses and phosphate transfer reactions. *The Journal of Biological Chemistry*, 254(14):6528–37, July 1979.

[Le Novere et al., 2006] N. Le Novere, B. Bornstein, A. Broicher, M. Courtot, M. Donizelli, H. Dharuri, L. Li, H. Sauro, M. Schilstra, B. Shapiro, J. L. Snoep, and M. Hucka. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Research*, 34 (Database issue):D689–91, 2006.

[Lee et al., 1994] A. C. Lee, M. Zizi, and M. Colombini. Beta-NADH decreases the permeability of the mitochondrial outer membrane to ADP by a factor of 6. *The Journal of Biological Chemistry*, 269(49):30974–30980, 1994.

[Lieber, 2002] R. R. Lieber. *Skeletal Muscle Structure, Function, and Plasticity: The Physiological Basis of Rehabilitation*. Lippincott Williams & Wilkins; Second edition, Baltimore, 2002. ISBN 0781730619.

[Liebermeister and Klipp, 2005] W. Liebermeister and E. Klipp. Biochemical networks with uncertain parameters. *Systems Biology (Stevenage)*, 152(3):97–107, 2005.

[Lipskaya and Savchenko, 2003] T. Y. Lipskaya and M. S. Savchenko. Once again about the functional coupling between mitochondrial creatine kinase and adenine nucleotide translocase. *Biochemistry (Moscow)*, 68(1):68–79, Jan. 2003.

[Lloyd et al., 2008] C. M. Lloyd, J. R. Lawson, P. J. Hunter, and P. F. Nielsen. The CellML Model Repository. *Bioinformatics*, 24(18):2122–2123, 2008.

[Lloyd et al., 2004] S. G. Lloyd, P. Wang, H. Zeng, and J. C. Chatham. Impact of low-flow ischemia on substrate oxidation and glycolysis in the isolated perfused rat heart. *American Journal of Physiology - Heart and Circulatory Physiology*, 287(1): H351—-62, July 2004.

[Lying-Tunell et al., 1980] U. Lying-Tunell, B. S. Lindblad, H. O. Malmlund, and B. Persson. Cerebral blood flow and metabolic rate of oxygen, glucose, lactate, pyruvate, ketone bodies and amino acids. *Acta neurologica Scandinavica*, 62(5): 265–75, Nov. 1980.

[Ma et al., 2007] H. Ma, A. Sorokin, A. Mazein, A. Selkov, E. Selkov, O. Demin, and I. Goryanin. The edinburgh human metabolic network reconstruction and its functional analysis. *Molecular Systems Biology*, 3:135, 2007.

[Maier et al., 2008] K. Maier, U. Hofmann, M. Reuss, and K. Mauch. Identification of metabolic fluxes in hepatic cells from transient 13C-labeling experiments: Part II. Flux estimation. *Biotechnology and bioengineering*, 100(2):355–70, June 2008.

[Malloy et al., 1990] C. R. Malloy, A. D. Sherry, and F. M. Jeffrey. Analysis of tricarboxylic acid cycle of the heart using 13C isotope isomers. *American Journal of Physiology - Heart and Circulatory Physiology*, 259(3):H987—-995, Sept. 1990.

[Malloy et al., 1996] C. R. Malloy, J. G. Jones, F. M. Jeffrey, M. E. Jessen, and A. D. Sherry. Contribution of various substrates to total citric acid cycle flux and anaplerosis as determined by13C isotopomer analysis and O2 consumption in the heart. *MAGMA Magnetic Resonance Materials in Physics, Biology, and Medicine*, 4(1):35–46, Mar. 1996.

[Martini et al., 2003] W. Z. Martini, W. C. Stanley, H. Huang, C. D. Rosiers, C. L. Hoppel, and H. Brunengraber. Quantitative assessment of anaplerosis from propionate in pig heart in vivo. *American journal of physiology - Endocrinology and metabolism*, 284(2):E351–6, Feb. 2003.

[Mason et al., 1995] G. F. Mason, R. Gruetter, D. L. Rothman, K. L. Behar, R. G. Shulman, and E. J. Novotny. Simultaneous determination of the rates of the TCA cycle, glucose utilization, alpha-ketoglutarate/glutamate exchange, and glutamine synthesis in human brain by NMR. *Journal of cerebral blood flow and metabolism*, 15(1):12–25, Jan. 1995.

[Matthews et al., 1998] R. T. Matthews, L. Yang, B. G. Jenkins, R. J. Ferrante, B. R. Rosen, R. Kaddurah-Daouk, and M. F. Beal. Neuroprotective Effects of Creatine and Cyclocreatine in Animal Models of Huntington's Disease. *The Journal of Neuroscience*, 18(1):156–163, 1998.

[McFarland et al., 1994] E. W. McFarland, M. J. Kushmerick, and T. S. Moerland. Activity of creatine kinase in a contracting mammalian muscle of uniform fiber type. *Biophysical journal*, 67(5):1912–1924, Nov. 1994.

[Meyer et al., 2006] L. E. Meyer, L. B. Machado, A. P. Santiago, W. S. Da-Silva, F. G. De Felice, O. Holub, M. F. Oliveira, and A. Galina. Mitochondrial creatine kinase activity prevents reactive oxygen species generation: antioxidant role of mitochondrial kinase-dependent ADP re-cycling activity. *The Journal of Biological Chemistry*, 281(49):37361–37371, 2006.

[Meyer et al., 1984] R. A. Meyer, H. L. Sweeney, and M. J. Kushmerick. A simple analysis of the "phosphocreatine shuttle". *The American journal of physiology*, 246(5 Pt 1): C365–77, May 1984.

[Murata, 1989] T. Murata. Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, 77(4):541–580, Apr. 1989.

[Murrell, 2013] P. Murrell. hyperdraw: Visualizing hypergaphs. *R package version 1.14.0. Available at* `http://www.bioconductor.org/`, 2013.

[Nelder and Mead, 1965] J. A. Nelder and R. Mead. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, Jan. 1965.

[Nuutinen et al., 1981] E. M. Nuutinen, K. J. Peuhkurinen, E. P. Pietiläinen, J. K. Hiltunen, and I. E. Hassinen. Elimination and replenishment of tricarboxylic acid-cycle intermediates in myocardium. *The Biochemical journal*, 194(3):867–875, Mar. 1981.

[Nybo et al., 2002] L. Nybo, N. H. Secher, and B. Nielsen. Inadequate heat release from the human brain during prolonged exercise with hyperthermia. *The Journal of Physiology*, 545(Pt 2):697–704, Dec. 2002.

[O'Donnell et al., 2004] J. M. O'Donnell, R. K. Kudej, K. F. LaNoue, S. F. Vatner, and E. D. Lewandowski. Limited transfer of cytosolic NADH into mitochondria at high cardiac workload. *American Journal of Physiology - Heart and Circulatory Physiology*, 286(6):H2237—-42, June 2004.

[Oevelen et al., 2009] D. Oevelen, K. Meersche, F. J. R. Meysman, K. Soetaert, J. J. Middelburg, and A. F. Vézina. Quantifying Food Web Flows Using Linear Inverse Models. *Ecosystems*, 13(1):32–45, Nov. 2009.

[Olds et al., 1993] T. S. Olds, K. I. Norton, and N. P. Craig. Mathematical model of cycling performance. *Journal of applied physiology*, 75(2):730–7, Aug. 1993.

[Oliveira and Kowaltowski, 2004] G. A. Oliveira and A. J. Kowaltowski. Phosphate increases mitochondrial reactive oxygen species release. *Free Radical Research*, 38 (10):1113–8, Oct. 2004.

[O'Reilly et al., 2010] J. O'Reilly, S. H. S. Wong, and Y. Chen. Glycaemic index, glycaemic load and exercise performance. *Sports medicine (Auckland, N.Z.)*, 40(1):27–39, Jan. 2010.

[Pan and Raftery, 2007] Z. Pan and D. Raftery. Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Analytical and bioanalytical chemistry*, 387(2):525–7, Jan. 2007.

[Panchal et al., 2000] A. R. Panchal, B. Comte, H. Huang, T. Kerwin, A. Darvish, C. des Rosiers, H. Brunengraber, and W. C. Stanley. Partitioning of pyruvate between oxidation and anaplerosis in swine hearts. *American Journal of Physiology - Heart and Circulatory Physiology*, 279(5):H2390–8, Nov. 2000.

[Panchal et al., 2001] A. R. Panchal, B. Comte, H. Huang, B. Dudar, B. Roth, M. Chandler, C. Des Rosiers, H. Brunengraber, and W. C. Stanley. Acute hibernation decreases myocardial pyruvate carboxylation and citrate release. *American Journal of Physiology - Heart and Circulatory Physiology*, 281(4):H1613–20, Oct. 2001.

[Perumal and Gunawan, 2011] T. M. Perumal and R. Gunawan. Understanding dynamics using sensitivity analysis: caveat and solution. *BMC systems biology*, 5(1):41, Jan. 2011.

[Petzold, 1983] L. Petzold. Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations. *SIAM Journal on Scientific and Statistical Computing*, 4(1):136 – 148, 1983.

[Quek et al., 2009] L.-E. Quek, C. Wittmann, L. K. Nielsen, and J. O. Krömer. OpenFLUX: efficient modelling software for 13C-based metabolic flux analysis. *Microbial cell factories*, 8(1):25, Jan. 2009.

[Quistorff et al., 2008] B. r. Quistorff, N. H. Secher, and J. J. Van Lieshout. Lactate fuels the human brain during exercise. *The FASEB journal*, 22(10):3443–9, Oct. 2008.

[Radrich et al., 2010] K. Radrich, Y. Tsuruoka, P. Dobson, A. Gevorgyan, N. Swainston, G. Baart, and J.-M. Schwartz. Integration of metabolic databases for the reconstruction of genome-scale metabolic networks. *BMC systems biology*, 4(1):114, Jan. 2010.

[Randle et al., 1970] P. J. Randle, P. J. England, and R. M. Denton. Control of the tricarboxylate cycle and its interactions with glycolysis during acetate utilization in rat heart. *The Biochemical journal*, 117(4):677–695, May 1970.

[Rasmussen et al., 2010] P. Rasmussen, N. Nyberg, J. W. Jaroszewski, R. Krogh-Madsen, N. H. Secher, and B. r. Quistorff. Brain nonoxidative carbohydrate consumption is not explained by export of an unknown carbon source: evaluation of the arterial and jugular venous metabolome. *Journal of cerebral blood flow and metabolism*, 30 (6):1240–6, June 2010.

[Redfield and Hull, 1986] R. Redfield and M. L. Hull. Prediction of pedal forces in bicycling using optimization methods. *Journal of biomechanics*, 19(7):523–40, Jan. 1986.

[Reeder et al., 1999] S. B. Reeder, A. A. Holmes, E. R. McVeigh, and J. R. Forder. Simultaneous noninvasive determination of regional myocardial perfusion and oxygen content in rabbits: toward direct measurement of myocardial oxygen consumption at MR imaging. *Radiology*, 212(3):739–47, Sept. 1999.

[Robitaille et al., 1993] P. M. Robitaille, D. P. Rath, A. M. Abduljalil, J. M. O'Donnell, Z. Jiang, H. Zhang, and R. L. Hamlin. Dynamic 13C NMR analysis of oxidative metabolism in the in vivo canine myocardium. *The Journal of Biological Chemistry*, 268(35):26296–301, Dec. 1993.

[Rosenthal, 2007] J. S. Rosenthal. AMCMC: An R interface for adaptive MCMC. *Computational Statistics & Data Analysis*, 51(12):5467–5470, Aug. 2007.

[Sahlin et al., 1987] K. Sahlin, A. Katz, and J. Henriksson. Redox state and lactate accumulation in human skeletal muscle during dynamic exercise. *The Biochemical journal*, 245(2):551–556, July 1987.

[Saks et al., 2008] V. Saks, N. Beraud, and T. Wallimann. Metabolic compartmentation - a system level property of muscle cells: real problems of diffusion in living cells. *International journal of molecular sciences*, 9(5):751–67, May 2008.

[Saks et al., 1976] V. A. Saks, G. B. Chernousova, R. Vetter, V. N. Smirnov, and E. I. Chazov. Kinetic properties and the functional role of particulate MM-isoenzyme of creatine phosphokinase bound to heart muscle myofibrils. *FEBS Letters*, 62(3): 293–296, 1976.

[Saks et al., 1996] V. A. Saks, R. Ventura-Clapier, and M. K. Aliev. Metabolic control and metabolic capacity: two aspects of creatine kinase functioning in the cells. *Biochimica et Biophysica Acta*, 1274(3):81–88, 1996.

[Saks et al., 2000] V. A. Saks, O. Kongas, M. Vendelin, and L. Kay. Role of the creatine/phosphocreatine system in the regulation of mitochondrial respiration. *Acta Physiologica Scandinavica*, 168(4):635–641, 2000.

[Sangar et al., 2012] V. Sangar, J. A. Eddy, E. Simeonidis, and N. D. Price. Mechanistic modeling of aberrant energy metabolism in human disease. *Frontiers in Physiology*, 3:404, Jan. 2012.

[Sauer, 2006] U. Sauer. Metabolic networks in motion: 13C-based flux analysis. *Molecular Systems Biology*, 2:62, Jan. 2006.

[Saupe et al., 1998] K. W. Saupe, M. Spindler, R. Tian, and J. S. Ingwall. Impaired cardiac energetics in mice lacking muscle-specific isoenzymes of creatine kinase. *Circulation Research*, 82(8):898–907, May 1998.

[Schaber et al., 2009] J. Schaber, W. Liebermeister, and E. Klipp. Nested uncertainties in biochemical models. *IET Systems Biology*, 3(1):1–9, 2009.

[Schellenberger and Palsson, 2009] J. Schellenberger and B. O. Palsson. Use of randomized sampling for analysis of metabolic networks. *The Journal of Biological Chemistry*, 284(9):5457–61, Feb. 2009.

[Schellenberger et al., 2010] J. Schellenberger, J. O. Park, T. M. Conrad, and B. O. Palsson. BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. *BMC Bioinformatics*, 11(1):213, Apr. 2010.

[Schellenberger et al., 2011] J. Schellenberger, R. Que, R. M. T. Fleming, I. Thiele, J. D. Orth, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis, S. Rahmanian, J. Kang, D. R. Hyduke, and B. O. Palsson. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature Protocols*, 6(9): 1290–307, Sept. 2011.

[Schellenberger et al., 2012] J. Schellenberger, D. C. Zielinski, W. Choi, S. Madireddi, V. Portnoy, D. A. Scott, J. L. Reed, A. L. Osterman, and B. O. Palsson. Predicting outcomes of steady-state 13C isotope tracing experiments with Monte Carlo sampling. *BMC Systems Biology*, 6(1):9, Jan. 2012.

[Schmitz et al., 2011] J. P. J. Schmitz, J. Vanlier, N. A. W. van Riel, and J. A. L. Jeneson. Computational modeling of mitochondrial energy transduction. *Critical reviews in biomedical engineering*, 39(5):363–77, Jan. 2011.

[Schroeder et al., 2009] M. A. Schroeder, H. J. Atherton, D. R. Ball, M. A. Cole, L. C. Heather, J. L. Griffin, K. Clarke, G. K. Radda, and D. J. Tyler. Real-time assessment of Krebs cycle metabolism using hyperpolarized 13C magnetic resonance spectroscopy. *The FASEB journal*, 23(8):2529–38, Aug. 2009.

[Sepp et al., 2010] M. Sepp, M. Vendelin, H. Vija, and R. Birkedal. ADP compartmentation analysis reveals coupling between pyruvate kinase and ATPases in heart muscle. *Biophysical journal*, 98(12):2785–93, June 2010.

[Sherry et al., 2004] A. D. Sherry, F. M. Jeffrey, and C. R. Malloy. Analytical solutions for 13c isotopomer analysis of complex metabolic conditions: substrate oxidation, multiple pyruvate cycles, and gluconeogenesis. *Metabolic Engineering*, 6(1):12–24, Jan. 2004.

[Sjogaard et al., 1985] G. Sjogaard, R. P. Adams, and B. Saltin. Water and ion shifts in skeletal muscle of humans with intense dynamic knee extension. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 248(2):R190–196, 1985.

[Soetaert et al., 2009] K. Soetaert, K. V. D. Meersche, and D. V. Oevelen. limsolve: solving linear inverse models in r. *R package version 1.5.5. Available at* `http://www.cran.r-project.org/`, 2009.

[Sorokina et al., 2007] N. Sorokina, J. M. O'Donnell, R. D. McKinney, K. M. Pound, G. Woldegiorgis, K. F. LaNoue, K. Ballal, H. Taegtmeyer, P. M. Buttrick, and E. D. Lewandowski. Recruitment of compensatory pathways to sustain oxidative flux with reduced carnitine palmitoyltransferase I activity characterizes inefficiency in energy metabolism in hypertrophied hearts. *Circulation*, 115(15):2033–2041, Apr. 2007.

[St Clair Gibson and Noakes, 2004] A. St Clair Gibson and T. D. Noakes. Evidence for complex system integration and dynamic neural regulation of skeletal muscle recruitment during exercise in humans. *British journal of sports medicine*, 38(6): 797–806, Dec. 2004.

[Stolwijk, 1971] J. A. Stolwijk. A mathematical model of physiological temperature regulation in man. Contractor Report CR-1855. Technical report, National Aeronautics and Space Administration, Washington DC, 1971.

[Stolwijk and Hardy, 1966a] J. A. Stolwijk and J. D. Hardy. Temperature regulation in man–a theoretical study. *Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere*, 291(2):129–62, Jan. 1966a.

[Stolwijk and Hardy, 1966b] J. A. Stolwijk and J. D. Hardy. Partitional calorimetric studies of responses of man to thermal transients. *Journal of Applied Physiology*, 21(3): 967–977, May 1966b.

[Stoner and Sirak, 1979] C. D. Stoner and H. D. Sirak. Steady-state kinetics of the overall oxidative phosphorylation reaction in heart mitochondria. *Journal of Bioenergetics and Biomembranes*, 11(5-6):113–146, 1979.

[Teague Jr. and Dobson, 1992] W. E. Teague Jr. and G. P. Dobson. Effect of temperature on the creatine kinase equilibrium. *The Journal of Biological Chemistry*, 267(20): 14084–14093, 1992.

[Teusink et al., 2000] B. Teusink, J. Passarge, C. A. Reijenga, E. Esgalhado, C. C. van der Weijden, M. Schepper, M. C. Walsh, B. M. Bakker, K. van Dam, H. V. Westerhoff, and J. L. Snoep. Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? testing biochemistry. *European Journal of Biochemistry*, 267(17):5313–5329, Sept. 2000.

[Thiele et al., 2013] I. Thiele, N. Swainston, R. M. T. Fleming, A. Hoppe, S. Sahoo, M. K. Aurich, H. Haraldsdottir, M. L. Mo, O. Rolfsson, M. D. Stobbe, S. G. Thorleifsson, R. Agren, C. Bölling, S. Bordel, A. K. Chavali, P. Dobson, W. B. Dunn, L. Endler, D. Hala, M. Hucka, D. Hull, D. Jameson, N. Jamshidi, J. J. Jonsson, N. Juty, S. Keating, I. Nookaew, N. Le Novère, N. Malys, A. Mazein, J. A. Papin, N. D. Price, E. Selkov, M. I. Sigurdsson, E. Simeonidis, N. Sonnenschein, K. Smallbone, A. Sorokin, J. H. G. M. van Beek, D. Weichart, I. Goryanin, J. Nielsen, H. V. Westerhoff, D. B. Kell, P. Mendes, and B. O. Palsson. A community-driven global reconstruction of human metabolism. *Nature Biotechnology*, 31(5):419–25, May 2013.

[Tipton, 2005] C. M. Tipton. *ACSM's Advanced Exercise Physiology*. Lippincott Williams & Wilkins, Baltimore, 1 edition, 2005. ISBN 0781747260.

[Van Beek, 2008] J. H. Van Beek. Multiscale and modular analysis of cardiac energy metabolism: repairing the broken interfaces of isolated system components. *Annals of the New York Academy of Sciences*, 1123:155–168, Mar. 2008.

[van Beek and Westerhof, 1990] J. H. van Beek and N. Westerhof. Response time of mitochondrial oxygen consumption following stepwise changes in cardiac energy demand. *Advances in Experimental Medicine and Biology*, 277:415–423, 1990.

[Van Beek and Westerhof, 1991] J. H. Van Beek and N. Westerhof. Response time of cardiac mitochondrial oxygen consumption to heart rate steps. *The American journal of physiology*, 260(2 Pt 2):H613–25, Feb. 1991.

[van Beek et al., 1998] J. H. van Beek, T. Csont, F. J. de Kanter, and J. Bussemaker. Simple model analysis of 13C NMR spectra to measure oxygen consumption using frozen tissue samples. *Advances in Experimental Medicine and Biology*, 454:475–85, Jan. 1998.

[van Beek et al., 1999] J. H. van Beek, H. G. van Mil, R. B. King, F. J. de Kanter, D. J. Alders, and J. Bussemaker. A (13)C NMR double-labeling method to quantitate local myocardial O(2) consumption using frozen tissue samples. *The American journal of physiology*, 277(4 Pt 2):H1630–40, Oct. 1999.

[van Beek, 1996] J. H. G. M. van Beek. Heat generation and transport in the heart. *Journal of Engineering Physics and Thermophysics*, 69(3):287–297, May 1996.

[van Beek, 2007] J. H. G. M. van Beek. Adenine nucleotide-creatine-phosphate module in myocardial metabolic system explains fast phase of dynamic regulation of oxidative phosphorylation. *American Journal of Physiology - Cell Physiology*, 293(3):C815–C829, 2007.

[van Beek et al., 2008] J. H. G. M. van Beek, H. Hettling, and T. W. Binsl. Computational methods to investigate and manage molecular networks in cells. In C. Backendorf, M. H. Noteborn, and M. Tavassoli, editors, *Proteins killing tumor cells*, page 317. Research Signpost, 2008. ISBN 978-81-7895-405-9.

[van Beek et al., 2009] J. H. G. M. van Beek, A.-C. Hauschild, H. Hettling, and T. W. Binsl. Robust modelling, measurement and analysis of human and animal metabolic systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1895):1971–1992, 2009.

[van Beek et al., 2011] J. H. G. M. van Beek, F. Supandi, A. K. Gavai, A. A. de Graaf, T. W. Binsl, and H. Hettling. Simulating the physiology of athletes during endurance sports events: modelling human energy conversion and metabolism. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1954):4295–4315, Oct. 2011.

[Veksler et al., 1995] V. I. Veksler, A. V. Kuznetsov, K. Anflous, P. Mateo, J. van Deursen, B. Wieringa, and R. Ventura-Clapier. Muscle creatine kinase-deficient mice. II. Cardiac and skeletal muscles exhibit tissue-specific adaptation of the mitochondrial function. *The Journal of Biological Chemistry*, 270(34):19921–9, Aug. 1995.

[Vendelin and Birkedal, 2008] M. Vendelin and R. Birkedal. Anisotropic diffusion of fluorescently labeled ATP in rat cardiomyocytes determined by raster image correlation spectroscopy. *American Journal of Physiology - Cell physiology*, 295(5):C1302–15, Nov. 2008.

[Vendelin et al., 2000] M. Vendelin, O. Kongas, and V. Saks. Regulation of mitochondrial respiration in heart cells analyzed by reaction-diffusion model of energy transfer. *American Journal of Physiology - Cell Physiology*, 278(4):C747–64, 2000.

[Vendelin et al., 2004] M. Vendelin, M. Eimre, E. Seppet, N. Peet, T. Andrienko, M. Lemba, J. Engelbrecht, E. K. Seppet, and V. A. Saks. Intracellular diffusion of adenosine phosphates is locally restricted in cardiac muscle. *Molecular and Cellular Biochemistry*, 256-257(1-2):229–241, 2004.

[Vendelin et al., 2010] M. Vendelin, J. A. Hoerter, P. Mateo, S. Soboll, B. Gillet, and J.-L. Mazet. Modulation of energy transfer pathways between mitochondria and myofibrils by changes in performance of perfused heart. *The Journal of Biological Chemistry*, 285(48):37240–50, Nov. 2010.

[Vicini and Kushmerick, 2000] P. Vicini and M. J. Kushmerick. Cellular energetics analysis by a mathematical model of energy balance: estimation of parameters in human skeletal muscle. *American Journal of Physiology - Cell physiology*, 279(1):C213—-24, July 2000.

[Wallimann et al., 2011] T. Wallimann, M. Tokarska-Schlattner, and U. Schlattner. The creatine kinase system and pleiotropic effects of creatine. *Amino acids*, 40(5):1271–96, May 2011.

[Weiss and Sinha, 1978] H. R. Weiss and A. K. Sinha. Regional oxygen saturation of small arteries and veins in the canine myocardium. *Circulation research*, 42(1):119–26, Jan. 1978.

[Weiss et al., 1989] R. G. Weiss, V. P. Chacko, J. D. Glickson, and G. Gerstenblith. Comparative 13C and 31P NMR assessment of altered metabolism during graded reductions in coronary flow in intact rat hearts. *Proceedings of the National Academy of Sciences of the United States of America*, 86(16):6426–6430, Aug. 1989.

[Weiss et al., 1993] R. G. Weiss, R. Kalil-Filho, A. Herskowitz, V. P. Chacko, M. Litt, M. D. Stern, and G. Gerstenblith. Tricarboxylic acid cycle activity in postischemic rat hearts. *Circulation*, 87(1):270–82, Jan. 1993.

[Weitzel et al., 2012] M. Weitzel, K. Nöh, T. Dalman, S. Niedenführ, B. Stute, and W. Wiechert. 13CFLUX2 - High-Performance Software Suite for 13C-Metabolic Flux Analysis. *Bioinformatics (Oxford, England)*, Oct. 2012.

[Werner, 2010] J. Werner. System properties, feedback control and effector coordination of human temperature regulation. *European Journal of Applied Physiology*, 109(1): 13–25, May 2010.

[Wiechert] W. Wiechert. 13c metabolic flux analysis. *Metabolic Engineering*, (3):195–206, July .

[Wishart et al., 2007] D. S. Wishart, D. Tzur, C. Knox, R. Eisner, A. C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M. A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D. D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G. E. Duggan, G. D. Macinnis, A. M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B. D. Sykes, H. J. Vogel, and L. Querengesser. Hmdb: the human metabolome database. *Nucleic Acids Res.*, 35 (Database issue):D521—-526, Jan. 2007.

[Wu and Beard, 2009] F. Wu and D. A. Beard. Roles of the creatine kinase system and myoglobin in maintaining energetic state in the working heart. *BMC Systems Biology*, 3:22, 2009.

[Wu et al., 2008] F. Wu, E. Y. Zhang, J. Zhang, R. J. Bache, and D. A. Beard. Phosphate metabolite concentrations and ATP hydrolysis potential in normal and ischaemic hearts. *The Journal of Physiology*, 586(Pt 17):4193–4208, 2008.

[Yang et al., 2005] J. Yang, S. Wongsa, V. Kadirkamanathan, S. A. Billings, and P. C. Wright. Metabolic flux distribution analysis by 13C-tracer experiments using the Markov chain-Monte Carlo method. *Biochemical Society Transactions*, 33(Pt 6): 1421–1422, 2005.

[Yu et al., 1995] X. Yu, L. T. White, C. Doumen, L. A. Damico, K. F. LaNoue, N. M. Alpert, and E. D. Lewandowski. Kinetic analysis of dynamic 13C NMR spectra: metabolic flux, regulation, and compartmentation in hearts. *Biophysical journal*, 69(5):2090–2102, Nov. 1995.

[Zawadzki et al., 1992] K. M. Zawadzki, B. B. Yaspelkis, and J. L. Ivy. Carbohydrate-protein complex increases the rate of muscle glycogen storage after exercise. *Journal of Applied Physiology*, 72(5):1854–9, May 1992.

Summary

Viable Models of Energy Metabolism
Contemplating Uncertainty in Measured Data, Parameter Estimates and Predictions

This thesis describes three computational approaches to analyse (mammalian) metabolic systems: kinetic modelling, flux estimation from stable isotope measurements and prediction of flux distribution in metabolic networks using constraint-based modelling. In all three approaches, emphasis is put on the quantification of uncertainty for parameters and model predictions.

The sustainment of life in any organism depends on the uptake and subsequent biochemical conversion of nutritional substances from the environment. The entirety of all biochemical reactions which transform nutrition into metabolites used for growth, proliferation or as a source of energy refers to metabolism. Investigation of metabolism is therefore essential in our understanding of life at a molecular level. Until now, the quantitative investigation of metabolism heavily relies on the application of computational models and methods so as to make sense out of data derived from laboratory experiments. The discipline of combining experimental and theoretical approaches in order to understand a biological system as a whole is defined as Systems Biology. A computational model in Systems Biology is essentially an abstraction of a biological process, formulated in mathematical terms. Oftentimes, models are calibrated by adjusting their parameter values such that the model dynamics can describe given experimental data. The calibrated models can then be used to test hypotheses about the biological system *in silico*. In the research of this thesis, mainly the metabolic components necessary for an organism to conduct mechanical work are investigated. In order to get further insight into the energy metabolism in animal and human muscle, various types of models encoded in different mathematical formalisms were used. The formalisms as well as techniques to complement computational models

with experimental data are reviewed in chapter 2.

Chapter 3 is dedicated to the creatine kinase system, an intriguing enzymatic process involved predominantly in the energy metabolism of muscle cells. Several metabolic functions have been attributed to the creatine kinase system, of which one has been significantly controversial and heavily debated. This controversial metabolic function known as 'phosphocreatine shuttle hypothesis' refers to the widespread assumption that the creatine kinase system in heart muscle is essential for the transport of energy-carrying molecules from sites of energy production (the mitochondria) to sites of energy consumption (the myofibrils). This hypothesis was challenged in chapter 3. We did this by quantifying energy transport, using a relevant mathematical model on the creatine kinase system integrated with *in vivo* experimental data from the isolated heart. In order to obtain model predictions that are statistically sound with respect to measurement error in the data, model parameters were estimated as posterior probability distributions rather than as single numerical values. Model predictions are therefore drawn from ensembles of probable parameter sets. The analysis suggests that the 'phosphocreatine shuttle' only plays a minor role in heart muscle. Instead, the creatine kinase system mainly acts as a buffering system which assures sufficient energy supply even when the heart beats fast and much energy is consumed. Further, a new hypothesis for the creatine kinase system is proposed based on the model analysis: creatine kinase lowers inorganic phosphate levels in the cytosol to protect from oxidative stress.

The next chapter (chapter 4) describes the simulation of energy conversion in a cyclist's body during a mountain time trial of the Tour de France. Besides the investigation of heat production and transport using a whole body model, the biochemical events during muscle contraction in the cyclist's leg are predicted. Again, emphasis was put on the functional role of the creatine kinase. The model used in the chapter 3 was therefore adapted to describe energy turnover in human muscle and extended to account for muscle acidity. For model calibration, measurements from human leg muscle during a bicycle exercise bout were used. Applying the calibrated model, it was predicted that the main physiological role of the creatine kinase enzyme is, similar than in cardiac tissue, the buffering of energy resources when the body is exposed to high bursts of sudden energy demand during the bicycle race.

Chapter 5 returns to the investigation of energy metabolism in cardiac tissue. This time, the focus is set on another essential process in mitochondrial energy production: the citric acid cycle. To quantify energy turnover in distinct small regions of the myocardium, we used *in vivo* data from carbon tracer experiments in the porcine heart. Small tissue samples were taken after isotopically labelled substrate was given to fuel aerobic energy metabolism. A sophisticated model that describes all transitions between carbon atoms of the different metabolites within the system was then used to analyse the data. As a result, the velocities of biochemical reactions around the citric acid cycle could be quantified, which constitutes important information on the physiology of the myocardium. To account for the relatively high noise level in the data, we adapted the approach of dealing with parameter uncertainty by gen-

erating ensembles of probable parameter combinations from chapter 3 to the carbon transition networks used in this chapter. This framework allowed us to analyse experimental data gained under different physiological and pharmacological conditions and to draw conclusions about the citric acid cycle fluxes in the different cardiac states. For instance, the citric acid cycle fluxes during ischaemia could be estimated, giving valuable information about metabolic energy fluxes in the failing heart.

The final chapter introduces the software tool BiGGR for constraint-based modelling of metabolism. Constraint based modelling is a technique to estimate metabolic reaction fluxes when *a priori* quantitative information on the network is sparse. Model constraints, for instance the steady-state constraint are used to limit the space of possible flux values. BiGGR allows the assembly of metabolic models from, among others, the BiGG database of metabolic reconstructions by querying for pathways, genes, metabolites or reactions of interest. Fluxes can then be estimated using a variety of linear inverse modelling algorithms. Finally, the models and estimated fluxes can be visualised in an intuitive way. Naturally, also constraint based models are subject to uncertainty in the predicted fluxes. BiGGR therefore uses sophisticated algorithms to enable the generation of flux ensembles respecting the constraints of the model. To demonstrate the practicality of the software, chapter 6 describes an entire work flow of flux estimation in an extensive model of energy metabolism in the human brain. BiGGR is implemented in the R programming language and publicly available to the scientific community in the Bioconductor repository.

In conclusion, computational strategies to quantitatively analyse metabolic systems were developed and applied during the course of this thesis. Special emphasis was put on the analysis of uncertainty in predictions derived from computational models of metabolism. The results described in this thesis thus contribute to new advancements in the investigation of energy metabolism using a Systems Biology approach.

Samenvatting

Realistische Modellen voor Energiemetabolisme
Beschouwingen over Onzekerheid in Data, Parameterschattingen en Voorspellingen

Dit proefschrift presenteert drie verschillende computationele benaderingen voor de analyse van metabole systemen van zoogdieren: kinetische modelleren, schatten van metabole fluxen middels metingen uit experimenten met stabiele isotopen en de voorspelling van fluxdistributies in metabole netwerken door middel van 'constraint-based modeling'.

In elk organisme is het (behoud van het) leven afhankelijk van de opname en de daaropvolgende biochemische conversie van voedingsstoffen. Het geheel van alle biochemische reacties verantwoordelijk voor de omzetting van voedingstoffen in metabolieten voor groei, proliferatie of een energiebron wordt 'metabolisme' genoemd. Onderzoek naar het metabolisme is daarom essentieel voor ons begrip van het leven op een moleculair niveau te vergroten. Tot nu toe is het kwantitatieve onderzoek naar metabolisme sterk afhankelijk van de toepassing van wiskundige modellen en methoden om de metingen afkomstig van laboratoriumexperimenten te verklaren. De discipline waarin experimentele en theoretische benaderingen worden gecombineerd om een biologisch systeem in het geheel te begrijpen wordt gedefinieerd als 'systeembiologie'. Een wiskundig model in de systeembiologie is in principe een abstractie van een biologisch proces, geformuleerd in wiskundige termen. Modellen worden vaak gekalibreerd door de aanpassing van parameterwaarden zodanig dat het model de dynamiek van de experimentele data kan beschrijven. Met de gekalibreerde modellen kunnen vervolgens hypotheses over het biologische systeem *in silico* getest worden. Het onderzoek in dit proefschrift behandeld voornamelijk het metabolisme dat noodzakelijk is voor het uitvoeren van mechanisch werk. Om een nader inzicht in het energiemetabolisme van de dierlijke (en menselijke) spier te krijgen, worden er verschillende soorten

modellen gecodeerd en verschillende wiskundige formalismen gebruikt. Deze formalismen alsook de technieken om de wiskundige modellen met experimentele data te combineren zijn beschreven in hoofdstuk 2.

Het derde hoofdstuk is gewijd aan het 'creatine kinase systeem', een fascinerende enzymatische proces dat hoofdzakelijk de energiestofwisselling in spiercellen betreft. Er zijn meerdere metabole functies toegeschreven aan dit creatine kinase systeem, waarvan vooral één functie zeer omstreden is. Deze omstreden functie, ook bekend als de 'fosfocreatine shuttle hypothese', is de wijdverbreide veronderstelling dat het creatine kinase systeem in de hartspier essentieel is voor het transport van energiedragende moleculen van plekken van de productie van energie (de mitochondriën) naar plekken van het verbruik van energie (de myofibrillen). Deze hypothese wordt uitgedaagd in hoofstuk 3. Hierbij is het energietransport met behulp van een relevant wiskundig model van de creatine kinase systeem gekwantificeerd en geïntegreerd met *in vivo* data van experimenten met een geïsoleerde hart. Om statistische zekerheid over de voorspellingen van ons model te verkrijgen met betrekking tot meetfouten in de data, worden de modelparameters niet als enkele waarden gezien, maar voor iedere parameter wordt een distributie van mogelijke waarden berekend. Voorspellingen van het model zijn daarom afgeleid van 'ensembles' van waarschijnlijke sets van de parameters. Onze berekeningen suggereren dat de 'fosfocreatine shuttle' slechts een kleine rol speelt in het metabolisme van de hartspier. Wij betogen dat het creatine kinase systeem voornamelijk een buffer functie heeft, die een voldoende energieverzorging garandeerd als het hart sneller klopt en er dus meer energie verbruikt wordt. Verder wordt een nieuwe hypothese voor het creatine kinase systeem voorgesteld op basis van onze modelanalyse: creatine kinase verlaagt het niveau van anorganisch fosfaat in het cytosol om de cel te beschermen tegen oxidatieve schade.

Het volgende hoofdstuk (hoofdstuk 4) beschrijft de simulatie van het energie metabolisme in het lichaam van een fietser tijdens een berg-tijdrit in de Tour de France. Naast het onderzoek naar de warmteproductie en -transport met behulp van wiskundig model van het gehele lichaam, zijn de biochemische gebeurtenissen tijdens spiercontractie in de benen van de fietser voorspeld. Opnieuw wordt de functie van het creatine kinase systeem onderzocht. Daarvoor wordt het model van hoofdstuk 3 aangepast om menselijke spieren te beschrijven en uitgebreid om rekening te houden met de zuurgraad van de spier. Het model wordt gekalibreerd met metingen in een menselijke beenspier tijdens een oefening op de fietsergometer. Met het gekalibreerde model wordt voorgespelt dat, net als in de hartspier, de belangrijkste fysiologische functie van het creatine kinase systeem de buffering is van de concentratie van energie-dragende moleculen als de energievraag van het lichaam verhoogt is tijdens de fietswedstrijd.

In hoofdstuk 5 gaat het weer over de metabolisme in hartspier. Hier ligt de focus op en andere essentiële voortgang in de mitochondriale energieverzorging: de citroenzuurcyclus. Om het energieverbruik in verschillende kleine plekken in de hartspier te kwantificeren hebben wij in vivo metingen van experimenten met een koolstof tracer in een varkenshart. Kleine biopsies van het weefsel zijn gebruikt, nadat een met isotopen gemarkeerd substraat is gegeven als energiebron. Met een gedetailleerd model,

dat alle overgangen tussen de koolstofatomen van de verschillende metabolieten beschrijft, hebben wij de data geanalyseerd. Daardoor konden wij de omzetting van de biochemische reacties in de citroenzuurcyclus kwantificeren, wat belangrijke informatie over de fysiologie van het hart oplevert. Om met het relatief hoge niveau van ruis in de metingen rekening te houden hebben wij de 'ensemble' methode uit hoofdstuk 3 aangepast, opdat zij met het model van de koolstof-overgangen werkt. Hiermee konden wij vervolgens de experimentele data over verschillende fysiologische en farmacologische condities analyseren en conclusies trekken over de fluxen van de citroenzuurcyclus bij verschillende hartbelasting. Het was bijvoorbeeld mogelijk de fluxen van de citroenzuurcyclus voor het ischemische hart te schatten, wat ons interessante informatie over het zieke hart geeft.

Het laatste hoofdstuk introduceert het software pakket 'BiGGR' voor constraint-based modelleren van het metabolisme. Constraint-based modelleren is een techniek voor het schatting van de fluxen van metabole reacties als er niet veel *a priori* informatie over de metabole netwerk is. Beperkingen van het model (constraints) bijvoorbeeld de steady-state beperking, helpen de ruimte van mogelijke fluxwaarden te beperken. BiGGR stelt ons in staat een systeem van metabole reacties te maken door de integratie met de BiGG database waar de gebruiker naar genen, metabolieten of metabole routes kan zoeken. Metabole fluxen kunnen geschat worden door gebruik te maken van verschillende lineaire inverse modelleringsalgoritmen. Modellen en geschatte fluxen kunnen vervolgens op een intuïtive manier gevisualiseerd worden. Natuurlijk zijn ook constraint-based modellen onderworpen aan onzekerheid van de gekwantificeerde fluxen. Daarom gebruikt BiGGR geraffineerde algorithmen om fluxensembles te berekenen die de beperkingen van het model respecteren. Wij tonen de toepasbaarheid van deze software in hoofdstuk 6, waarin een gehele workflow met een uitgebreid model van het metabolisme in de menselijke hersenen geanalyseerd wordt. BiGGR is geïmplementeerd in programmeertaal R en beschikbaar via de Bioconductor archief.

Concluderend, in dit proefschrift worden computationele strategiën ontwikkeld voor en toegepast op metabole systemen. Speciale nadruk wordt gelegd op de analyse van de onzekerheid in de voorspellingen gemaakt door deze computationele modellen. De resultaten die in dit proefschrift beschreven zijn dragen daarom bij aan nieuwe ontwikkelingen in het onderzoek van het metabolisme met methodologie uit de systeembiologie.

Acknowledgements

The long term project of a doctoral thesis is never an effort of only a single person. The success of conducting the research required for this thesis therefore naturally depended on my scientific mentors, collaborators and co-workers. And, of course, friends and family were essential for reminding that life does not consist entirely of research, manuscripts and deadlines.

Here I would like to thank everyone who supported me on the way of obtaining my degree. First of all, I am greatly indebted to my supervisors Hans van Beek and Jaap Heringa. Hans' kind and patient guidance and his excellent scientific input was the essential factor for the successful completion of my Ph.D. Throughout my time at the Vrije Universiteit, I was always reassured that Hans would be there for me and I would like to express my gratitude for that. My promotor Jaap's door was always open for me to help with any concern. Jaap's knowledge, his keen intellect and also his genuine interest in my person make it a pleasure to have a conversation with him. I hope to continue the professional and private contact with both of them in the future.

My time at the IBIVU would not have been the same without my great colleagues. Thomas, Bernd, Walter and Bart, each one of them was a colleague at the beginning but has become a very good friend. The time we spend together in room P.138 will always remain unforgettable. I am very thankful to Bernd and Thomas for all their scientific input and support and I am happy to have both as my paranymphs. I would also like to express my gratidute to all other members of the IBIVU (and to some people in the department of clinical genetics at the VUmc) that I had the pleasure to work with: Anton, Nicola, Sanne, Anand, Hilal, Farah, René, and as well to the 'younger generation' of Ph.D. students: Ali, Mohammed, Punto, Qingzhen, Annika and especially Erik who gave valuable input for this thesis.

During my time as a Ph.D. student, I had the privilege to be in a stable employment at the Vrije Universiteit Amsterdam in positions funded by the Centre for Medical Systems Biology and the Netherlands Consortium for Systems Biology, which I am very greateful for. I would also like to thank all members of the committee that

took the time and effort to read this thesis and to be part of the defence ceremony: Bas Teusink, Bret Olivier, Albert de Graaf, Jaap Molenaar, Marko Vendelin, and Martijn Huynen.

I also want to thank everyone who helped me to finalize this book: Bernd for helping me with the dutch summary, Brian and Anika for designing the cover, my sister Anna for helping me with the figures and Irisa for professional proofreading and language editing.

A special thanks goes to everyone that I had the pleasure to live with in the shared flat on Singel 6C. There is certainly never a dull moment in this house.

Whatever life throws at me, I am greatful that I always am able to count on my family. I would like to share this gratitude with my parents, Gerd and Bettina, my siblings Anna and Benni and my grandmother, Nora.

Finally, I want to express my genuine gratitute for my girlfriend Irisa Ono. We met at the university when I just had started my Ph.D. and the time that followed was the best in my life. We have been inseperable ever since and I hope this will never change.