METABOLIC MODELING OF DIFFERENT DEGREES OF STEATOHEPATITIS IN MICE

Dissertation

submitted to the

Department of Mathematics and Computer Science

of the Freie Universität Berlin

for the degree of

Doctor of Natural Sciences (Dr. rer. nat.)

by

Vikash Kumar Pandey

July 2014

This work was accomplished during the period from August 2010 till July 2014 at the Max Planck Institute for Molecular Genetics in the department of Prof. Dr. Hans Lehrach in the systems biology group of Dr. Christoph Wierling.

First reviewer: Prof. Dr. Alexander Bockmayr
Second reviewer: Prof. Dr. Hans Lehrach
Third reviewer: Prof. Dr. Matteo Barberis
Day of the disputation: 27 October 2014

Contents

Acknowledgements

I would like to express my deepest gratitude to Prof. Dr. Hans Lehrach for providing me the chance to pursue my PhD studies in his department of vertebrate genomics.

My sincere thanks go to Prof. Dr. Alexander Bockmayr for giving me the great opportunity to work in the highly interesting field of constraint-based modeling.

This work was inspired and substantially influenced by the ideas of Dr. Christoph Wierling, who also provided me his continuous comprehensive support and constructive advice on all aspects of my scientific research.

I am indebted to the Systems biology group of Dr. Christoph Wierling in the department of Prof. Lehrach, namely Dr. Hendrik Hache, Dr. Thomas Kessler, and Alexey Shadrin.

I am specially grateful to David DeMiglio, Alexey Shadrin, and Thomas Kessler for proofreading and commenting on the manuscript of my thesis.

My sincere thanks go to members of IMGuS project for providing data sets.

Finally, I want to thank my parents and friends for their support throughout my thesis.

This work was supported by the Max Planck Society.

List of Figures

List of Tables

Summary

The rate of nonalcoholic fatty liver disease (NAFLD) such as steatosis and nonalcoholic steatohepatitis (NASH) in populations is continuing to grow vigorously and became a worldwide public health issue. To understand the liver disease progression one needs to investigate complex interactions occurring within biological systems. Systems biology tries to understand the interactions within biological systems by means of mathematical models. Exploiting this approach I want to describe interactions of genes, proteins and metabolites that are involved in nonalcoholic fatty liver disease. Molecular data from liver tissue samples of three mouse strains (A/J, C57Bl6 and PWD) under two different conditions: 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC)-treated and untreated (control) were analyzed. Each of these mouse strains shows different degrees of the disease under DDC treatment displaying high, low, and no steatohepatitis-like phenotypes for A/J, C57Bl6 and PWD, respectively.

In this work I performed pathway analysis using gene expression data of mouse liver samples and identified metabolism of histidine, beta-alanine, purine along with glycolysis and gluconeogenesis pathways as top hit candidates that may be involved in liver dysfunction. Furthermore, gene expression and metabolite data of the arachidonic acid metabolism were found to be deregulated and this pathway was used for kinetic modeling. Genes and metabolites of S-adenosylmethionine (SAMe) metabolism were found to be perturbed under DDC-treatment. In addition, I have developed a novel enrichment analysis approach that may be used for identification of the most relevant fully coupled modules in a disease context. The approach includes three steps: 1) obtain fully coupled reactions which represent a module, 2) use gene expression data of a disease context to obtained marked correlated modules and 3) select modules in which at least one gene is differentially expressed between normal and disease conditions. Aforementioned steps are used to identify liver disease specific modules such as modules of pentose phosphate pathway and hepatic SAMe metabolism which are linked to oxidative stress. Furthermore, I also identified a module of cholesterol metabolism which is linked to apoptosis along with a module of pyrimidine catabolism, for which experimentally measured genes and metabolites were also found to be deregulated. The goal was to identify modules for which genes and metabolites are perturbed under DDC-treatment. The identified modules may be involved in liver disease and they can be used to build kinetic models for better understanding of the liver disease progression. Thus, in addition to enrichment analysis of fully coupled modules, I developed an approach which is based on elementary flux modes (EFMs). In this approach initially differentially regulated metabolites due to DDC-treatment were identified. Then reactions which can produce differentially regulated metabolites were

used as a target set. For each reaction in the target set, 50 EFMs that contain the reaction were identified. After that, gene expression data of mouse liver samples were used to select important EFMs that may be involved in the liver disease progression. I identified two EFMs: one EFM comprises differentially regulated metabolites L-arginine, ornithine and putrescine and another EFM comprises differentially regulated metabolites D-glucose, L-glutamine and L-asparagine.

I introduced a mGX-FBA method which is a modified version of the previously published GX-FBA method. Differences between metabolic flux levels among mouse strains may provide a better understanding of the reasons behind NAFLD. To address this, I performed an *in silico* flux-based analysis using E-Flux and the modified mGX-FBA method by incorporating gene expression data of the mouse model. Furthermore, during the course of my thesis I compared the results of both methods, E-Flux and mGX-FBA, and validated the results with experimental data. The change of flux through metabolic pathways may change metabolic concentrations. Due to the absence of experimental flux data for mouse liver samples it is difficult to assess *in silico* predicted flux regulation. However, *in silico* flux regulation may give a hint about the regulation of metabolic concentrations. Hence, to observe the flux regulation I colored metabolic maps with the deviation of fluxes between DDC-treated versus control. Different degrees of flux regulation was identified through cholesterol biosynthesis among all three mouse strains. The concentration of desmosterol that is a downstream metabolite of cholesterol biosynthesis was found to be regulated at different degrees. The regulation of desmosterol concentration is inline with the flux regulation of cholesterol biosynthesis among all three strains. In addition, to understand whether one can speculate about the prediction of flux regulation in metabolic pathways based on gene regulation I used bile acid synthesis and cholesterol biosynthesis pathways. For these pathways, I observed that based on gene expression data it is difficult to estimate flux regulation, but the integration of gene expression data using E-Flux improves the prediction of flux regulation.

I introduced novel objective functions for measuring the readout of steatosis and steatohepatitis. To construct an objective function for steatosis I used metabolites which are involved in the formation of lipid droplets (LDs), while for steatohepatitis an objective function integrating both metabolites that are involved in LDs formation and metabolites involved in oxidative stress was used. Gene expression data of liver samples of all three mouse strains were incorporated to a mouse metabolic model and obtained objective values were validated with strain's phenotypic data. I also constructed an objective function with metabolites which may be involved in cell proliferation. Cell proliferation is used as a readout of hepatocellular carcinoma (HCC). Applying these functions I performed an *in silico* drug target analysis in which

potential drug candidates for steatosis, steatohepatitis and HCC were identified. Cholesterol metabolism and triacylglycerol synthesis were found as top hits that contain the largest number of potential drug target candidates. Out of 78 identified potential drug candidates, 7 were found to be approved by the Food and Drug Administration (FDA) as anticancer drugs.

Metabolite concentrations can be viewed as end points of perturbations occurring at the gene level, so that changes of gene expression might explain changes in metabolite concentrations. I proposed a novel hypothesis to predict changes in metabolite concentrations between two conditions based on gene expression data. To address this, I have developed a Petri net-based method (MPN) and used it to simulate the arachidonic acid model. As an alternative to MPN, Monte Carlo ODE-based simulation was used, but both methods cannot predict the metabolic concentrations in the real range of experimental data. To overcome this, I have developed a fitted detailed kinetic model of the arachidonic acid metabolism that comprises metabolites that are markedly deregulated due to DDC-treatment in all three mouse strains.

Zusammenfassung (German Summary)

Die Rate der nichtalkoholischen Fettlebererkrankungen (NAFLD), wie Steatose und nichtalkoholische Steatohepatitis (NASH) nimmt kontinuierlich zu und ist zu einem weltweiten Gesundheitsproblem geworden. Um die Entstehung dieser Lebererkrankungen besser zu verstehen bedarf es einer detaillierten Untersuchung der komplexen Interaktionsnetzwerke lebender Systeme. Systembiologie versucht die komplexen Netzwerke biologischer Systeme mit Hilfe mathematischer Modelle besser zu verstehen. In der vorliegenden Arbeit versuche ich unter Verwendung dieses Ansatzes die Interaktionen zwischen Genen, Proteinen und Metaboliten in der nichtalkoholischen Fettlebererkrankung besser zu beschreiben. Für diese Analyse lagen molekulare Daten zu Leberproben dreier Mausstämme (A/J, C57Bl6 und PWD) vor, die einer Behandlung mit DDC (3,5-diethoxycarbonyl-1,4-dihydrocollidine) unterzogen wurden, die als Modellsystem für die nichtalkoholische Fettlebererkrankung dient, wobei die drei Mausstämme unterschiedliche Ausprägungen der Erkrankung zeigen, die von hoch in A/J über niedrig in C57Bl6 bis zu keiner Steatohepatitis in PWD reicht.

Im Rahmen dieser Arbeit habe ich auf Basis entsprechender Transkriptomdaten eine Pathway-Analyse durchgeführt, bei der unter anderem der Arachidonsäurestoffwechsel als dereguliert gefunden wurde, was auch durch weitere Metabolitdaten untermauert werden konnte, und daher Gegenstand einer entsprechenden mathematischen Modellierung war. Des Weiteren wurden Transkriptom- und Metabolitdaten des S-Adenosylmethionin (SAMe)-Metabolismus als dereguliert durch DDC-Behandlung gefunden. Ferner habe ich eine neue Methode zur Anreicherungsanalyse entwickelt, die für die Identifizierung vollständig gekoppelter Module genutzt werden kann. Der Ansatz beinhaltet drei Schritte: 1) Identifizierung vollständig gekoppelter Reaktionen, die ein Modul darstellen, 2) die Filterung dieser Module auf Basis von Genexpressionsdaten, und 3) die Selektion relevanter Module, in denen wenigstens ein Gen differentiell reguliert ist zwischen gesundem und krankem Zustand. Dieser Ansatz wurde verwendet um spezifische Module der Lebererkrankung im Modellsystem zu finden. So wurde z.B. ein Modul des Pentosephosphatwegs oder des SAMe-Metabolismus identifiziert, die mit oxidativem Stress assoziiert sind. Das Ziel war es somit hier Module zu identifizieren, die sowohl in den Expressionsdaten als auch in den metabolischen Daten im DDC-Modellsystem verändert sind, und somit am Krankheitsprozess beteiligt sein können. Die identifizierten Module tragen daher zu einem besseren Verständnis der Krankheitsentstehung bei und können als Ausgangspunkt für weitere kinetische Modellierungen dienen. Zudem habe ich einen entsprechenden Ansatz auch auf Basis von elementaren Flussmodi entwickelt.

Des Weiteren habe ich eine modifizierte Form der auf der Flux-Balance-Analyse basierenden GX-FBA Methode entwickelt. Änderungen im metabolischen Flussverhalten können zu einem besseren Verständnis der Ursachen von NAFLD beitragen. Hierfür habe ich eine vergleichende Flussanalyse mit Hilfe der Methoden E-Flux und der von mir modifizierten GX-FBA Methode auf Basis der vorliegenden Genexpressionsdaten durchgeführt und diese Methoden diesbezüglich auch verglichen. Mit dieser Analyse konnten unter anderem Unterschiede im Cholesterol- und

Gallensäure-Metabolismus identifiziert werden.

Um den Grad von Steatose und Steatohepatitis *in silico* beurteilen zu können, habe ich neue Zielfunktionen definiert, die die Bildung von Lipidtröpfchen berücksichtigen. Des Weiteren wurde eine Zielfunktion erstellt, die den Grad der Zellproliferation beschreibt und als Maßstab für die Leberkrebsentstehung verwendet werden kann. Diese Zielfunktionen wurden genutzt, um neue Ziele für potentielle Medikamente für Steatosis, Steatohepatitis oder Leberkrebs zu identifizieren. Von den 78 hierbei identifizierten Kandidaten gibt es schon zu 7 entsprechende Medikamente, die von der FDA (Food and Drug Administration) zugelassen sind.

Änderungen in den Konzentrationen von Metaboliten können als Folge einer geänderten Genexpression interpretiert werden. Um diesen Zusammenhang für die qualitative und quantitative Modellierung zu nutzen, schlage ich hier eine auf einer Petri Netz-Simulation beruhenden Analyse vor und vergleiche diese mit einer Monte Carlo-basierten am Beispiel des hier entwickelten Arachidonsäurestoffwechselmodells, sowie einem Modell, bei dem die kinetischen Parameter mit Hilfe einer Parameteroptimierung an die experimentellen Daten angepasst wurden. Dabei zeigt das parameteroptimierte Modell deutlich die besten Vorhersagen.

1 Introduction

Liver diseases, such as nonalcoholic fatty liver disease (NAFLD) and alcoholic liver disease (ALD) are very common worldwide. NAFLD includes a disease spectrum ranging from steatosis, to nonalcoholic steatohepatitis (NASH), to advanced fibrosis and cirrhosis. ALD develops due to alcohol abuse and it runs a spectrum similar to NAFLD. NAFLD affects 10 to 24% of the population in various countries (Angulo, 2002). The prevalence of NAFLD varies with ethnicity (45% in Hispanics; 33% in whites; 24% in blacks) and sex (42% in white men; 24% in white women) in Western adults (Browning et al., 2004). The prevalence increases 57.5% to 74% in obese persons along with 90% in morbidly obese persons (those weighting more than 200 % of their ideal body weight) (Angulo, 2002; Machado et al., 2006). NAFLD affects 2.6% of children and 22.5% to 52.8% of obese children (Angulo, 2002). NASH, an advanced form of NAFLD is less common in the normal population. The prevalence of NASH is estimated to be 2% to 3% in adults, and 37% in morbidly obese (Dowman et al., 2011; Neuschwander-Tetri, 2005). NAFLD is closely associated with obesity and insulin resistance. Accordingly, it represents the hepatic manifestation of the metabolic syndrome. Approximately 20% of alcoholics and heavy drinkers develop fatty liver, or steatosis (Mann et al., 2003). Fatty liver can be reversed by reducing alcohol consumption substantially. After excessive use of alcohol the fatty liver shows inflammatory changes that are an early sign of more serious liver disease, and may lead to death. Both NAFLD and ALD have become issues of public health due to these serious health concerns.

Complex diseases such as NAFLD or cirrhosis develop as a result of the interaction between the human genome and the environment. The development of complex diseases is difficult to understand based on individual genetic variants of the human genome. In contrast, genetic variation is usually causative of Mendelian disorders. Complex disease phenotypes are controlled not only by genes but by the dynamics of interaction networks of genes, proteins and metabolites (Strohman, 2002). These networks range from signaling to metabolic pathways regulating enzymes and metabolites. Thus, to understand the progression and development of complex diseases one needs first to understand complex interactions of biological systems instead of single genes in isolation. As such, systems biology can be seen as an approach to understand the complex behavior of biological systems.

The term systems biology was introduced a decade ago to understand biology according to a system's structure and dynamics (Kitano, 2002a). In recent years, the field has made considerable progress in technologies for global cell measurement, as well as in computational analysis of these data in order to study and model cell function. The ability to make genome-wide (proteome- or transcriptome-wide) measurements on a system is the driving force behind the rise of systems biology. Only genome-wide measured data is not sufficient to study the real biological systems. Thus, systems biology comprises methods which can be used to study with a systems hypothesis-driven approach for the experimental

design (Ideker et al., 2001).

Detailed knowledge of signaling and metabolic pathways interactions is needed for many fields relevant to medical research, such as hepatology and cancer biology. Quantitative proteomics enables a broader and more unbiased view of the cellular signaling event. Proteomics data was used to model the epidermal growth factor receptor (EGFR) signaling network intended for proliferation and migration (Wolf-Yadlin et al., 2006). Furthermore, diagnostic biomarkers were identified based on profiling gene and protein expression for complex diseases such as cancer and steatohepatitis (Cheang et al., 2008; Quackenbush, 2006; Starmann et al., 2012). Therefore, high-throughput data, pathway related information (Croft et al., 2011; Kanehisa et al., 2012), and computational algorithms open an approach to view human diseases at the system-level. To address this approach, networks of biological reaction systems can be created by integrating genetic, genomic, biochemical, cellular, physiological, and clinical data that can be used to understand disease progression and therapy. In order to understand disease therapy, one needs not only to define the topology of disease related networks or modules, but also to explore the dynamic response to perturbations.

A disease might be a cause of abnormalities in genomic or environmental factors; however, currently, only approximately 10% of the known genes can be associated with a disease (Amberger et al., 2009). Mutations of hub genes or proteins can be responsible for a disease. Disease related genes are likely to form a subnetwork which is formed by disease modules (Loscalzo and Barabasi, 2011). These modules can be identified computationally using experimentally data (Loscalzo and Barabasi, 2011). Most biological networks show a high degree of clustering that implies the existence of topological modules. Modules can be obtained based on the similar structure and function among elements of the global network. However, computing modules is computationally challenging (Ahn et al., 2010; Girvan and Newman, 2002).

Systems biology is the study of the interaction between potentially thousands of genes as a whole system instead of a single gene. We need to perform a quantitative study to understand the mechanisms of components interacting within the regulatory system. Such a study would rely on an experimental design of complex biological systems. The interpretation of experimental results requires mathematical particularly often statistical approaches. The data obtained by large-scale experiments, such as mRNA microarrays or protein-protein interaction screening can be analyzed using statistical methods. These statistical methods provide initial insight into large systems, e.g., how many genes that are differentially regulated can be found by statistical testing methods. A cellular network constructed according to graph theory provides a good starting point for mapping large-scale experimental data within the cells. The undirectional (the link has no direction between the nodes) network is less informative, whereas the directional (the link has a direction between nodes, e.g. A ->B) network can be used to identify motifs, such as feedforward or feedback

loops, and to infer regulatory mechanisms. However, these statistical and graphical methods are limited to studying the dynamics of cellular regulation. To this end, predictive models are much needed to understand the regulation of the system.

Predictive models can be built using quantitative data on cellular levels, such as the concentration of proteins and metabolites, and on the reaction rates for interactions between components. Quantitative biochemical models are difficult to build because of the limitations of experimental biochemistry and quantitative molecular biology data. Metabolic concentrations and other components are likely to vary in different cell types. In order to model this, it is necessary to measure concentrations of those components in multiple cell types. Similarly, kinetic parameters require us to determine them under multiple conditions. Although there is information on metabolic concentrations (Wishart et al., 2007) and kinetic parameters (Scheer et al., 2011; Wittig et al., 2012), this information is limited to model complex biological systems. In short, predictive models are needed to understand a biological system, but are least likely to be possible due to the often limited availability of sufficient experimental data.

An approach is proposed to study the parameter space of the model using a Monte Carlo-based simulation in which unknown kinetic parameters are sampled from an appropriate distribution for simulations (Wierling et al., 2012). This approach may be used to predict complex and variable phenotypes from genotypes in personalized medicine. Furthermore, kinetic parameters can be estimated or optimized using an objective function of quantitative experimental data. Therefore, optimization methods can be used to build realistic and reliable mathematical models (Raue et al., 2013); however, these are challenging in computation and limited to small-scale modeling. Moreover, a Petri net-based non-parametric strategy is proposed for characterizing the dynamics of a signaling network (Ruths et al., 2008) which may be used for large-scale modeling but still remains to be tested.

Constraint-based modeling is a mathematical approach for analyzing the flow of metabolites through a metabolic network (Varma and Palsson, 1994; Orth et al., 2010). It can be applied to genome-scale metabolic networks (Schellenberger et al., 2010). Furthermore, it can be used to identify coupled reactions (David et al., 2011; Larhlimi et al., 2012) in the metabolic network. These coupled reactions can be seen as a module that can be used to explore a complex disease and its related data. Taken together, one can study the development and therapy of a complex disease, such as NAFLD, using several systems biology approaches. Furthermore, several studies can be done in the context of NAFLD, e.g, identification of deregulated modules and pathways, in addition to metabolic drug targets.

Outline

In this work I present liver functions and related diseases (section 1.3 and 1.4). In the following section I will outline the experimental data and systems biology approach for studying biological systems (section 1.5 and 1.6). Later, I present the computational

modeling approaches.

Furthermore, I present published methods in detail such as computation of shortest elementary flux modes (section 2.3), integration of gene expression data into a metabolic model (section 2.4 and 2.5), simulation based on Petri nets (section 2.8). I introduce objective functions for the readout of steatohepatitis and steatosis (section 2.6).

Pathway analysis is performed using gene expression data of mouse liver samples (section 3.1). I introduce modules identification in metabolic networks using fully coupled reactions as well as elementary flux modes (section 3.2). Furthermore, to predict steady state fluxes I performed the integration of gene expression data into a metabolic model (section 3.3). I identified *in silico* potential candidates of drugs in liver disease through metabolic models (section 3.4). Arachidonic acid metabolism is identified as a candidate in pathway analysis and it is used for kinetic modeling (section 3.5). Results of section 3 are compared with other studies in section 4.

1.1 The Liver

The liver is an important organ that is involved in metabolic and clearance functions. It fulfills diverse but closely connected functions in the metabolism of carbohydrates, proteins and lipids, the clearance of toxins and pathogens, and the regulation of immune response. The food we eat is eventually processed through the liver before the body can use it; our food is broken down into small particles/molecular subunits, e.g., fats, proteins and carbohydrates. Proteins can be involved in the catalyzation of metabolic reactions. The liver stores fuels for our body, e.g., storage of glycogen, fat, cholesterol, minerals, vitamins and carbohydrates. The liver sends these fuels through our blood for use as energy and in building healthy tissue. Finally, the liver makes bile which dissolves fats.

The liver lobule is the structural unit of the liver. The entire liver consists of thousands of lobules. Each lobule consists of a roughly hexagonal arrangement of hundreds of hepatocytes and other types of cells (Figure 1.1). Nutrients- and oxygen-rich blood enters the liver lobule from the hepatic portal vein and hepatic artery, and is then distributed throughout the lobule. Within the lobules blood flows between the hepatocytes through small capillary channels called sinusoids. Toxic substances from the blood exit the lobule through the central vein (i.e., the hepatic venule). The blood close to the hepatic portal vein and hepatic artery has higher nutrients and oxygen content than blood close to the hepatic venule. This is due to much of the oxygen and nutrients already having been extracted from the blood by the hepatocytes close to the hepatic portal vein and hepatic artery. As a result, the blood close to hepatic venule contains less oxygen and nutrients (Cunningham and Van Horn, 2003).

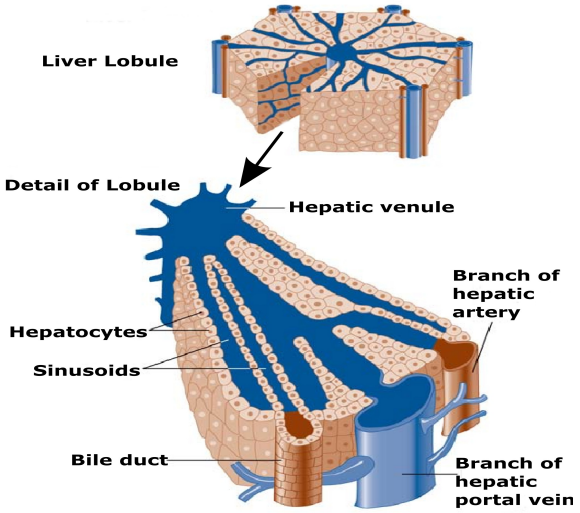

Figure 1.1 The structure of the liver's functional units, or lobules (from: Cunningham and Van Horn, 2003). Blood enters the lobules through branches of the portal vein and hepatic artery, then flows between the hepatocytes through small channels called sinusoids. Blood exits the lobules through a small vein located in the center of each lobule called the hepatic venule.

Some of the better well known functions of the liver include the following:

- Regulation of glucose metabolism
- Regulation of lipoprotein metabolism
- Regulation of levels of amino acids in blood, which form the building blocks of proteins
- Processing of hemoglobin for use of its iron content
- Conversion of poisonous ammonia to urea
- Clearing the blood of drugs and other poisonous substances

The regulatory mechanisms concerning glucose and lipoprotein metabolism are of a particular interest due to these pathways and modules are identified in our analysis. Therefore, a detailed overview of both glucose and lipoprotein metabolism is given below.

1.1.1 Regulation of glucose metabolism

Blood glucose in the human body maintains a narrow range between 4 to 7mM (Saltiel and Kahn, 2001). Insulin and glucagon are the hormones that control glucose homeostasis. In response to high glucose levels, insulin is released from pancreatic beta cells and converted to an active form in the blood. Insulin stimulates glycogen accumulation through the increase in glucose transport (via Glut2) and glycogen synthesis (Figure 1.2). In case of high level of glucose, insulin inhibits the production of glucose by repressing gluconeogenesis (pyruvate to glucose) and glycogenolysis (glycogen to glucose). In contrast, low glucose levels cause secretion of the pancreatic hormone glucagon from alpha cells. Glucagon stimulates the

conversion of the liver glycogen to glucose and the release of glucose back into the blood. Glucose can also be generated by gluconeogenesis during periods of starvation and intense exercise. The disruption of glucose homeostasis can cause diabetes, a metabolic syndrome in which patients do not produce balanced levels of glucose.

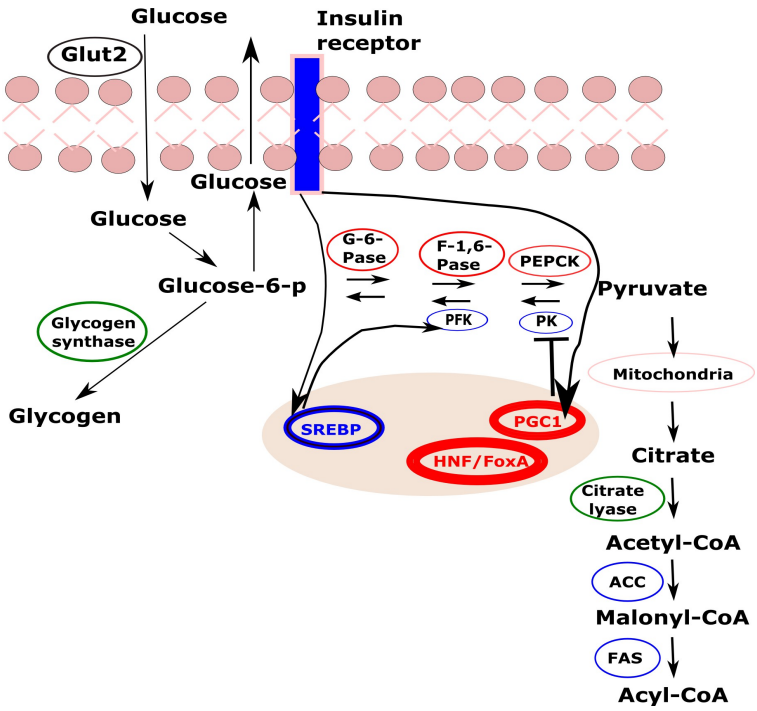

Figure 1.2 The regulation of glucose metabolism by the liver (modified after Saltiel and Kahn, 2001). In the liver, insulin stimulates the synthesis of glycogen from glucose, while repressing glucose synthesis and release. Insulin stimulates the expression of genes encoding glycolysis (in red; G-6-Pase, F-1-6-pase and PEPCK), while inhibiting the expression of those encoding gluconeogenesis enzymes (in blue; PK and PFK). Insulin stimulates enzymes of fatty acid pathway (in blue; FAS). These effects are determined by transcription factors and co-factors, i.e. sterol regulatory element binding protein (SREBP-1), hepatic nuclear factor (HNF-4), the forkhead protein family (Fox), Peroxisome proliferator-activated receptor gamma coactivator (PGC1). Insulin also regulates the activities of some enzymes such as glycogen synthase and citrate lyase (in green), through changes in the phosphorylation state. GK, glucokinase; Glucose-6-p, Glucose-6-phosphate; G-6-Pase, Glucose-6-phosphatase; F-1,6-Pase, fructose-1,6-biphosphatase; PEPCK, phosphoenolpyruvate carboxykinase; PFK, phosphofructokinase; PK pyruvate kinase; ACC, acetyl-CoA carboxylase, FAS, fatty-acid synthase.

1.1.2 Regulation of lipoprotein metabolism

The lipoproteins transport lipids, including cholesterol, to blood streams. The liver plays an important role maintaining cholesterol levels through the lipoprotein metabolism. The regulation of lipoprotein metabolism involves three steps (Figure 1.3). First, we see the metabolism chylomicrons in which chylomicrons are formed from dietary cholesterol and the intestine triglyceride. Chylomicrons form chylomicron remnants after secretion of fatty acids. Subsequently, fatty acids and chylomicron remnants are uptaken by the peripheral tissues and the liver, respectively. Second, the liver secrets very low density lipoproteins (VLDLs) after loading lipids and cholesterol to the apolipoprotein B (apoB), and VLDLs are then converted to low density lipoproteins (LDLs). LDLs can be uptaken by the liver through the LDL receptor. When the free cholesterol is found in higher amounts in peripheral tissues, then as a third step the reverse transport of cholesterol happens from peripheral tissues to the liver through high-density lipoprotein (HDL) metabolism.

1.2 Liver diseases

Hepatic diseases are a type of damage to or disease of the liver. Some kinds of liver diseases are caused by infective viruses, e.g., the Hepatitis virus (A, B, and C). Liver disease can happen due to long-term drug or alcohol intake. Some examples of alcohol related diseases are alcoholic hepatitis and alcoholic cirrhosis. Fatty liver disease can also occur with very little alcohol intake, and then is called nonalcoholic fatty liver disease (NAFLD). NAFLD represents a disease progression ranging from simple steatosis to NASH, fibrosis, cirrhosis and eventually liver cancer.

1.2.1 Steatosis and steatohepatitis

Steatosis, also known as fatty liver, is a process characterized by the accumulation of fat in the liver. Hepatic steatosis is a condition in which intracellular lipids are accumulated and subsequently form lipid droplets in the cytoplasm of hepatocytes which is associated with an enlargement of the liver. When steatosis of the liver is further combined with inflammation, the condition is known as steatohepatitis. Both pathological conditions are included under the term non alcoholic fatty liver disease (NAFLD) if alcohol can be excluded as a primary cause.

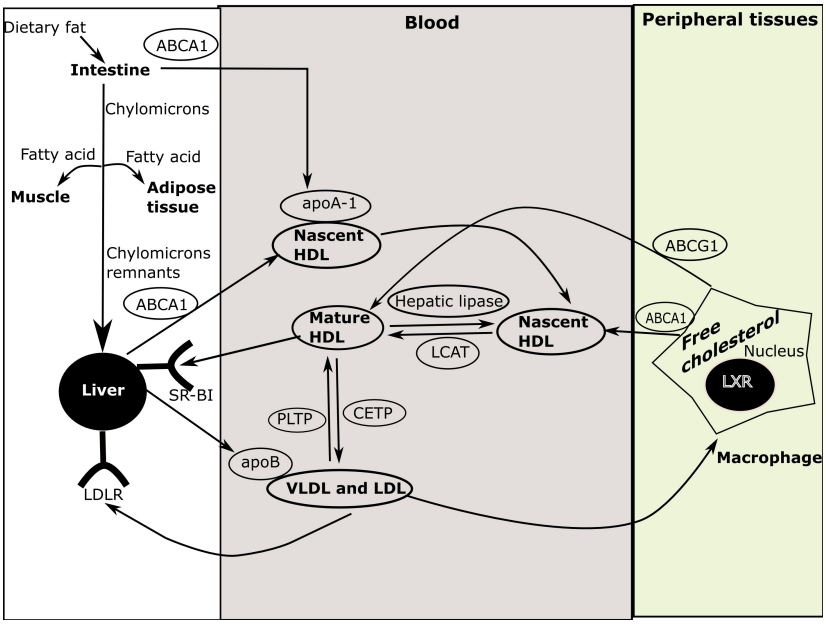

Figure 1.3 Regulation of lipoprotein metabolism (modified after Rader and Daugherty, 2008). Lipoprotein metabolism involves the transport of lipids, e.g., cholesterol and triglyceride, in the blood. Dietary fat from the intestine forms chylomicrons (large triglyceride-rich proteins), which are transported to peripheral tissue such as muscle and adipose tissue, through the blood. The chylomicrons form chylomicron remnants and secrete fatty acids to the peripheral tissue. The chylomicron remnants are taken up by the liver and bind to apoB. This complex secrete very-low-density lipoproteins (VLDLs), which further develop low-density lipoproteins (LDLs) by activation of lipoprotein lipase. LDLs are then taken up by the liver through LDL receptors (LDLR) and also it can pass through blood to other peripheral tissues. In contrast, high density lipoproteins (HDLs) are generated by the liver and the intestine through secretion of lipid free apoA-1. apoA-I then form nascent HDLs from cholesterols of these organs by action of the transporter ABCA1. In peripheral tissue and macrophage, ABCA1 transports free cholesterol to nascent HDLs, while ABCG1 transports free cholesterol to mature HDLs. The free (unesterified) cholesterol in nascent HDLs is esterified to cholesteryl ester by the enzyme lecithin cholesterol acyltransferase (LCAT), creating mature HDLs. The cholesterol in HDLs is returned to the liver both directly, through uptake by the receptor scavenger receptor class B member 1 (SR-BI), and indirectly, by transfer to LDLs and VLDLs through the cholesteryl ester transfer protein (CETP).

An incidence of NAFLD of 15 to 20% is observed in western populations, and its most common cause of development is overnutrition (Amarapurkar et al., 2007; Bedogni et al., 2005). Overnutrition- and obesity-related NAFLD is observed in patients with metabolic syndrome, e.g., obesity and insulin resistance (Higuchi and Gores, 2003). It is remarkable that the prevalence of steatosis in patients with obesity is about 75% (Adams et al., 2005; Browning et al., 2004), and nearly 35% of these develop NASH (Ong et al., 2005; Xanthakos et al., 2006). It is noteworthy that a considerable fraction of patients with steatohepatitis will progress to advanced disease stages, e.g., fibrosis, cirrhosis and hepatocellular carcinoma.

Typically, middle-aged patients have an increased risk of developing NAFLD. In NAFLD patients, > 80% of patients are overweight or obese, 80% have hyperlipidemia, 30-70% are hypertensive and 20% have type 2 diabetes mellitus (Grattagliano et al., 2007). There do exist pharmacological means to counteract this development, which are weight loss or an improvement of insulin sensitivity. Nevertheless, these treatments are only partially effective (Mishra and Younossi, 2007). Due to absence of effective treatment for NASH, the disease is still to date the subject of studies on morbidity and mortality. The diagnosis of NASH is the following: 12% death within 10 years (Schuppan et al., 2010). Clearly, liver diseases such as NASH, are an important and growing public health issue. Cirrhosis due to NASH may now account for up to 20% of cirrhosis cases and may also play a major role in cryptogenic cirrhosis. Furthermore up to 10% of hepatocellular carcinomas may result from NASH (Angulo, 2002; Bugianesi et al., 2002; Day, 2005; Farrell and Larter, 2006).

The molecular mechanism of intrahepatic lipid accumulation and growth of lipid droplets may arise from 1) increased uptake of lipids, 2) elevated *de novo* synthesis of fatty acids, 3) impaired lipoprotein synthesis, or secretion and /or 4) reduced fatty acid oxidation (Anderson and Borlak, 2008). In this regard, the liver plays a central role in energy homeostasis by storing glucose as glycogen and distributing fuels to peripheral tissue in form of glucose and lipids. Dietary lipids in form of chylomicrons are transported from the intestine to the liver where they are processed with apolipoprotein B 100 to form VLDL. These particles are secreted and then distribute lipids to adipose tissue. Consequently, dietary increase of lipids and glucose may result in hepatic lipid accumulation (Bray et al., 2004).

One of most exciting questions in this study pertains to why and how hepatic lipid accumulation leads to the development of liver inflammation. A potential subject of research to study might look for the factors responsible for switching between steatosis and steatohepatitis. To provide pathophysiologic principles for the progression to steatohepatitis, the ’two-hit’ model is proposed (Day and James, 1998). The ’First-hit’ refers to the intracellular deposition of triacylglycerols which leads to metabolic and molecular alteration that sensitizes the liver, while the ’second-hit’ refers to oxidative stress and cytokine-induced liver injury. Lipid accumulation and short- and long-term effects of non-esterified fatty acids (NEFA) increase oxidative stress, cellular stress, e.g., activation of protein kinase C (PKC), mitogen-activated protein kinase (MAPK), jun N-terminal kinase (JNK), nuclear factor-κB (NF-κB), and expression of pro-inflammatory cytokines, such as tumor necrosis factor α (TNF-α) and interleukin 6 (IL-6) (Reddy and Rao, 2006). In conclusion, steatosis of the liver may arise from an excess supply of fatty acids and/or glucose, and insulin resistance; its progression to NASH is linked to the paracrine effect of pro-inflammatory cytokines (Anderson and Borlak, 2008).

1.2.2 Histological features of nonalcoholic steatohepatitis

Histology is an anatomical study of the microscopic structure of cells and tissues of plants and animals. It is an essential tool in biology and medicine. In a histological study, cells and tissues are sectioned, stained and examined under a light microscope or electron microscope. Hematoxylin and eosin staining (H & E stain or HE stain) is a popular method in histology. It is widely used by pathologists to examine the biopsy of patients related to diseases such as cancer and steatohepatitis. Histologic features of NASH are a) mixed macrovesicular and microvesicular steatosis, b) zonal distribution of steatosis (zone three) in NASH, c) foci of lobular inflammation composed of lymphocytes, d) Mallory's hyaline stained with immunohistochemistry for Ubiquitin, e) nuclear vacuolation in hepatocytes, and f) Apoptosis (Singh et al., 2010).

1.2.3 Biological model systems for nonalcoholic fatty liver disease

The mechanism of progression of NAFLD/NASH has not yet been completely elucidated (Cohen et al., 2011). This may be a reason for the availability of very few therapeutic modalities for NAFLD/NASH that include weight loss agents, bariatric surgery, insulin-sensitizing agents, lipid-lowering agents, antioxidants, probiotics and other agents (Lam and Younossi, 2010). Applying this study of NAFLD/NASH to humans has limitations, however, due to genetic variation among individuals of various ethnic groups and long time required for the occurrence of NAFLD/NASH. Thus, large populations and long periods of time are required for the study incidence of NASH/NAFLD in humans. To overcome this problem, animal models of NAFLD/NASH can be used to demonstrate the pathogenesis of NAFLD/NASH along with therapeutic effects of various agents.

Histological features of NASH, such as steatosis, ballooning of hepatocytes, apoptosis, and Mallory bodies, can be reproduced to a variable extent in different mouse models by various treatments like chronic intoxication with 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC), methionine-choline deficient (MCD) diet, alcohol or high fat diet (HFD) (Anstee and Goldin, 2006; Denk et al., 2000). Previous studies demonstrated that the disease phenotypes of these animal models were influenced by the genetic background (Haluzik et al., 2004; Kirsch et al., 2003).

1.2.4 Molecular causes of steatosis and steatohepatitis

Steatosis and steatohepatitis are associated with several molecular causes and mechanisms such as nuclear receptors, inflammation mediated by signaling pathways, oxidative stress, and lipotoxicity (Anderson and Borlak, 2008; Marra et al., 2008). Peroxisome proliferator-activated receptor alpha (PPARα), a nuclear receptor, has been identified as a master regulator of liver inflammation and for the metabolism of hepatic lipid metabolism, glucose, lipoprotein and amino acids (Rakhshandehroo et al., 2010). Some molecular causes are

described below.

1.2.4.1 Nuclear receptors and transcription factors

Fatty acid is a metabolic regulator of fatty acid oxidation, and also is known to act as a ligand for nuclear transcription factors, i.e., PPARα and hepatic nuclear factors (HNFs) (Rasmussen and Wolfe, 1999). The lipid metabolism is regulated by major transcription factors, i.e., sterol regulatory element binding protein 1 (SREBP-1), PPARα, carbohydrate-responsive element-binding protein (ChREBP), lipogenic liver X receptor (LXR), forkhead box 01 (Foxo1), and PPARγ, while the lipoprotein metabolism is controlled by apolipoprotein AI regulatory protein-1, v-*erbA*-related protein (EAR-2), EAR-3, and HNF-4 (Canbay et al., 2007; Cha and Repa, 2007; Ide et al., 2003; Ladias et al., 1992). PPARs are the key player behind lipid accumulation. For instance, liver-specific expression of PPARα is activated by different ligands as xenobiotics (e.g., fibrates), eicosanoids, and fatty acids (Mehendale, 2000; Motojima and Hirai, 2006). PPARα shows a negative regulation of inflammation, possibly inhibiting the translocation of nuclear factor-κB (NF-κB) to the nucleus by interacting with p65 (Delerive et al., 2001). PPARα can prevent the induction of pro-inflammatory cytokine and enzyme expression, i.e., tumor necrosis factor α (TNF-α) and cyclooxygenase (COX) II through the inhibition of NF-κB (Yu et al., 2006). A deactivation of PPARα may reduce the liver's ability to uptake, utilize, and catabolize fatty acids and thereby it may be causally involved in the development of steatosis (Rakhshandehroo et al., 2010; Reddy, 2001).

1.2.4.2 Lipotoxicity

Lipotoxicity is a metabolic syndrome that refers to a cellular dysfunction due to intracellular accumulation of lipids. Elevated levels of free fatty acid cause inflammation observed in adipose tissue, obesity, and NASH, and that is characterized by elevated plasma levels of TNF-α (Crespo et al., 2001; Kern et al., 1995; Valenti et al., 2002). Fatty acid induces synthesis of ceramide via L-serine and palmitoyl-CoA, catalyzed by serine palmitoyltransferase (Unger and Orci, 2002). Ceramide upregulates the expression of inducible nitric oxide synthase (iNOS), which leads to apoptosis. Lipoapoptotic cell death provides an explanation for elevated apoptosis rates in NASH and links hepatic lipid accumulation to inflammation. As for the involvement of free fatty acids in activation of TNF-α, a hypothesis is proposed (Feldstein et al., 2004) that suggests a high-fat diet causing the accumulation of lipids and translocation of the pro-apoptotic factor bax to lysosomes. Bax induces a release of lysosomal cysteine protease cathepsin B (cath B), which is responsible for degradation of intercellular adhesion molecule (IKK-β) and the activation of NF-κB, and which subsequently via NF-κB induces TNF-α (Figure 1.4).

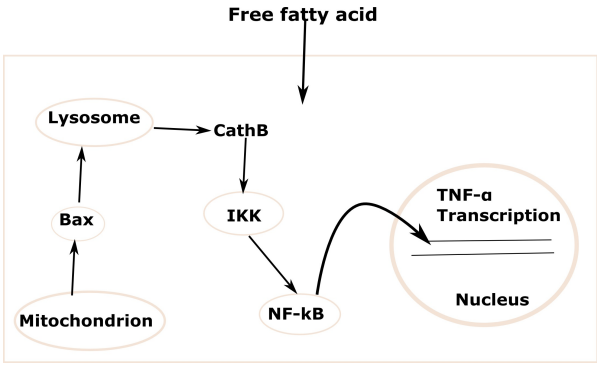

Figure 1.4 Mechanism of lipid formation. A simplified scheme of the mechanism by which lipids lead to activation of TNF-α (Anderson and Borlak, 2008).

1.2.4.3 Oxidative stress

In the 'two-hit' model, the first hit represents lipid accumulation, while the second hit represents an increase in oxidative stress (Day and James, 1998). Oxidative stress can be caused by reactive oxygen species (ROS) (Pessayre et al., 2002). ROS are chemically reactive molecules containing oxygen, e.g., oxygen ion and peroxide. Either increases ROS directly or secondarily (impaired respiration leads to the formation of ROS) can be a cause for the development of steatohepatitis. Higher ROS formation in steatohepatitis can cause more lipid peroxidation, cytokine induction, and fibrogenesis than in steatosis (Pessayre et al., 2002). Excessive generation of ROS may affect the maintenance of an appropriate redox environment to avoid oxidative modification in proteins, lipids and DNA within mitochondria (Garcia-Ruiz and Fernandez-Checa, 2006). Mitochondrial Glutathione (mGSH) plays an essential role in the mitochondrial defense against constant ROS generation. Thus, a marked depletion of mGSH is a critical factor in the development of steatohepatitis through sensitization of hepatocytes to inflammatory cytokines (Garcia-Ruiz and Fernandez-Checa, 2006).

1.2.4.4 Inflammation mediated through signalling and metabolic pathways

Activation of the c-Jun N-terminal kinase (JNK) pathway can lead to lipotoxicity and pro-apoptotic mechanisms in NASH through inflammation and oxidative stress (Leamy et al., 2013). JNK stimulation has been observed in NASH patients as well as murine models of steatohepatitis (Puri et al., 2008; Singh et al., 2009).

Arachidonic acid (AA)-derived lipid, such as lipoxin (LXA_4) and aspirin-triggered LXA_4 (ATL) act as mediators to reduce inflammation and promote resolution (Schwab and Serhan, 2006). Biosynthesis of the LXA_4 and ATL involve the conversion of hepatocyte-derived 15-

hydroxyeicosatetraenoic acid by the sequential interaction of a 15-lipoxygenase (LO) and a 5-LO expressed in Kupffer cells (Clària and Planagumà, 2005). ATL regulates the expression of key transcriptional factors (e.g., peroxisome proliferator-activated receptor/PPAR) along with the cytokine–chemokine axes in hepatic cells (Clària and Planagumà, 2005). An altered availability of arachidonic acid may shift the balance of lipid signaling towards the production of inflammatory eicosanoids, thereby possibly contributing to the development of steatohepatitis (Anderson and Borlak, 2008).

1.3 Molecular data for modeling

In this study, several sources of experimental data such as genomics, proteomics, and metabolomics were used. Such omics data are useful to determine how the cell is perturbed by the actions of external signals (e.g. drug treatment), or how cell functions differ between a normal and a diseased state.

1.3.1 Method for gene expression analysis

Within the past decade high-throughput technologies for transcriptome analysis, such as DNA microarray and RNA-Seq, have been developed.

A DNA microarray is a collection of microscopic DNA spots attached to a solid surface. Using such array, the expression levels of thousands of genes can be measured simultaneously. First, thousands of spots known as probes are immobilized on a solid surface (e.g., a nylon filter or a glass slide). Next, the hybridization experiment is performed with the labeled mRNA molecules. After attaching to the reverse complementary sequence the amount of bound labeled material is quantified by a scanning device. Finally, spot intensities are transformed into numerical values that reflect the abundance of each specific probe.

In contrast to DNA microarrays, RNA-Seq, a sequence-based high-throughput approach that directly determines the cDNA amount, is a new method for mapping and quantifying transcriptomes (Wang et al., 2009). RNA-Seq can be used to measure transcriptome composition and to discover new exons or genes by direct ultra-high-throughput sequencing of cDNA (Mortazavi et al., 2008). In general, a library of cDNA fragments with adaptors attached to one or both ends, is prepared by conversion of a population of RNA (total or fractionated, such as poly (A)+) (Wang et al., 2009). A high-throughput manner sequences each molecule, with or without amplification to obtain short sequences from one end (single-end sequencing) or both ends (pair-end sequencing) (Wang et al., 2009). The readings can be 30-400 bp depending on the type of DNA-sequencing technology used (Wang et al., 2009). Some high-throughput sequencing systems used in the wet-lab are for instance Illumina IG, Applied Biosystems SOLiD and Roche 454 Life Science (Wang et al., 2009). The readings obtained after sequencing can be mapped to a reference genome or transcriptome, or assembled *de novo* without the genomic sequence to produce a genome-scale transcriptional

structure and/or level of expression for each gene (Wang et al., 2009).

RNA-Seq offers several key advantages over DNA microarrays. First of all, RNA-Seq can be used to analyze model organisms with existing genomic sequence as well as non-model organisms with genomic sequences that are yet to be determined. In contrast, a DNA microarray is limited to model organism, since you have to know transcripts that correspond to existing genomic sequence (Wang et al., 2009). For example, using 454 pyrosequencing data a *de novo* assembly of the Glanville fritillary butterfly transcriptome was presented, which had no previous genomic data (Vera et al., 2008). Second, RNA-Seq has very low background noise compared to DNA microarrays (Wang et al., 2009). Third, RNA-Seq can be used for more than an 8000-fold dynamic range to quantify gene expression level, while a DNA microarray has a much smaller dynamic range of 100 to a few hundred-fold (Wang et al., 2009).

1.3.2 Proteomics data using reverse-phase protein arrays

Protein arrays are designed to detect proteins, monitor their expression levels, screen for molecular markers and pathway targets, and investigate protein interactions and functions. In classical protein arrays only the probes are immobilized; however in a reverse phase protein array (RPPA) immobilizes the whole repertoire of proteins. RPPA can be used to measure proteins that represent the state of individual tissue cell populations undergoing disease transitions (Paweletz et al., 2001). The objective behind the RPPA development was to achieve sensitive high throughput protein analysis as well as precision and linearity (Paweletz et al., 2001). Reverse-phase protein array (RPPA) is a micro-format dot western blot that allows measurement of protein expression levels of hundreds to thousands of samples simultaneously in a qualitative manner (Spurrier et al., 2008). When specific antibodies are used, RPPA can generate 1000 times more data points using 10,000 times less sample volume than an ordinary western blot. RPPA enables us to monitor quantitative proteomic responses for various time-scale and input-dose gradients simultaneously. Hence, the RPPA method can be an excellent method for experimental validation of theoretical protein network models. The use of RPPA may make it possible to answer questions such as a) is the candidate molecule present or altered in the human disease?; b) is the alteration of the candidate molecule correlated with the molecular phenotype altered in a majority of patients?; c) can the altered candidate molecule in the disease state be chosen for treatment?

1.3.3 Metabolomics data using mass spectrometric analysis

Mass spectrometry (MS) is used in drug-metabolite analysis and metabolic research. The advantages of mass spectrometry include a wide dynamic range, the ability to observe a diverse number of molecular species, and reproducible quantitative analysis. Liquid chromatography mass spectrometry (LC-MS) is the most common approach for metabolite-profiling studies (Want et al., 2005). MS can be used from basic biochemistry to clinical

biomarker discovery. It generates a comprehensive profile and data analysis, and structurally characterize physiologically important metabolites.

1.4 Systems biology

The term “systems biology” was introduce about a decade ago, and calls for the examination of the structure and dynamics of biological systems by either perturbing their genes, proteins and pathways, or formulating a mathematical model using molecular informations that describes the structure of the system and its response to individual perturbations (Ideker et al., 2001; Kitano, 2002a). In recent years, systems biology has welcomed many new systematic technologies and methods in both experimental and computational areas that were not available a decade ago (Chuang et al., 2010).

A multitude of biological facts, such as genome sequences and protein properties has been discovered by molecular systems biology; however, these facts alone are not sufficient for an interpretation of biological systems (Kitano, 2002b). Genes, proteins, metabolites, cells, tissues, organs, organisms and ecological webs are components of an interacting system whose interactions have been defined by evolution (Kitano, 2002b). Thus, understanding at a system-level should be the prime goal of biology rather than studying a part of the system in isolation. In recent years, experimental approaches and technologies have been advanced in terms of accuracy, quantitative measurement and high-throughput, and this trend will doubtlessly continue. However, these methods are still limited to the study of intrinsic complexity in biological systems. Thus, a combined approach, computational biology, which comprises experimental and computational data is expected to resolve this problem (Kitano, 2002b). Wolkenhauer, (2014) proposed an iterative approach of data-driven modeling and model-driven experimentation, in which alternative hypotheses are postulated and refined until they are validated. This approach can help in identifying new regulatory interactions in the biological systems.

Systems biology has been widely applied in the identification of pathways based biomarkers, disease genes and drug targets (Chuang et al., 2010; Folger et al., 2011). It has also been used in biomedicine to better understand a particular disease and its molecular basis along with the mechanism of drug action (Zou et al., 2013). Three computational methods have been commonly used in drug discovery, such as, network analysis, constraint-based modeling and dynamical modeling (Folger et al., 2011; Zou et al., 2013). These computational methods and their applications provide a framework for addressing disease mechanism and approaching drug discovery that can be useful for clinical benefits, such as novel biomarkers and promising therapies (Zou et al., 2013).

A systems biology approach was used to identify pathways involved in liver toxicity that is induced by free fatty acids and TNF-α in human hepatoblastoma cells (HepG2/C3A) (Li and Chan, 2009). This was done by integrating multiple-level information, i.e., microarray gene

expression, metabolite profile, toxicity measurements and pathways information (Li and Chan, 2009).

1.5 Pathway analysis

High-throughput techniques of gene and protein profiling enable a comprehensive understanding of biological systems, and have revolutionized the biological research. Researchers in biology may share interest in the identification of a differentially expressed list of genes or proteins from analysis of high-throughput data in a disease context. Such a list is extremely useful to identify a pathway that may have an important role in a given phenomenon, or phenotypes, or disease. However, for many researchers, this list usually fails to give insight into the underlying mechanism of disease or phenotypes. Thus, the emergence of high-throughput profiling technologies introduced a new challenge in extracting meaningful knowledge from a long list of differentially expressed gene and proteins. An approach has been used to address this challenge, by which a long list of genes is divided into smaller sets to reduce the complexity. Using the knowledge of biological processes, groups of genes and proteins have already been defined, which represents genes related to pathways. This makes analysis easier because thousands of genes are reduced to a few hundred pathways. One use of this idea is to identify a list of pathways based on a longer list of differentially expressed genes, a process termed pathway analysis.

Pathway analysis can give us some basic insight into the underlying biology of differentially expressed genes and proteins, as it reduces complexity and increases explanatory power. A pathway over-representation analysis (ORA) statistically evaluates the fraction of genes in a particular pathway found among the set of genes showing changes in expression. There are several tool for pathways ORA that include differentially expressed genes and proteins (Huang et al., 2009; Kamburov et al., 2011a). Based on genes and metabolites a combined pathway analysis can be done (Kamburov et al., 2011b).

1.5.1 Pathway databases

Pathway databases store the knowledge on the molecular interactions and reaction networks. There are several pathways databases such as KEGG (Kanehisa et al., 2012), Reactome (Croft et al., 2011), and ConsensusPathDB (Kamburov et al., 2011a).

The KEGG database. KEGG is integrated with 15 data objects which are known as KEGG objects for computer representation of molecular systems (Kanehisa et al., 2012). KEGG objects are broadly categorized into systems information (PATHWAY, BRITE, MODULE, DISEASE, DRUG and ENVIRON), genomic information (ORTHOLOGY, GENOME and GENES) and chemical information (COMPOUND, GLYCAN, REACTION, RPAIR, RCLASS and ENZYME). KEGG includes a collection of tools for KEGG PATHWAY, BRITE and MODULE mapping, which enables integration and interpretation of large-scale

data sets. Different type of informations, such as disease genes and drug targets can be integrated as a part of molecular networks of KEGG.

The Reactome database. Reactome is an open online pathway and manually curated database of human pathways and processes (Croft et al., 2011). It includes compartment information of entities and transport reactions that transport the entities from one cellular compartment to another. Entities represent nucleic acids, small molecules, metabolites, proteins (with or without post-translational modifications) and macromolecular complexes and this generalization allows us to cover several biological processes such as signaling, metabolism, transcriptional regulation, and apoptosis which can be downloaded in a computationally navigable format. A visualization of full pathways is represented in Systems Biology Graphical Notation which is a standard graphical representation in systems biology, i.e., every molecule and reaction has a particular shape (Croft et al., 2011; Le Novère et al., 2009). Reactome offers the Pathway Analysis tool which can be used in ID mapping, pathway assignment, and over-representation analysis by user-supplied lists of genes, proteins and small molecules. Furthermore, it is a resource of human pathways that can be used for pathway modeling, systems biology, genome analysis and education.

The ConsensusPathDB database. ConsensusPathDB is a database that integrates different types of functional, molecular and reaction interaction networks from several databases such as KEGG, Reactome, etc. (Kamburov et al., 2011a). Currently the database comprises information of physical protein interactions, metabolic and signaling reactions and gene regulatory interactions for three different organisms: human, mouse and yeast. The database integrates currently (June 2014) 155,432 functional interactions and 2,205 biological pathways in human, 194,480 functional interactions and 734 biological pathways in yeast and 13648 functional interactions and 1,381 biological pathways in mouse, which originate from 18 databases on human and eight databases on yeast and mouse interactions each (Kamburov et al., 2011a). This database features an advanced visualization by which network nodes (physical entities/functional interactions) are movable and can be rearranged automatically using different layout methods; thus this has advantages for visualizing larger networks. Using the Web interface of ConsensusPathDB, pathways as well as network module analysis can be done by user defined gene lists.

Pathway analysis has become the first choice for obtaining meaningful insight into high-throughput molecular measurements. Nearly every recent bioinformatics study in a disease context looked for statistically significant pathways for the biological interpretation of a disease. In a recent study several pathway analysis tools of the last ten years have been compared and discussed along with their methodological challenges (Khatri et al., 2012). A first challenge is in the restriction of the development of next-generation pathway analysis methods due to a low resolution knowledge base, missing condition- and cell-specific information, and incomplete annotations. Secondly, they found limitations in the utility of existing methods due to the fact that these methods do not incorporate the dynamical nature

inherent to biological systems (Khatri et al., 2012). However, pathway analysis is still useful for identifying deregulated pathways in a disease context that can be further used for modeling purposes to better understand the dynamics of biological systems.

1.6 Computational modeling techniques of biological systems

Systems biology is a field of science that studies complex interactions of biological systems for example at the cellular level (Kitano, 2002a). This field is essential for research in several areas such as drug development, cellular interactions, and biotechnology. Computer models can be built for biological processes using computational tools and high-throughput experimental data. These models might cover signaling, gene regulatory or metabolic networks. To study cellular systems requires a whole-cell modeling system. However, it is difficult to develop a whole-cell computational model that integrates and simulates all the components of a living cell. The reduction modeling approach (that tends to focus on the application of interest) is popular due to the enormous size and complexity of computational cell models. Several modeling frameworks have been used in systems biology including constraint-based and dynamic modeling formalisms such as differential equations and Petri nets.

1.6.1 Constraint-based modeling

Constraint-based modeling of cellular metabolism was influenced by the work of Palsson and contributors (Varma and Palsson, 1994). Constraint-based methods, such as flux balance analysis, only require information about metabolic reaction stoichiometry, and metabolic requirements for growth. These methods overcome the limitations in the lack of experimental data for parameter estimation inherent to fully detailed dynamic models, but limit the dynamic study of biological systems.

1.6.1.1 Flux balance analysis

Flux balance analysis (FBA) is a constraint-based method for studying biochemical networks, genome-scale metabolic networks in particular (Orth et al., 2010; Schellenberger et al., 2010). These networks contain all known metabolic reactions of an organism. The first step of FBA is to form a stoichiometry matrix S of size $m \times n$, where m and n are number of metabolites and reactions of the network, respectively. The entries of each column are the stoichiometric coefficients of the metabolites participating in the reaction. The negative coefficient determines a metabolite as a substrate while the positive coefficient determines a metabolite as a product (Figure 1.5a & b). The steady-state flux through all of the reactions in a network is represented as vector $\boldsymbol{v}$ of length n.

In FBA, the representation of constraints can be in two ways: 1) equations that balance

reaction inputs and outputs and 2) inequalities that impose bounds on the system. The system of mass balance equations at steady state ($d\boldsymbol{x}/dt=0$) is given in Figure 1.5c: $\boldsymbol{Sv}=0$, where $\boldsymbol{S}$ is the stoichiometry matrix, $\boldsymbol{v}$ is the vector of reaction fluxes, and $\boldsymbol{x}$ is a vector of metabolic concentrations. Usually, genome-scale metabolic networks have more reactions than compounds ($n>m$) resulting in an equation system with more unknown variables than equations. Thus there exists no unique solution to this system of equations. Upper and lower bounds can also be set for every reaction (Figure 1.5d). Upper and lower bounds define the maximum and minimum allowable fluxes of the reactions.

FBA is used to maximize or minimize an objective function $Z=\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}$, which can be any linear combination of fluxes, where $\boldsymbol{c}$ is the weight vector, indicating how much each reaction contributes to the objective function. For example, the production of b2 in Fig. 1.5 is maximized by using vector $\boldsymbol{c}$ of zeros with one at the position of reaction or production of b2 (Figure 1.5e). Thus, FBA uses linear programming to solve the equation $\boldsymbol{Sv}=0$ given a set of upper and lower bounds on $\boldsymbol{v}$ and an objective function Z. The output of FBA is a particular flux distribution $\boldsymbol{v}$ which maximizes or minimizes the objective function but need not to be unique.

1.6.1.2 Integration of gene expression data with flux balance analysis

The FBA framework has been applied to many genome-scale models with great success (Duarte et al., 2007; Sigurdsson et al., 2010) along with the systematic prediction of genetic knockout phenotypes (Deutscher et al., 2006). However, methods are still missing that can be performed on genome-scale models as well as integrate data from all cellular levels (e.g. genomic, transcriptomic, proteomic, metabolomic), and that can accurately predict metabolic phenomena under different environmental conditions.

The information flows from DNA to mRNA and to enzymes that catalyze biological reactions and regulate various cellular functions. Hence, one might expect the expression levels of mRNA to correlate directly with the amount of enzymes and thus with the flux through associated reactions. A method called Gene Inactivity Moderated by Metabolism and Expression (GIMME) proposed to produce the context-specific reconstruction using the quantitative gene expression data, the genome-scale reconstruction, and one or more required metabolic functionalities (Becker and Palsson, 2008). To get a context-specific reconstruction, first, a reduced model is obtained by removing reactions from genome-scale reconstruction with gene expression data below the pre-defined threshold. Next, reactions are reinserted in a reduced model in order to ensure that the FBA model achieves its required metabolic functionalities.

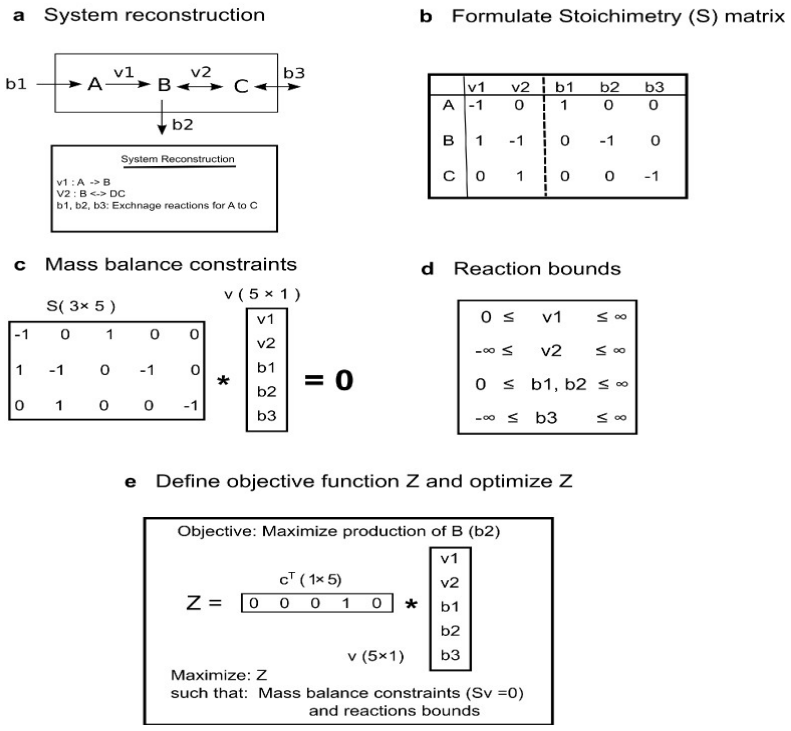

Figure 1.5 Flux balance analysis formulation. a) A metabolic network reconstruction is built that consists of a list of stoichiometrically balanced biochemical reactions. **b)** This reconstructed metabolic model is formulated in a mathematical model by forming a stoichiometry matrix ***S***, in which each row represents a metabolite and each column a reaction, respectively. **c)** At steady state, mass balance can be represented as $\boldsymbol{Sv}=0$, where ***S*** is stoichiometry matrix and ***v*** is flux vector of reactions of the network. **d)** Additional flux constraints can be imposed by lower and upper bounds. **e)** The objective function is defined as $Z=\boldsymbol{c}^{T}\boldsymbol{v}$, where ***c*** is vector of weights indicating the contribution of reactions to the objective function.

Another method is E-Flux (as a combination of flux and expression) that extends the technique of FBA by using maximum flux constraints as a function of measured gene expression (section 2.3; Colijn et al., 2009). The method is similar to setting the width of "pipes" around particular reactions that represent the function of an expression state. Low expression of genes represent thin pipes around reactions that can limit the maximum flux through the reactions. Thus, normalized gene expression data is used to constrain the reaction flux. This method can be used to obtain a metabolic state under a particular condition using condition dependent gene expression data.

A further method has been proposed that combines gene-expression data with FBA (GX-

FBA) (section 2.4; Navid and Almaas, 2012). The method allows modeling of genome-scale metabolic responses to a broad range of environmental perturbations. The ratio (normal vs perturbed) of gene expression is used to formulate the objective function for understanding perturbation effects on metabolic states. Although the method does not account for capacity limitations in various enzymes, it is able to impose pseudo-metabolic regulation due to mass balance and network connectivity.

1.6.2 ODE-based modeling

In a cell, abundances of cell molecules such as mRNAs, proteins, or metabolites can range from a few molecules to thousands of molecules or even more, and they vary in time and in response to external or internal stimuli. The changes of cell molecules can be captured by a cellular interaction network that needs to be developed into a dynamic network model. Dynamic network models have as input the interaction network in which nodes represent cell molecules, the set of equations indicating how the state of each node changes in response to changes of the state of its regulators, and the initial state of each node in the system (Albert, 2007). For example, suppose x is the concentration of a substrate S (e.g., mRNA, protein and metabolite); $\frac{dx}{dt}=kx$, where x is the state variable and k is a parameter. It describe a reaction rate, i.e., how the concentration of S varies with time.

A dynamic model is needed to study the dynamics of biological systems. Differential equations describe the rate of change of continuous variables. These equations are typically used for modeling dynamical systems in systems biology as well as in other areas. Systems of non-linear ordinary differential equations (ODEs) have been used to obtain the amount of species, such as proteins and metabolites, in the modeled system as a function of time. They have been applied to pathways related to metabolism (Yang et al., 2007), signal processes (Tyson et al., 2003) and transcriptional regulation (Elowitz and Leibler, 2000).

A detailed kinetic model can be used to perform time-course simulations, to predict the response to different inputs, and to design system controllers. However, building kinetic models requires reactions rate laws and its kinetic parameters. Kinetic parameters are limited because they either have to be measured experimentally or estimated based on experimental data. The lack of kinetic data has limited the scale of the modeled networks.

1.6.3 Petri net-based modeling

Petri net is a graphical and mathematical modeling language developed in the early 1960s by Carl Adam Petri. Petri nets consist of two types of nodes (places and transitions) and arcs. Arcs join a place to a transition or a transition to a place. In the Petri net modeling of biochemical reactions, places represent e.g. reactants, products and enzymes whereas transitions represent biochemical reactions.

A static Petri net N can be written as $N = (P, T, F, W, M_0)$ where,

- P is the set of places,
- T is the set of transitions
- $F \subseteq (P \times T) \cup (T \times P)$, where (p, t) or (t, p) are arcs and p is an element of set P and t is an element of set T.
- A mapping $W : F \rightarrow \mathbb{R}$, is the weight function W(p, t) assigning each arc (p, t) a real number.
- M_0 is the initial marking, which contains initial tokens of places of P.

In Petri net-based modeling, biochemical systems evolve in a sequence of steps. A step converts a given state into a new state upon the occurrence of reactions. A corresponding Petri net models a state as a marking of places, and steps as an occurrences of transitions. The sequences of steps describe the dynamic behavior of nets.

The Petri net-based modeling has subsequently been adapted and extended in many fields such as systems biology. A qualitative analysis of a metabolic pathway is performed using the Petri net-based modeling. Discrete and continuous events can be modeled with Petri nets. Many studies have modeled discrete and continuous events in pathways using the concept of Petri nets. Petri net-based methods have already been used for the modeling of different biological systems (Doi et al., 2011; Voss et al., 2011). An extension of Petri nets, hybrid functional petri net, was for example used to model apoptosis induced by Fas ligand, and *Drosophila* circadian oscillator (Matsuno et al., 2003). A non-parametric Petri net-based strategy is proposed for characterizing the dynamics of signaling pathway (Ruths et al., 2008). A Petri net framework is introduced to investigate biomolecular networks (Rohr et al., 2010).

1.7 Modules identification in biological networks

Biological networks can be seen as organized in a modular, hierarchical manner. There is growing need to identify modules for a better understanding of the modularity and organization principle of a large-scale biochemical network that continues to grow in size (Kessler et al., 2013). Such modules represent some biological processes, which conceptually simplify the function of complex biological networks. This functional module analysis of metabolism can be performed using pathway analysis, such as elementary flux mode identification. However use of this method is often hampered by combinatorial explosion due to the complexity of metabolic networks (Schuster et al., 1999).

There is usually intuitive reasoning behind defining the network modules and there are several ways to identify modules. For example, a decomposition method uses a distance definition derived from the path length between two reactions (Ma et al., 2004). Furthermore, a network-clustering method has been developed to identify modules in protein-interaction network of the yeast cell (Rives and Galitski, 2003). In addition, Müller

and Bockmayr, (2013) presented a direct method for computing flux modules of the thermodynamically constrained optimal flux space of a metabolic network. Moreover, modules in biochemical reaction networks can be defined by correlated reaction sets (Co-Sets) (Papin et al., 2004). The group of reactions of Co-Sets always appear together in functional states of that network. Thus, they represent a functional module of the reaction network. In this study, modules are identified using Co-Sets formalism and elementary modes. Hence, flux coupling analysis, a method for obtaining Co-Sets along with elementary flux mode are described in given the following.

1.7.1 Flux coupling analysis

Flux coupling analysis (FCA) is a method to identify dependencies between reaction fluxes. Blocked reaction can not carry any flux while unblocked reactions can carry flux under the steady state condition. To find coupled reaction sets and blocked reactions in genome-scale metabolic networks, the flux coupling finder procedure was introduced (Burgard et al., 2004). The set of blocked reactions is identified by maximizing flux through the reactions of a given network under the steady state condition. If the maximum value of flux through a reaction is zero, then this reaction is said to be unusable or blocked because it cannot carry any flux. Let v1 and v2 be fluxes of two unblocked reactions. The coupling relationships are defined as following:

1. Directional coupling, if a non-zero flux for v1 implies a non-zero flux for v2 but not necessarily the reverse.

2. Partial coupling, if a non-zero flux for v1 implies a non-zero, though variable, flux for v2 and vice versa.

3. Full coupling, if a non-zero flux for v1 implies not only a non-zero but also a fixed flux for v2 and vice versa

Reaction pairs not characterized by one of these categories are classified as uncoupled.

1.7.1.1 Fully coupled reactions as a module

In recent years, genome-scale metabolic networks of various organisms have been reconstructed (Schellenberger et al., 2010) and kinetic modeling of these models is either impractical or infeasible. Thus, the modularity of a biochemical network gives a way to build useful and simplified models of complex biological systems. Functional modules of biological networks can be a mechanism by which a cell coordinates the functional states (phenotype) of metabolic interactions (Resendis-Antonio et al., 2012). The cellular phenotype expressed under certain conditions can be represented by the metabolic fluxes (Nielsen, 2003). The correlation among reaction fluxes across network states can be a measure of functional dependence among enzyme-coding genes (Papin et al., 2004). Recently, the functional associations between metabolic genes (also referred to as “correlated

reaction sets" or "flux coupling") of genome-scale metabolic networks can be determined using the developed computational framework (Burgard et al., 2004; David et al., 2011; Larhlimi et al., 2012). There is also another method implemented in COBRA for finding coupled reaction sets using an artificial centering hit-and-run (ACHR) sampling strategy (Becker et al., 2007). A combined analysis of gene expressions and flux coupling suggests that genes with correlated reactions often show signs of co-regulation (Price et al., 2004). Hence, fully coupled reactions might be taken as modules to understand the co-regulation of gene expressions in two different conditions.

1.7.2 Elementary modes

A pathway, e.g. glycolysis as a part of a metabolic network, can be defined as a set of consecutive reactions that are linked by common metabolites. Different levels of large-scale datasets, such as genomics, transcriptomics, proteomics, and metabolomics have been mapped to pathway maps, e.g., KEGG (Kanehisa et al., 2012), for systematic biological interpretation. Identification of alternative paths in complex networks has been elaborated by introducing the concept of elementary flux mode (Schuster and Hilgetag, 1994). To get elementary flux modes (EFMs), the stoichiometry matrix of a metabolic network can be investigated to identify direct routes that start and end at external metabolites.
A flux mode $\boldsymbol{M}$ is a set of flux vectors that represents direct routes through a metabolic network. Mathematically, it is defined as the set

$$M=\{\boldsymbol{v}\ \epsilon\ R^r / v=\lambda\,\boldsymbol{v}^*, \lambda>0\}, \tag{1.1}$$

where $\boldsymbol{v}^*$ is an r-dimensional vector fulfilling two conditions: i) steady state $\boldsymbol{Sv}=0$, where $\boldsymbol{S}$ is stoichiometry matrix and ii) sign restriction, i.e., the flux direction in $\boldsymbol{v}^*$ fulfill the prescribed irreversibility reactions (Klipp et al., 2009).

A flux mode is called an EFM if it uses a minimum set of reactions and cannot be further decomposed. In other words, an EFM is a minimal set of reactions that can operate at steady state with all irreversible reactions used in the appropriate direction. There are software tools for the computation of EFMs, e.g., METATOOL (Pfeiffer et al., 1999), COPASI (Hoops et al., 2006), FluxAnalyzer (Klamt et al., 2003) and PySCeS (Olivier et al., 2005). An example of elementary flux modes of a metabolic system is illustrated in Figure 1.6.

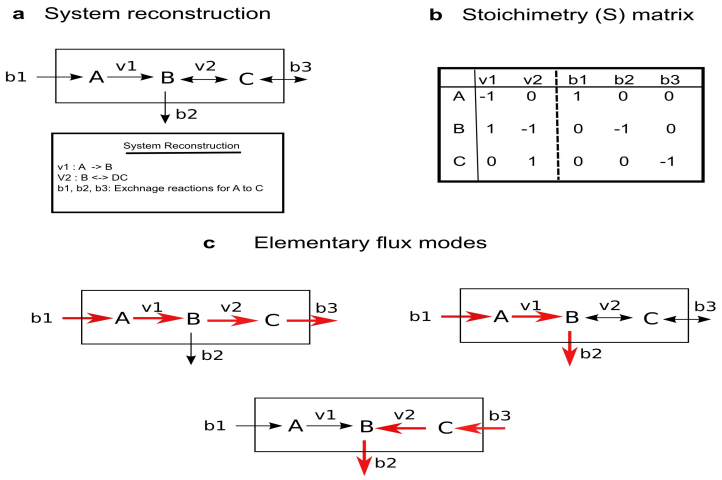

Figure 1.6 An example of elementary flux modes. a) Metabolic system **b)** stoichiometry matrix **c)** all possible three elementary flux modes represented by red arrows.

1.8 Objectives

The main objective of my thesis is to comprehensively describe and model the interactions of susceptibility genes, proteins, and metabolites in metabolic liver diseases, such as steatosis and steatohepatitis. There can be a great variability in human liver diseases such as steatohepatitis due to a broad genetic spectrum of individuals and the environmental risk. To study the development of steatohepatitis, different mouse models were used. Three mouse strains A/J, C57Bl6 and PWD were fed a 5-diethoxycarbonyl-1,4-dihydrocollidine (DDC) supplemented diet, which mimicked different degrees (A/J, high; C57Bl6 low; PWD, unspecific) of the steatohepatitis-like phenotype. Gene expression, proteomics and metabolic profiles were available for mouse liver samples and these data were used in this study to identify and model disease-related molecular processes.

Given this data I want to 1) understand the different degree of steatohepatitis using metabolic fluxes, 2) identify statistically significant pathways related to liver disease, 3) identify interactions or modules of susceptibility genes, proteins, and metabolites, 4) find novel potential metabolic drug targets for liver disease, and 5) build a model of deregulated pathways in liver disease as well as to use this model for predicting qualitative and quantitative changes in metabolite concentration based on changes in expression data.

Understanding the different degrees of steatohepatitis using metabolic fluxes. This goal can be achieved by integration of gene expression data to the mouse genome-scale metabolic

model (Sigurdsson et al., 2010). Expression data of mice under normal as well as treated states can be studied (Colijn et al., 2009; Navid and Almaas, 2012). One can identify consistent results and ask questions such as, are fluxes regulated differently under normal and treated conditions? Furthermore, the relative fluxes (normal *vs* treated) can be visualized by the map of BiGG, and therefore, one can observe whether the three mice exhibit different degree of flux regulation that may be linked to their degree of steatohepatitis like phenotypes (Schellenberger et al., 2010).

Identifying statistical significant pathways related to liver disease. Pathways analysis allows to gain primary insights into the underlying biology of a disease from large sets of, e.g., expression data in a disease context. Differentially expressed genes and differentially regulated metabolites can be identified between treated *vs* control mice (Robinson et al., 2010). The list of genes and metabolites can be used to identify deregulated pathways using pathway analysis tools such as ConsensusPathDB (Kamburov et al., 2011a). Based on the results of pathway analysis of gene expression and metabolic profiles, statistical significant pathways can be identified that can be used as a model system to study their dynamics.

Identifying interactions or modules of susceptibility genes, proteins, and metabolites. Cellular functions might be assumed to be composed in a modular way and each module comprises many species of interacting molecules which are participating in a common biological function (Hartwell et al., 1999). Using a mouse genome-scale metabolic network, functional modules can be identified based on correlation among reaction fluxes (Papin et al., 2004). Correlation among reaction fluxes can be obtained using F2C2 and COBRA (Becker et al., 2007; Larhlimi et al., 2012). Expression data can be used to identified flux modules that may demonstrate different degree of regulation of genes, proteins and metabolites that may be linked to mice phenotypes. Elementary flux mode analysis is a tool for pathway or module identification, but obtaining a complete set of elementary flux modes is computationally intensive for large-scale networks and especially genome-scale networks. However, there is an algorithm to obtain k-shortest elementary flux modes that can be used for large-scale networks (de Figueiredo et al., 2009). Differentially regulated metabolites (control vs treated) can be used further as an intermediate node of the shortest elementary flux modes of the metabolic model. Such elementary flux modes that pass through differentially regulated metabolites can be further investigated whether genes of the modules are also differentially expressed. Thus, using gene expression and metabolic profiles, one can identify deregulated pathways in a disease context, which can be used for building and validating small-scale dynamic model helping to understand the biology of the disease progress.

Finding novel metabolic drug target candidates for liver disease. There are very few therapeutic drugs for hepatic steatosis and steatohepatitis, which are related to weight loss,

lowering the blood sugar and lowering the lipid count (Anderson and Borlak, 2008). The decreasing number of newly released drugs of hepatic steatosis, steatohepatitis and liver cancer urges the need to find novel metabolic drug targets for liver disease. An objective function that mimics liver disease progression can be formulated. This function can be used for a biomass reaction in metabolic models for finding potential metabolic drug targets.

Building a model of deregulated pathways in liver disease to predict qualitative and quantitative changes in metabolite concentrations based on changes in expression data. As I mentioned, pathways and module analysis can be used to identify deregulated pathways using gene expression data in a disease context. Furthermore, these pathways or modules can be used to construct kinetic models. The availability of kinetic parameters is limited because they are normally measured in specific conditions. However, there are some proposed methods, such as Monte Carlo ODE-based and Petri net-based approaches that are limited to a qualitative or semi-quantitative level (Ruths et al., 2008; Wierling et al., 2012). Here, I want to test the prediction of changes in metabolic concentration based on changes of gene expression using both aforementioned methods. Then results can be compared to the experimental data. These methods can only explain changes of metabolic concentrations at a qualitative level. However, one can also build a detailed kinetic model by estimating kinetic parameters and including regulatory links. Hence, I envisage to build a model that can be used in various ways such as understanding the effect on the metabolic concentrations after changing the enzyme concentrations and single- and multi-drug targets identification.

2 Materials and Methods

In this section I describe published models and methods which are used in analysis of this work, for example, a genome-scale mouse metabolic model (section 2.1) and the algorithm of shortest elementary flux modes (section 2.2). Methods which incorporates gene expression data with FBA are illustrated (section 2.3 and 2.4). I introduce the Monte Carlo-based simulation strategy (section 2.5), Petri net simulator (section 2.8), kinetic parameter optimization (section 2.9), and the formulation of an objective or biomass function which mimics the readout of steatosis, steatohepatitis and HCC (section 2.6). Experimental data of mouse liver samples are described in section 2.7.

2.1 Genome-scale mouse model

A genome-scale mouse metabolic model that comprises 3727 reactions, 2775 metabolites, and 1415 genes is used in several studies, such as elementary mode, flux balance analysis, flux coupling analysis and drug targets analysis (Sigurdsson et al., 2010). The model is reconstructed based on Recon1 (Duarte et al., 2007) and denoted in the following as GMM. The GMM has a pre-defined biomass reaction that mimics cellular growth, has eight cellular compartments, and has been tested in 260 validation test cases (Sigurdsson et al., 2010). Deletion of essential genes of the GMM model resulted in zero biomass production. These essential genes are identified *in silico* and confirmed with high accuracy *in vivo* (Sigurdsson et al., 2010).

2.2 K-shortest elementary flux mode

Elementary flux mode (EFM) analysis is a useful tool, e.g., to elucidate novel metabolic pathways in addition to the standard biochemistry text book knowledge. For example, a new metabolic pathway has been found that catalyzes glucose oxidation via the phosphoenolpyruvate-glyoxylate cycle (Fischer and Sauer, 2003). However, EFM analysis has a severe drawback: the number of EFMs grows exponentially with network size and this leads to a large number of possible EFMs even for medium-sized models (Klamt and Stelling, 2002). For instance, a metabolic network of central metabolism, 112 reactions and 89 metabolites, has 2,450,787 EFMs (Gagneur and Klamt, 2004). To generate an entire set of EFMs of genome-scale network is computationally prohibitive. To overcome such intensive computation, some methods have been developed, e.g., to generate random samples of EFMs without computing the whole set (Machado et al., 2012) and enumerating the K-shortest EFMs (de Figueiredo et al., 2009). Technically, the K-shortest EFM method produces EFMs in an increasing order of reactions by solving a sequence of discrete optimization problems. Thus, we have: 1-shortest EFM, the EFM containing the minimum number of reactions; 2-shortest EFM for the EFM containing the second minimum number of reactions, and so on. It may be possible for multiple EFM to contain the same minimum number of reactions. If

this happens then multiple EFM are associated with different K values.

The detection of K-shortest EFMs can be of interest for several biological applications. For example, the highest increase in pathways flux is achieved if the identified enzymes (Kacser and Acerenza, 1993) or a considerable number of enzymes (Fell and Thomas, 1995; Niederberger et al., 1992) are overexpressed, and shorter pathways can carry higher fluxes (Meléndez-Hevia et al., 1994; Pfeiffer and Bonhoeffer, 2004). Given below, some constraints are introduced for detecting K-shortest EFMs (de Figueiredo et al., 2009).

$$t_r \leq M\, z_r \quad r = 1, .., R \tag{2.1}$$

$$z_r \leq t_r \quad r = 1, .., R \tag{2.2}$$

$$z_\alpha + z_\beta \leq 1 \;\forall\, \alpha, \beta \in \mathrm{B} \tag{2.3}$$

$$\sum_{r=1}^{R} S_{cr} t_r = 0 \;\; \forall\, c \in I \tag{2.4}$$

$$\sum_{r=1}^{R} z_r \geq 1 \tag{2.5}$$

$$minimize \sum_{r=1}^{R} z_r \tag{2.6}$$

$$\sum_{r=1}^{R} Z_r^k z_r \leq \left(\sum_{r=1}^{R} Z_r^k\right) - 1,\; k = 1, .., K-1 \tag{2.7}$$

Let a metabolic network comprise R reactions and C compounds, and s_{cr} be the stoichiometric coefficient associated with compound c (c=1-----, C) in a reaction r (r=1----,R). z_r is a binary integer variable, assigned 1 or 0 depending on whether reaction r is active in an EFM or not. In addition, t_r is a non-negative integer flux associated with each reaction. Equation (2.1) ensures that no flux pass through the reaction r if $z_r = 0$. Equation

(2.2) indicates that t_r is non-zero if z_r =1. M is a large constant integer (M>=1). In the model, reverse reactions are decomposed into two irreversible reactions. A set B={(α, β)} | reaction α and reaction β are the forward and backward reaction of a reverse reaction} is defined. Equation (2.3) ensures that backward and forward reaction do no appear at an EFM. Equation (2.4) is defined for steady state condition where I is the set of internal compounds. In order to avoid the trivial solution ($z_r = t_r$ =0, r=1-----,R) a constraint is defined in equation (2.5). The shortest EM is obtained by using the mathematical optimization of the objective function (2.6) in subject to equations (2.1)-(2.6). A K-shortest EFM can be obtained by adding further a constraint to eliminate (K-1)-shortest EFMs from the set of solutions. Constraint (2.7) is used to get the K-th shortest EFM, where the constraint of the first (K-1) shortest EFMs are accumulated that makes sure that the K-th shortest EFMs are different from the (K-1) shortest EFMs.

2.3 E-Flux method

E-Flux is a method that integrates quantitative gene expression to genome-scale models (Colijn et al., 2009). Conceptually, the E-flux method can be understood as setting the width of "pipes" around the reaction and the quantity of widths is evaluated by a function of gene expression. Figure 2.1 describes a simple metabolic model that comprises 4 metabolites, 4 internal reactions catalyzed by enzymes corresponding to 4 genes, an uptake reaction, and a biomass reaction. Two different sets of gene expression data are shown in the two panels. In the top panel, G_1 is a poorly expressed gene that is interpreted as a thin pipe around reaction R_1, and limits the maximum flux through reaction R_1. In contrast, in the bottom panel, G_1 is a highly expressed gene corresponding to a more possible flux (a wider pipe). After integrating the expression to the model one could expect more flux through reactions R_1 and R_2 in the bottom panel than in the top panel.

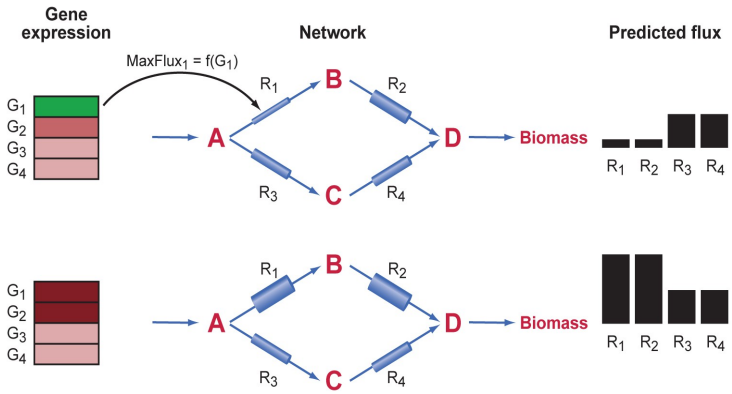

Figure 2.1. Illustration of E-Flux (source, Colijn et al., 2009). In E-flux, gene expression is used to set the maximum flux constraints on individual reactions. The quantity of gene expression can be interpreted as pipes of different widths around each reaction. Here, a simple model is shown that comprises of 4 metabolites (A-D), 4 internal reactions, an uptake reaction for A, and a reaction converting D into biomass. Gene expression data for 4 genes whose enzymes catalyze the 4 internal reactions (green-lower expression, red-higher expression) is shown on the left. G1 is a poorly expressed gene in the top panel that can be conceptualized as a thin pipe around reaction R_1 as shown. G1 and G2 are highly expressed genes that correspond to a wider pipe for these reactions as shown in upper panel. The method would predict more flux through R_1 and R_2 in the bottom panel relative to the top panel and R_3 and R_4. This is shown by the bars on the right.

Mathematically, E-Flux modifies FBA as follows. FBA involves solving the following optimization problem:

$$\max_{v} \quad c^T \cdot v \qquad (2.8)$$

$$where \quad \begin{cases} S \cdot v = 0 \\ a_j \le v_j \le b_j \end{cases}$$

where v represents a flux vector for a particular flux configuration, S is the stoichiometric matrix, c is a vector of coefficients that defines a linear objective function $c^T v$, and a_j and b_j are the minimal and maximal fluxes through the reaction j. In the E-flux method, the maximum flux b_j of j^{th} reaction is determined according to the expression level of associated genes of reaction j. a_j is determined according to negative value of b_j (i.e., $a_j = -b_j$) if the reaction is reversible, otherwise $a_j = 0$.

2.4 Integration of gene expression data into flux balance analysis

A constraint-based approach, GX-FBA, has been proposed for integrating gene expression data with FBA (Navid and Almaas, 2012). Perturbation experiments can be studied using

GX-FBA. For example, the change in metabolic states has been studied using expression ratio (perturbed *vs.* control). I introduced a "mGX-FBA" method which is slightly modified from GX-FBA. Steps of mGX-FBA are:

1) Assume that the flux distribution of the genome-scale model represents a reference state (control experiment). Thus, maximize flux through the biomass reaction of a metabolic model and obtain flux distribution (v^{cont}) which represents the control experiment.

2) Perform flux variability analysis (Müller and Bockmayr, 2013; Mahadevan and Schilling, 2003) to calculate the lower and upper fluxes of each model reaction i (v_i^{min} and v_i^{max} respectively) based on environmental limitations and network connectivity. Then, calculate the mean possible flux of each reaction *i* by averaging v_i^{min} and v_i^{max} .

3) Identify the set of reactions **S** for which expression data can be integrated. Note that reversible reactions are excluded from **S**. Calculate the ratio (perturbed *vs.* control) of gene expression. For those reactions for which catalyzing enzymes are associated with more than one gene, the maximal up- or down-regulation value is used unless the gene expression ratio is inconsistent (mixture of up- and down-regulation). In the inconsistent case, the reaction is excluded from **S**.

4) For each reaction *i* in **S**, calculate $b_i = r_i^{RNA} * v_i^{cont}$ if the gene expression is upregulated, and $a_i = r_i^{RNA} * v_i^{cont}$ if it is downregulated, where r_i^{RNA} is the gene expression ratio, and use b_i as an upper bound and v_i^{min} as a lower bound or a_i as a lower bound and v_i^{max} as an upper bound.

5) Construct an objective function as defined by the GX-FBA method: $\sum_{i \in T} log2 \frac{(r_i^{RNA}) * v_i}{\bar{v}_i}$

where $\bar{v}_i$ is the mean flux of reaction *i* in set **S**. Objective function aims to maximize the correlation between differential changes in gene-expression and reaction fluxes.

2.4.1 Comparison of fluxes between control *vs* perturbed

The relative deviation between the fluxes of a reaction in two conditions is calculated as:

$$d_r = \frac{x - y}{|x| + |y|} \tag{2.9}$$

where x and y are the flux of a reaction r in condition 1 and condition 2, respectively.

2.5. Perturbation based on a Monte Carlo simulation strategy

A Monte Carlo method is proposed in which kinetic parameters were sampled from a given distribution (Wierling et al., 2012). Models were simulated into its steady state. Monte Carlo simulations can be performed as given in the following:

Suppose a metabolic model **M** (X_1, X_2, …, X_N; R_1, R_2, …, R_M; E_1, E_2 …, E_M) where,

X_1, X_2, …, X_N are metabolites, R_1, R_2, …, R_M are reactions, and E_1, E_2 …, E_M are enzymes.

1. The model **M** is initialized with control and perturbed/treated data from the experiments, such as enzyme and metabolite concentrations, and simulated into its steady state.
2. **M** was simulated with a given number of sampled kinetic parameters sets eventually resulting in **S** steady state concentration vectors.
3. $(X_{11}, X_{21}, \ldots, X_{N1})^C$, $(X_{12}, X_{22}, \ldots, X_{N2})^C$, …, $(X_{1S}, X_{2S}, \ldots, X_{NS})^C$ and $(X_{11}, X_{21}, \ldots, X_{N1})^T$, $(X_{12}, X_{22}, \ldots, X_{N2})^T$, …, $(X_{1S}, X_{2S}, \ldots, X_{NS})^T$ are steady state concentration vectors of metabolites for the control and perturbed/treated states, respectively.
4. Geometric means over all steady state concentration vectors were calculated for each metabolite.

2.6 Biomass reaction for liver disease

Three biomass reactions are formulated in order to mimic phenotypes of steatosis, steatohepatitis and hepatocellular carcinoma (HCC).

2.6.1 Lipid droplet compositions

Lipid droplets (LDs) can formed through the accumulation of lipids. LDs comprise phospholipid, sterol, diacylglycerol, triacylglycerol and sterol ester (Guo et al., 2009). The compositions of LDs include cholesteryl esters (34%), triacylglycerol (44%), cholesterol and diacylglycerol (1.6%), and phospholipid (1.7%) (Bartz et al., 2007). The GMM metabolic model contains 4 phospholipids: phosphatidylserine, phosphatidic acid, phosphatidylethanolamine and phosphatidylinositol. I hypothesized that LDs are formed by triacylglycerol, cholesteryl esters, cholesterol and diacylglycerol, and aforementioned 4 phospholipids. The percentage of composition of LDs are scaled on 100%, and new compositions are triacylglycerol (54.2%), cholesteryl esters (41.8%), cholesterol (1%), diacylglycerol (1%) and each 4 phospholipids (0.5%).

To obtain a medium growth of the LDs formation the coefficients of the components of LDs were estimated using aforementioned percentage of the LDs compositions. Substrates are cholesterol/chsterol(c), diacylglycerol/dag_hs(c), phosphatidic acid/pa_hs(c), phosphatidylinositol/pail_hs(c), phosphatidylethanolamine/pe_hs(c), phosphatidylserine/ps_hs(c), triacylglycerol/tag_hs(c) and cholesteryl esters /xolest_hs(c)

and their coefficients are 0.001, 0.001, 0.0005, 0.0005, 0.0005, 0.0005, 0.0418, and 0.00542, respectively. Range of the coefficients is estimated based on the study of Sigurdsson et al., (2010). The flux through the lipid droplet formation reaction is interpreted as the steatosis phenotype.

2.6.2 Biomass reaction of steatohepatitis

A review was done to identify key metabolites that may participate in characterizing the steatohepatitis phenotypes. A “two hit” model was proposed for the progression of steatohepatitis whereby the first “hit”, steatosis, sensitizes the liver and the second “hit” that leads to oxidative stress (OS) and cytokine-induced liver injury (Day and James, 1998). OS is caused by reactive oxygen species, such as, superoxide anion, hydrogen peroxide, that damage to hepatic membranes, proteins and DNA and play an important role in the development of NASH (Koek et al., 2011). The generation of OS mediates palmitate-induced lipotoxicity and cellular stress (Srivastava and Chan, 2007) and the metabolite ceramide, which belongs to the class of sphingolipids, plays an important role in lipoapoptosis (Unger and Orci, 2002). The metabolites alanine and glutamate can be used as biomarkers for apoptosis (Halama et al., 2013). Elevated triacylglycerol levels and ceramides are found in lipidomics analysis of obesity related hepatic steatosis (Yetukuri et al., 2007). A study proposed that a deficiency of nutrients, such as glucose can activate the pro-apoptotic proteins Bax and Bak and cause cell death (Mason and Rathmell, 2011). Based on aforementioned study, I formulated an objective function to mimic the steatohepatitis. To formulated an objective function I used a combination of phenotypes that includes lipid droplets (40%), oxidative stress (20%) which may be caused by hydrogen peroxide and superoxide, and apoptosis (40%) which may be caused by alanine, glutamate, palamite and ceramide. Substrates and their stoichiometry coefficients of the biomass reaction for steatohepatitis are given in Table 2.1.

Table 2.1. Biomass reaction for steatohepatitis. Substrates and their coefficients are given in the table. 'c' represents cytosol and short form of metabolites of GMM model are given in the table.

Metabolites	Stoichiometric coefficients
ala-L(c)	1.00E-002
chsterol(c)	4.00E-004
crm_hs(c)	1.00E-002
dag_hs(c)	4.00E-004
glu-L(c)	1.00E-002
hdca(c)	1.00E-002
pa_hs(c)	2.00E-004
pail_hs(c)	2.00E-004
pe_hs(c)	2.00E-004
ps_hs(c)	2.00E-004
tag_hs(c)	1.67E-002
xolest_hs(c)	2.17E-002
h2o2(m)	1.00E-002
o2s(m)	1.00E-002

For cancer an increase in the overall biomass flux was assumed that reflects the increase in proliferation. Hence, a biomass reaction was introduced to measure the proliferation in cancer (Folger et al., 2011). I formulated a biomass reaction to mimic the readout of HCC. Substrates of the biomass reaction were used by a proposed study of Folger et al. and their coefficients were used form the study of Sigurdsson et al., (2010) (Table 2.2).

Table 2.2. Biomass reaction for proliferation. Substrates and their coefficients are given in the table. Coefficients are used from the biomass reaction of the GMM model. 'c' represents cytosol and short form of metabolites of GMM model are given in the table.

Metabolites	Stoichiometric coefficient
alaL(c)	6.00E-001
amp(c)	3.30E-002
argL(c)	3.77E-001
asnL(c)	2.88E-001
aspL(c)	3.59E-001
chsterol(c)	1.80E-002
clpn_hs(c)	3.00E-003
cmp(c)	5.51E-002
cysL(c)	1.45E-001
damp(c)	1.48E-002
dcmp(c)	9.90E-003
dgmp(c)	9.90E-003
dtmp(c)	1.48E-002
glnL(c)	3.22E-001
gluL(c)	3.86E-001
gly(c)	5.38E-001
gmp(c)	6.24E-002
hisL(c)	1.43E-001
ileL(c)	3.24E-001
leuL(c)	5.64E-001
lysL(c)	5.70E-001
metL(c)	1.38E-001
pail_hs(c)	1.00E-002
pchol_hs(c)	6.90E-002
pe_hs(c)	2.60E-002
pglyc_hs(c)	1.00E-003
pheL(c)	2.19E-001
proL(c)	3.13E-001
ps_hs(c)	3.00E-003
serL(c)	4.30E-001
sphmyln_hs(c)	8.00E-003
thrL(c)	3.86E-001
trpL(c)	4.40E-002
tyrL(c)	1.82E-001
ump(c)	3.30E-002
valL(c)	4.16E-001

2.7 Mouse experimental data

Data from mouse experiments were used to understand NAFLD (e.g., steatosis and steatohepatitis; unpublished data). Three different mouse strains A/J, C57Bl6 and PWD were used to study steatohepatitis. A histological analysis of the liver samples of three mice was conducted after 8 weeks of feeding with a DDC supplement diet. Mice A/J, C57Bl6 and PWD show high, low and no steatohepatitis-like phenotype, respectively, but PWD mice show high fatty liver (steatosis) than A/J and C57Bl6 mice. RNA-Seq experiments were performed to obtain expression data. mRNA expression was generated using liver samples of three biological replicates of each mouse strain in control and DDC-treated conditions. Mass spectrometry techniques were performed to obtain proteomics and metabolomics data. Proteomics and metabolomics data are measured for selected proteins and metabolites.

Differentially expressed genes between DDC-treated and control mice were obtained using the edgeR package (Robinson et al., 2010). Three biological replicates of each mouse strain A/J, C57Bl6, and PWD for control and DDC-treated states, respectively, were used.

2.8 The signaling Petri net-based simulator

Several studies argued the connectivity of a network alone can provide significant insights into its dynamics (Aldana and Cluzel, 2003; Kauffman et al., 2004; Klemm and Bornholdt, 2005). Motivated by these studies, Ruths et al. (2008) proposed a Petri net based modeling and simulation approach for characterizing the dynamics of signal flow through a signaling network using token distribution and sampling. This signaling Petri net-based simulator (SPN) is a non-parametric model of cellular signaling networks. Based solely on the network's connectivity the SPN can analyze large-scale networks, and provide insights into the changes in biomolecule levels generated in response to an external stimulus.

Here, a description is presented to obtain the SPN formalism from the Petri net formalism. Formally, a Petri net is a 4-tuple $Q=<P,T,I,O>$

1. $P=\{p_1, p_2, \ldots, p_m\}$ is the set of places,
2. $T=\{t_1, t_2, \ldots, t_n\}$ is the set of transitions,
3. $I=\{i_1, i_2, \ldots, i_k\}$ is the set of input arcs where for all $(u,v)\in I$, $u\in P$ and $v\in T$,
4. $O=\{o_1, o_2, \ldots, o_l\}$ is the set of output arcs where for all $(u,v)\in O$, $u\in T$ and $v\in P$.

A number of tokens are assigned to each place that encodes the state of the system and is

called a marking, denoted **m**. A *marked Petri net*, $R=<Q, \mathbf{m}_0>$, is a Petri net with a marking $\mathbf{m}_0$, called the initial marking.

The state of the system changed after *executing* the Petri net. This execution accomplished via changes in marking that are induced by sequential *firing* one or more transitions. In Petri net formalism, a transition should be enabled before *firing*. The enabling criteria of a transition are determined when the number of tokens in each of its input places is at least equal to the arc weight going from the place to the transition. When an enabled transition is fired, the number of tokens removed from input places and added to output places according to arc weights and place capacities. However, the SPN does not use arc weight which considers in the very common Petri net formulation.

The SPN is an extension of the synchronized Petri net paradigm (Ruths et al., 2008). In synchronized Petri net paradigm, the firing of a transition is triggered by a specific event that occurs in the environment. Formally, a synchronized Petri net is a 3-tuple (*R*,*E*,*Sync*), where (Ruths et al., 2008):

R is a marked Petri net,

$E=\{e_1,e_2,\ldots,e_s\}$ is a set of events, and

Sync: $T \rightarrow E \cup \{\mathbf{e}\}$ is a mapping function from each transition in the Petri net to an event, and event **e** is the *always occurring event*. Any enabled transition associated with **e** is always fired immediately.

A pseudo code is given below for executing the SPN (Ruths et al., 2008).

1) Set initial marking of S

2) for i =1 to r

a) Generate random sequence of the events in E.

b) Simulate the network by executing the transitions associated with the event in the generated sequence.

c) Record the number of tokens at each node, for each time step

3) for each node, compute the number of tokens at each time unit t, averaged over r.

Places of the SPN are initialized by the number of tokens (step 2 in pseudo code). The SPN is executed by two time-scale simulation as illustrated in Figure 2.2 (Ruths et al., 2008). The smaller time scale referred to as the *firing* time scale in which a single transition is fired. The larger time-scale called time *blocks* in which, firing steps are nested, and each transition is fired exactly once. Hence, there are $|T|$ firings per block. The simulation need to be run $B|T|$ firing steps for the specified number of time blocks, B. The SPN method executes many simulation runs (Step 2 in pseudo code) in order to sample the space of possible transition rate parameters. The final markings are obtained by averaging over these runs (Step 3 in

pseudo code).

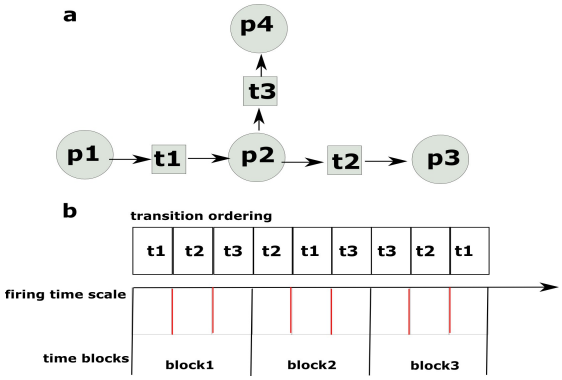

Figure 2.2. Illustration of time structure. a) A toy example of the Petri net (Places p1, p2, p3 and p4; transitions t1, t2, and t3). In the Petri net model, two different firing order of the transitions can give two different tokens distribution of places. **b)** The evaluation of transitions during a run. In each time block, every transition is fired once. The firing of each transition corresponds to one firing time scale. The order of the transition firing is shuffled during each time block in order to sample the space of possible transition firing rates.

2.9 Kinetic parameter optimization

In the kinetic parameter optimization, parameter values which are optimal in some desirable sense (providing a minimum to some objective function) are searched. An objective expresses a desired goal in mathematical terms. A desired goal can be to obtain kinetic parameters which can provide concentrations of the metabolites close to that observed in the experiment. An objective function f can be formulated to achieve the goal.

$$f_{\mathbf{p}} = \sqrt{\sum_{x \in M} \left([x]_e - [x]_{s,\mathbf{p}} \right)^2} \qquad (2.10)$$

where x is a given metabolite of a set of metabolites M of our metabolic model, $[x]_e$ and $[x]_s$ are concentration values of a metabolite x from experiments and simulations, respectively, and $\mathbf{p}$ is a set of kinetic parameters. $[x]_{s,\mathbf{p}}$ is the steady state concentration of metabolite x during simulation which depends on $\mathbf{p}$ and $[x]_e$ is the experimentally measured concentration of x. $\min f_{\mathbf{p}}$ is an optimization problem which minimizes f over $\mathbf{p}$.

Hoops et al., (2006) introduced the tool COPASI which is equipped with a number of diverse optimization algorithms that can be used for optimization of a predefined objective function depending on model parameters. COPASI comprises algorithms such as steepest descent and Levenberg–Marquardt which are based on estimating derivatives of the objective function

the Hooke–Jeeves method which is based on geometric concepts and a genetic algorithm with stochastic ranking and a simple random search. I used the genetic algorithm of COPASI in my analysis. Furthermore, Dierkes et al., (2011) developed a package, BioPARKIN, for parameter identification of large kinetic models.

3 Results

A major objective of my thesis is to identify pathways or modules of susceptibility genes, proteins, and metabolites that may be linked to the liver disease steatosis and steatohepatitis. In the first part of this chapter I determine relevant pathways that may be involved in liver disease using pathway databases (section 3.1). In the second and third part of this chapter I pinpoint most relevant metabolic modules related to liver diseases using constraint-based modeling techniques (section 3.2 and 3.3). Other aims are to identify candidates for drug targets in liver diseases, to model most relevant pathways that may be involved in liver diseases for a better understanding of the disease progression, to describe new potential metabolic drug targets for liver disease, and to elaborate on the prediction of changes of metabolic concentrations based on changes in expression data. I have identified *in silico* candidates for drug targets in liver disease using constraint-based modeling techniques (section 3.4). Furthermore, I model the arachidonic acid pathway that is identified by pathway analysis of gene expression data as relevant to liver disease (section 3.5). The model is used to study changes in metabolic concentrations at qualitative and quantitative level (section 3.5).

To understand the above mentioned human liver diseases one would benefit from analysis of a broad range of molecular data of normal as well as disease liver samples. However, such data may not be readily accessible as a large cohort of patients would be required for proper analysis. As a model system of human liver disease, I have used data from a mouse model system to understand liver disease such as steatosis and steatohepatitis. Three different mouse strains A/J, C57Bl6 and PWD were used in this study. Three biological replicates of each mouse strain were fed a DDC-supplemented diet that can reproduce steatohepatitis-like phenotypes. Histological analysis of liver samples of mice of all three different strains were used to correlate the results. Steatohepatitis-like phenotypes were found to be high in the mouse strain A/J, low in the mouse strain C57Bl6, and the mouse strain PWD did not show such a phenotype. The mouse strain PWD was found to be of higher degree of steatosis or fatty liver than the mouse strains C57Bl6 and PWD. RNA-Seq expression data along with some selected proteomics and metabolomics data were measured. These data sets were generated within the IMGuS project and used in this study for pathway analysis and computational modeling to achieve the aforementioned objectives.

3.1 Pathway analysis

High-throughput experimental data by itself does not produce biological findings. However, the data can for example be interpreted in the context of biological processes using pathway analysis. Pathway analysis can be used to identify genes, proteins and metabolites associated

with the etiology of a specific disease. Therefore, I have applied pathway analysis using gene expression data of mouse liver samples to identify relevant pathways that may be involved in liver disease.

3.1.1 Phenotype, susceptibility and metabolic model associated genes

Phenotypes of liver samples of all three mouse strains after 8 weeks of a DDC-supplemented diet were characterized by histological analysis (Figure 3.1a). Based on the degree of the phenotypes of the mouse strains, A/J and C57Bl6 can be categorized as high and low steatohepatitis-like phenotypes, while the mouse strain PWD shows no steatohepatitis-like phenotype. The degree of steatosis in PWD mice was higher than in C57Bl6 and A/J mice.

First, to address the DDC-treated effect irrespective of strain-specific effects a principle component analysis (PCA) was performed based on 2813 genes, which were differentially expressed in at least one mouse strain due to DDC-treatment (Figure 3.1c). While healthy mice show strain specific expression profiles, DDC-treated mice profiles cluster together in PCA, implying a similar overall response due to DDC treatment irrespective of the strain. Albeit of these similarities in gene expression changes upon DDC-treatment, it was expected that severity of phenotypic response is linked to gene sets specifically activated ot repressed in individual strains. To address this question, differentially expressed genes were identified between control *vs* DDC-treated mice for each strain to address the strain-specific response (section 2.7). The overlap of strain-specific differentially expressed genes is shown in Figure 3.1b. 471 differentially expressed genes were found only in A/J and I call them susceptibility genes because this strain shows the most significant steatohepatitis phenotype.

NAFLD has been defined as the metabolic syndrome that is associated with the insulin-resistance syndrome (Marchesini et al., 2003). Hence, the GMM model was (section 2.1) used as a reference to identify metabolic pathways of related genes. Irrespective of strain-specific effects, a group of 288 genes coding for metabolic enzymes were identified as differentially expressed (p-value < 1e-06) after DDC treatment. This set of 288 genes was subsequently used in the pathway analysis to identify affected metabolic pathways due to DDC treatment.

3.1.2 Pathway analysis using ConsensusPathDB

A pathway analysis of the metabolic profiles and gene expression data was performed with ConsensusPathDB to identify steatohepatitis-specific pathways (Kamburov et al., 2011a). Pathway over-representation analysis of the aforementioned 471 susceptibility genes obtained only in A/J identified nucleotide, histidine, beta-alanine and purine metabolism upon the top-ranked deregulated pathways. Within the beta-alanine pathway I found Srm encoding spermidine synthase (SPDS) as being upregulated in A/J (2.4 fold in A/J compared to 0.77, 0.75 for C57Bl6 and PWD, respectively) as well as the related metabolites

spermidine (SPD) and spermine (SPM) that form a module of the hepatic S-adenosylmethionine (SAMe) metabolism (Figure 3.1d). SAMe is needed for methylations of DNA, RNA and lipids, synthesis and catabolism of SAMe is tightly regulated and changes in SAMe level might lead to fatty liver disease and the development of HCC (Mato et al., 2013).

Measured transcriptomics values of genes and metabolite concentrations of hepatic SAMe metabolism were found to be affected in the DDC-treated state (Figure 3.1e and 3.1f). Expression of *Mat1a*, *Srm*, *Sms*, *Dnmt1*, *Ahcy*, and *Bhmt* and concentration changes of spermidine, spermine and putrescine show a different behavior between A/J, C57Bl6 and PWD. Gene expression and metabolic profiles of hepatic SAMe metabolism imply that PWD has an opposite response than A/J and C57Bl6 after feeding DDC supplemented diet which resembles the differences in steatohepatitis phenotypes of the strains. The enzyme SAMe decarboxylase (SAMDC) is activated by putrescine and upregulation of spermidine synthase (SPDS) and spermine synthase (SPMS) might explain the increased concentration of spermidine (SPD) and spermine (SPM) in A/J. A higher production of putrescine in A/J lowers the K_m of SAMDC activating polyamine synthesis that may affect the concentration of SAMe (Mato et al., 2013). This might affect methylation of various substrates such as DNA, RNA and lipids, that might be one reason in disease development, such as NASH. However, it remains difficult to judge the difference in degree of steatohepatitis because expression and metabolic profiles show similar response to DDC-supplemented diet in A/J and C57BL6.

A pathway over-representation analysis was done based on 288 metabolic genes which were obtained by aforementioned metabolic model and 19 differentially regulated metabolites which were obtained by a t-test (p-value < 0.05) analysis. The arachidonic acid pathway was found in both analysis to be deregulated (p-value 1.2e-5 and 2.5e-4) after DDC-treatment. Analysis of measured metabolic concentrations of the arachidonic acid / eicosanoid pathway yield also significant changes due to DDC treatment of some metabolites of this pathway. The abundances of four metabolites of this pathway are significantly (p-value<0.05) altered due to DDC-treatment: PGD2, 5-HPETE, 15-HETE, and 15-HPETE. Based on these observations I selected the arachidonic acid metabolism as a candidate for further *in silico* analysis.

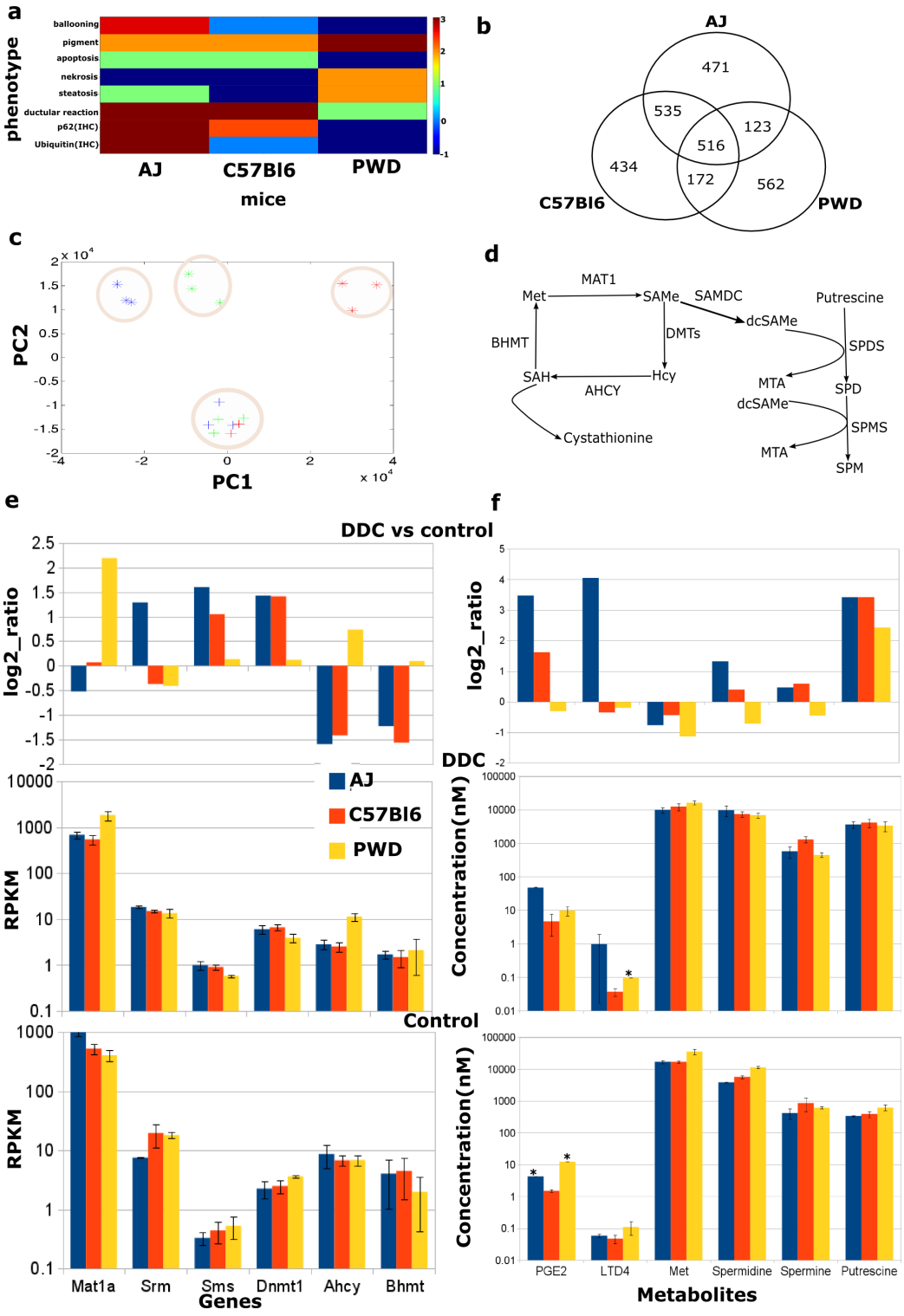
a
phenotype
ballooning
pigment
apoptosis
nekrosis
steatosis
ductular reaction
p62(IHC)
Ubiquitin(IHC)
AJ
C57Bl6
PWD
mice
b
AJ
471
535
123
516
434
172
562
C57Bl6
PWD
c
PC2
PC1
x 10^4
d
MAT1
Met
SAMe
SAMDC
Putrescine
dcSAMe
BHMT
DMTs
SPDS
SAH
Hcy
AHCY
MTA
SPD
dcSAMe
SPMS
Cystathionine
MTA
SPM
e
DDC vs control
log2_ratio
AJ
C57Bl6
PWD
RPKM
Control
RPKM
Mat1a
Srm
Sms
Dnmt1
Ahcy
Bhmt
Genes
f
log2_ratio
DDC
Concentration(nM)
Control
Concentration(nM)
PGE2
LTD4
Met
Spermidine
Spermine
Putrescine
Metabolites

Figure 3.1 Phenotypes and SAMe metabolism. a) Qualitative rating of histological phenotypes of the mouse liver samples. -1 represents the control or negative phenotype, 0 - 3 indicates different degrees of the phenotype. Immunohistochemistry is abbreviated as IHC. **b)** Venn diagram of differentially expressed genes due to DDC treatment in A/J, C57Bl6, and PWD mice. **c)** Principle component analysis (PCA) of 2813 genes that were found differentially expressed for at least one mouse strain due to DDC-treatment. * and + indicate control and DDC mice, respectively; red, green and blue represent A/J, C57Bl6, and PWD mice, respectively. Principle component 1 (PC1) explains 43% and PC2 29% of the data. **d)** S-adenosylmethionine (SAMe) metabolism. Methionine (Met) is converted to SAMe by the enzyme methionine adenosyltransferase (MAT1). SAMe is converted into S-adenosylhomocysteine (SAH) by DNA-methyltransferase (DMTs) that is subsequently converted by SAH hydrolase (AHCY) into homocysteine (Hcy) which is a substrate for Met formation by betaine-homocysteine methyltransferase (BHMT). SAMe can also be converted into spermine (SPM) via decarboxylated SAMe (dcSAMe) and spermidine (SPD). SAMDC, SAMe decarboxylase; SPDS, SPD synthase; SPMS, SPM synthase. This pathway is regulated by putrescine, which activates SAMDC. **e)** Means of RPKM values of aforementioned genes for the control and DDC-treated state of mouse liver samples. The error-bars indicate standard deviations over biological replicates. The bar chart shows log2 ratios of RPKM values of DDC-treated vs control. The genes Mat1a, Srm, Sms, Dnmt1, Ahcy, and Bhmt encode the enzymes MAT, SPDS, SPMS, DMTs, AHCY and BHMT, respectively. **f)** The bar chart shows the median concentrations of the metabolites prostaglandin D2 (PGD2), leukotriene D4 (LTD4), methionine, spermidine, spermine and putrescine. The error-bars represent the median absolute deviations. * indicates samples without a replicate.

Taken together, based on susceptible genes histidine, beta-alanine and purine metabolism were identified as relevent pathways that may be involved in liver disease. Furthermore, gene expression data and metabolic concentrations of hepatic SAMe metabolism were found to be deregulated at different levels among mice. Based on pathway analysis of gene expression data and metabolite data the arachidonic acid metabolism was found to be deregulated under DDC-treated condition. Furthermore, the metabolites PGD2, 5-HPETE, 15-HETE, and 15-HPETE of arachidonic acid pathway were found to be deregulated significantly (p-value <0.05) due to DDC-treatment.

3.2 Identifying modules and elementary flux modes

The complex functions of a living cell may be carried out in a modular way (Zhao et al., 2007). A module that shares a common function is known as a functional module. For example, a group of reactions involved in the glycolysis pathway may form a functional module and a function of this module may be the production of ATP. Identifying functional modules is crucial for understanding the basic building blocks of cellular organizations as well as cellular responses to internal and external signals. One aim of this work was to identify modules of liver metabolism. Functional modules can be defined as groups of reactions with correlated fluxes. Furthermore, liver tissue specific gene expression data were used to identify most relevant liver-specific modules which may be helpful for a better understanding of liver disease. Another aim of this work was to identify modules that comprise deregulated metabolites (and genes) between two conditions. Such modules were identified using elementary flux mode (EFM) analysis.

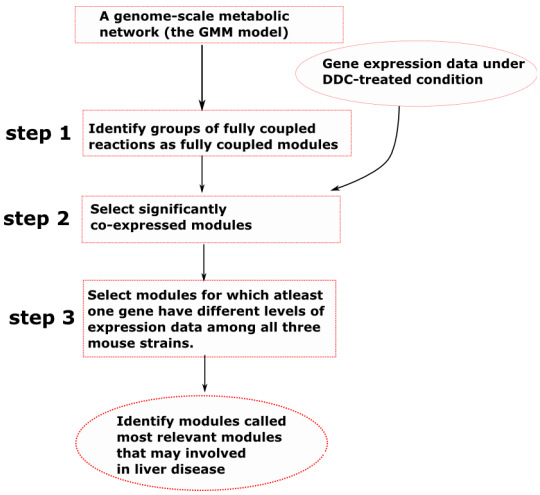

Figure 3.2 Flow chart of module identification. Flow chart diagram for the identification of relevant modules that may be related to liver diseases.

3.2.1 Identification of modules

Functional modules of metabolic networks may help to get a better understanding of the metabolism of an organism. Here, the goal was to identify most relevant modules that may be involved in the liver metabolism or in the context of liver disease. For better understanding of liver functions and liver disease progression, first of all, modules based on correlated fluxes can be identified from a large-scale metabolic model. Then, liver tissue-specific gene expression data measured in disease phenotype can be used to identify the

modules which may be involved in the disease phenotype (Figure 3.2).

A group of reactions with fully correlated fluxes or a group of fully coupled reactions is termed a fully coupled module. To achieve the aforementioned goal, first of all, fully coupled modules were identified using the framework F2C2 (Larhlimi et al., 2012). This framework resulted in 785 blocked reactions in the GMM model. In addition to F2C2, fully coupled modules were also identified by the method identifyCorrelSets that is implemented in the COBRA toolbox (Becker et al., 2007). Both methods F2C2 and identifyCorrelSets were applied to the GMM model and they resulted in 248 and 247 fully coupled modules comprising at least 2 reactions and 2 genes. 246 common modules with both methods were obtained (Figure 3.3a). Note that fully coupled modules were identified based on the structure of the GMM model and the constraint that the fluxes of reactions of these modules were fully correlated. For identification of the most relevant modules in a context of liver disease from the aforementioned 246 fully coupled modules, a kind of enrichment analysis was implemented using gene expression data of mouse liver samples under DDC-treatment conditions.

As a perquisite for enrichment analysis, I tested the hypothesis that genes associated with fully coupled modules may show co-expression under a given condition. In this study, co-expressed genes are defined as genes that have similar expression patterns in the disease phenotype of interest, e.g.,. liver disease. To test the aforementioned hypothesis, a correlation analysis was applied using gene expression data of mice under DDC-treatment conditions and the fully coupled module.

Out of 1415 genes given by entrez gene ids of the GMM model 1398 ensemble ids could be mapped by BioMart (Kasprzyk, 2011). I have applied a correlation analysis by using expression data of these 1398 genes measured under DDC-treatment conditions. For each of the aforementioned 246 fully coupled modules that were identified using F2C2, average pairwise correlation coefficients were computed for the expression data of genes associated with each fully coupled module. This pairwise correlation coefficient computation was performed using 9 expression values (3 biological replicates of each of the mouse strain A/J, C57Bl6 and PWD under DDC-treatment condition) for each of pairwise gene. A statistical analysis was performed to estimate the significance of pairwise Pearson correlation among genes. In this statistical analysis, first of all, the random distribution of the average pairwise correlation coefficients was obtained by computing the average pairwise correlation coefficient of 1,000,000 random sets of the aforementioned 1398 genes of the GMM model. For each of the identified 246 modules, p-values (expected value based on random distribution) were computed. Average pairwise correlation coefficients and p-values of 246 fully coupled modules are shown in Figure 3.3c. Out of 246 fully coupled modules that were identified using F2C2, 146 fully coupled modules have significant (p-value<0.05) average

pairwise correlation coefficient (Figure 3.3c). This result indicates that fully coupled modules show a moderate degree of co-expression.

Regulation of the expression of genes associated to the fully coupled modules may change the throughout flux through these modules. Thus, different degrees of flux regulation may be expected due to different degrees of regulation of gene expression data in the fully coupled modules. Therefore, using gene expression data of all three mice, most relevant fully coupled modules related to mouse phenotypes may be identified. Taken together, the aim is to identify the fully coupled modules for which associated genes are significantly (p-value < 0.05) co-expressed as well as regulated at different degrees among mice. Therefore, I did a kind of enrichment analysis to obtain modules that satisfy three criteria (Figure 3.2 step1, 2 & 3): i) reactions of modules are fully coupled, ii) genes of the modules are significantly correlated or co-expressed (p-value<0.05), and ii) at least one gene of the module was regulated at different degree among three mouse strains. To obtain different degrees of regulations of genes among mice an analysis was done to identify the of number of differentially expressed (DDC-treated vs control) genes of fully coupled modules (Figure 3.3b). Applying aforementioned 3 criteria most probable modules that are relevant to mouse phenotypes are illustrated in Figure 3.4. For some of the genes of the aforementioned most probable modules, different degrees of regulation among mice were observed (Figure 3.5). These different degrees of regulation may alter throughout the flux of the most probable modules which might be linked to mouse phenotypes (Figure 3.4). Thus, these modules were studied and explored in more detail to understand liver disease progression.

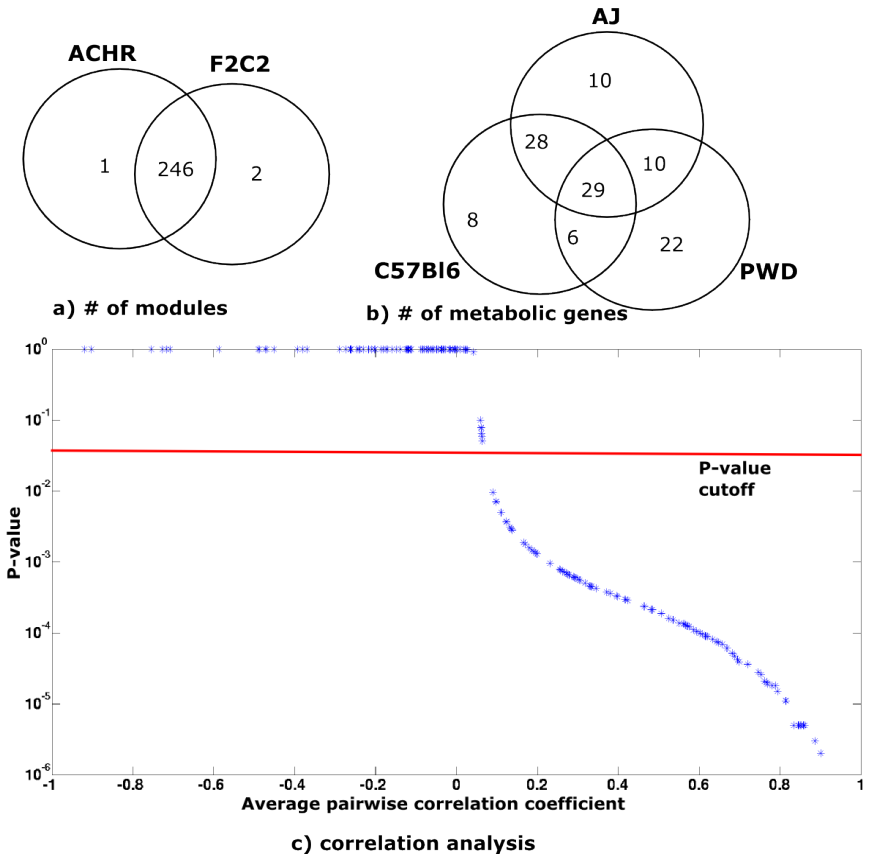

Figure 3.3. Module analysis. a) Two methods COBRA toolbox (e.g. identifyCorrelSets) and F2C2 are used to compute fully coupled modules. A fully coupled module is defined as a set of fully coupled reactions. **b)** Number of differentially expressed genes related to metabolism are identified between DDC-treated vs control for all three mouse strains. **c)** Average pairwise correlation coefficient and p-values of 248 fully coupled modules are shown. ' *' represents a fully coupled module.

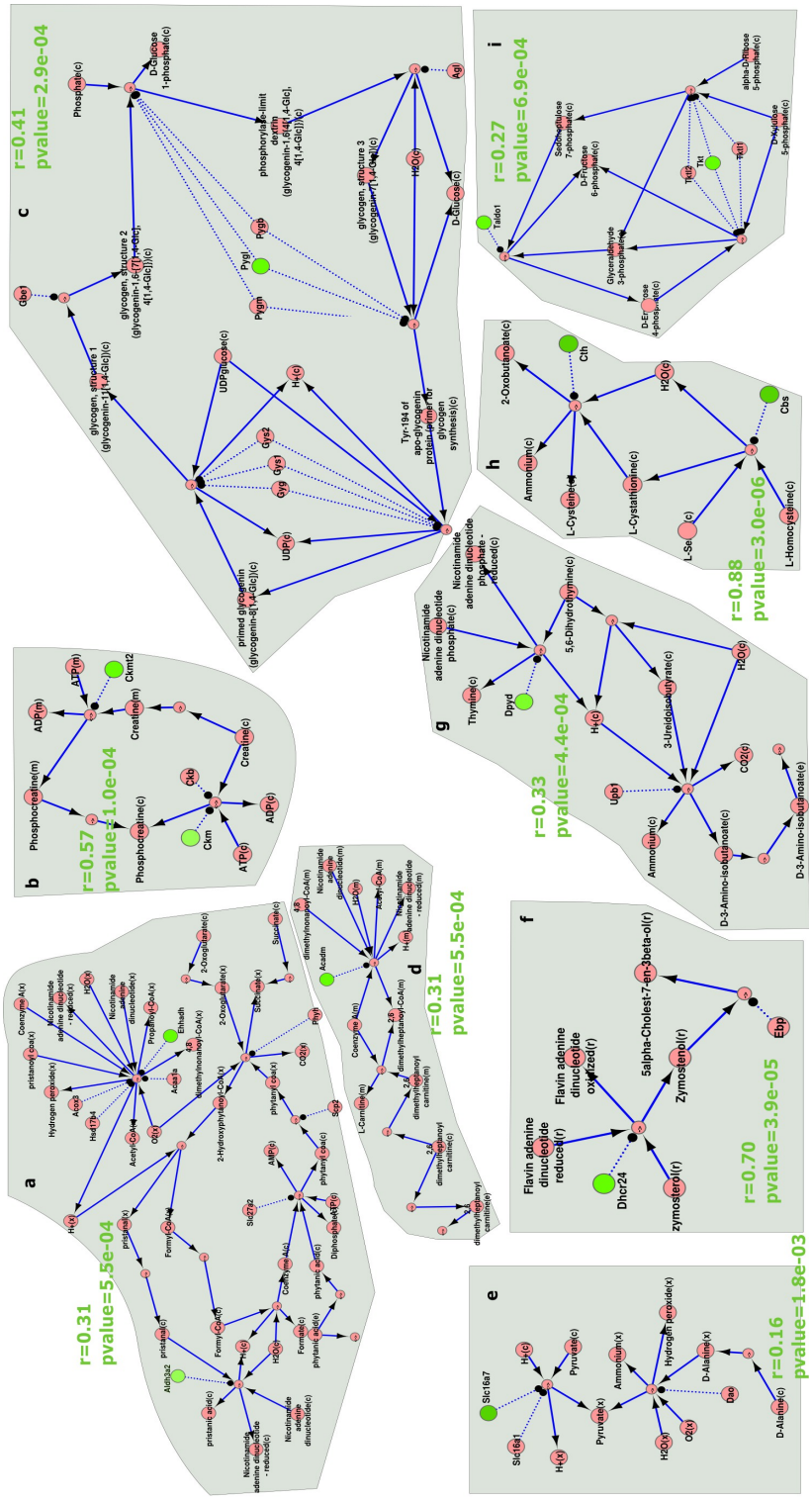
a
r=0.31
pvalue=5.5e-04
b
r=0.57
pvalue=1.0e-04
c
r=0.41
pvalue=2.9e-04
d
r=0.31
pvalue=5.5e-04
e
r=0.16
pvalue=1.8e-03
f
r=0.70
pvalue=3.9e-05
g
r=0.33
pvalue=4.4e-04
h
r=0.88
pvalue=3.0e-06
i
r=0.27
pvalue=6.9e-04

Figure 3.4. Most relevant modules that may be involved in liver disease. Green nodes indicate genes with different degrees of regulation among mice. 'r' represents average pairwise Pearson correlation coefficient of genes associated with the corresponding module. Pvalue is an expected value of average pairwise correlation coefficient of a module based on random distribution of average pairwise correlation coefficient.

a) Module of fatty acid metabolism. Phytanic acid is a branched-chain fatty acid which can be obtain through the intake of dairy products (Brink and Wanders, 2006). Phytanic acid rapidly converts to pristanic acid in rat (Avigan et al., 1966). Carbon chain of pristanic acid is one carbon atom shorter than phytanic acid, which means that phytanic acid undergoes one round of α-oxidation (Brink and Wanders, 2006). The process of α-oxidation consists of the removal of one carbon atom. In contrast, two carbon atoms are removed in β-oxidation (Brink and Wanders, 2006). External phytanic acid is transported to cytosol and then converted to phytanyl CoA by solute carrier family 27 member 2 *(Slc27a2*). Sterol Carrier Protein 2 (*Scp2)* is involved in the transport of phytanyl CoA from cytosol to peroxisome. Phytanyl CoA formed pristanal by phytanoyl-CoA 2-hydroxylase (*Phyh)*. Pristanal, is transported from peroxisome to cytosol, and formed pristanic acid by aldehyde dehydrogenase 3 family, member A2 (*Aldh3a2*). Pristanic acid derived from perxisomal pristanoyl CoA undergoes the process of β-oxidation in the presence of group of genes *Acox3*, *Hsd17b4*, *Accaa1a* and *Ehhadh*.

b) Module of creatine metabolism. Creatine is phosphorylated in the presence of ample supply of ATP and converted to the high energy compound phosphocreatine by creatine kinase. This reaction encounters in cytosol and mitochondria matrix.

c) Module of sucrose metabolism. Glycogen, a polymeric storage form of glucose, is used as a form of energy storage in animals. The synthesis of a glycogen molecule begins with "Tyr-194 of apo-glycogenin" protein which is formed by covalently attachment of a Tyr residue in a specialized initiator protein called glycogenin (Mu and Roach, 1998). Tyr-194 of apo-glycogenin protein, acts as primer, and is converted to primed glycogenin that then mediates the bulk synthesis of glycogen (glycogen, structure 1) by glycogen synthase (*Gys1* and *Gys2*). Structure of glycogen is modified and finally formed Tyr-194 of apo-glycogenin by the action of glucan (1,4-alpha-), branching enzyme 1 (*Gbe1*) along with muscle- , liver-, and brain- glycogen phosphorylase which are encoded by genes Pygm, Pygl and Pygb, respectively.

d) Fatty acid oxidation and carnitine shuttle. A medium-chain fatty acyl-CoA, 4,8-dimethylnonanoyl-CoA, that forms from a fatty acid, 4,8-dimethylnonanoic acid, by condensation of the thiol group of CoA with the carboxy group. 4,8-dimethylnonanoyl-CoA is converted to 2,6-dimethylnonanoyl-CoA in the presence of acyl-coenzyme A dehydrogenase, C-4 to C-12 straight chain (*Acadm*). 2,6-dimethylnonanoyl-CoA and L-carnitine may form 2,6-dimethylnonanoyl carnitine (since reaction is reversible; ↔) that can be transported from mitochondria, to cytoplasm, and to external medium.

e) D-alanine metabolism. This module comprised a transport reaction of each metabolite D-alanine and pyruvate along with a reaction catalyzed by D-amino-acid oxidase (*Dao*). Pyruvate is transported from cytoplasm to peroxisome by solute carrier family 16 (*Slc16a1* and *Slc16a7*).

f) Cholesterol module. This module is a part of cholesterol metabolism and comprises two reactions. In first reaction, zymosterol is converted to zymostenol by action of 24-dehydrocholesterol reductase (*Dhcr24*). In second reaction, zymostenol is converted to 5-alpha-cholest-7-en-3beta-ol by emopamil binding protein (*Ebp*).

g) Pyrimidine catabolism. This module is a part of pyrimidine catabolism in which thymine is converted to 5,6-dihydrothymine by dihydropyrimidine dehydrogenase (*Dpyd*). And then 5,6-dihydrothymine is converted to 3-ureido-isobutyrate which can be further converted to 3-amino-isobutanate by ureidopropionase, beta (*Ubp1*). 3-amino-isobutanate is transported from cytosol to extracellular region.

h) Cystein and methionine module. This module is a part of biosynthesis of L-methionine from L-cysteine. L-cysteine is converted to L-cystathionine by action of cystathionase (Cth). L-cystathionine is converted to L-homocysteine by cystathionine beta-synthase (Cbs). L-homocysteine is further converted to L-methionine which is not shown in module.

i) Pentose module. This module comprises non-oxidative phase of pentose phosphate pathway, in which alpha-D-ribose 5 phosphate (5 carbon sugar) is formed by interconversion of D-fructose 6-phosphate (6 carbon sugar) and glyceraldehyde 3-phosphate (3 carbon sugar). This interconversion process completes in 3 steps: 1) interconversion of D-erythrose 4-phosphate (4 carbon sugar) and xylulose 5-phosphate (5 carbon sugar) from D-fructose 6-phosphate and glyceraldehyde 3-phosphate from by enzyme transketolase that is encoded by genes *Tkt, Tktl1*, and *Tktl2*. 2) interconversion of glyceraldehyde 3-phosphate (3 carbon sugar) and sedoheptulose 7-phosphate (7 carbon sugar) form D-erythrose 4-phosphate and D-fructose 6-phosphate by enzyme transaldolase 1 which is encoded by gene *Taldoa1*. 3) interconversion of alpha-D-ribose 5 phosphate and xylulose 5-phosphate from glyceraldehyde 3-phosphate and sedoheptulose 7-phosphate by enzyme transketolase.

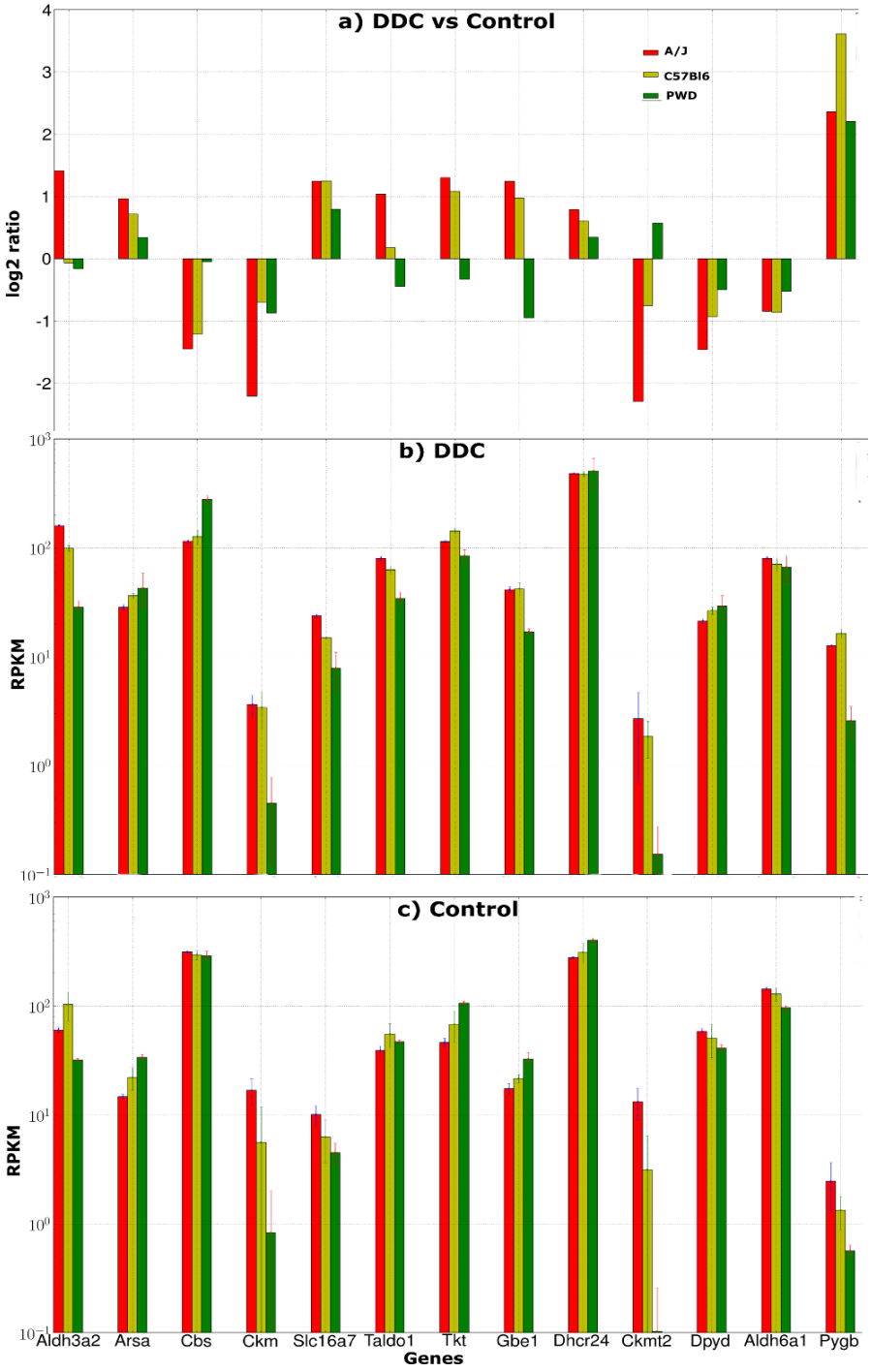

Figure 3.5. Differentially expressed genes in fully coupled modules. These genes are involved in fully coupled modules (Figure 3.4). **a)** Ratio (DDC-treated vs control). Gene expression data for **b)** DDC-treated mice and **c)** control mice.

Module of the fatty acid metabolism. Applying this approach that I propose here a module of the fatty acid metabolism was identified in most relevant modules that may involved in liver disease (Figure 3.4a). In this module, Aldh3a2 (green node) catalyzes a reaction that is involved in conversion of pristanic acid from pristanal. The expression of *Aldh3a2* was upregulated only in A/J mice (Figure 3.5). One may expect that the change of expression of *Aldh3a2* in A/J mice may alter the throughout fluxes of this module that may lead to a deregulation of metabolites. This deregulation of this module may be associated to only A/J mice of a high steatohepatitis-like phenotype. This fatty acid module comprises a reaction of peroxisomal beta-oxidation pathway that is catalyzed by enoyl-CoA, hydratase/3-hydroxyacyl CoA dehydrogenase (*Ehhadh*), acyl-CoA oxidase 3, pristanoyl (*Acox3*), Acetyl-CoA Acyltransferase 1 (*Acaa1a*), and hydroxysteroid (17-beta) dehydrogenase 4 (*Hsd17b4*) (Figure 3.4a). The expression of *Ehhadh* was found to be upregulated 1.4 and 1.3 fold in A/J and C57Bl6 mice with a DDC supplemented diet, while to be 0.6 fold downregulated in PWD mice. Upregulation of *Ehhadh* may shorten very long chain fatty acids (e.g. pristanoyl CoA) and produce more H_2O_2 during this shortening process. Accumulation of H_2O_2 may lead to oxidative stress (Rao and Reddy, 2001). Thus, one can speculate the elevated expression of *Ehhadh* in A/J and C57Bl6 mice may lead to oxidative stress that may be linked to mouse phenotypes.

Module of the creatine metabolism. Another module of the creatine metabolism is identified that has marked co-expression (Figure 3.4b). In this module, the expression of *Ckmt2* was found to be downregulated in A/J and C57Bl6 mice, and to be upregulated in PWD mice (Figure 3.5). The expression of *Ckm* was found to be downregulated for all three mouse strains between DDC-treatment and control (Figure 3.5), the degree of downregulation of *Ckm* being higher in A/J mice compared to PWD mice. The different degrees of regulation of *Ckm* and *Ckm2* among mice may alter the fluxes along with the creatine concentration of the module at different levels. Creatine plays a crucial role in reducing the toxic effects that occur by increasing production of reactive oxygen species (ROS) (Lawler et al., 2002). In other words, creatine acts as an antioxidant in the prevention of oxidative stress. Decline in hepatic functional capacity leads to decreased creatine production and lowers serum creatine levels (MacAulay et al., 2006). A creatine pool is required to be maintained by its continuous synthesis and spontaneous breakdown to creatinine. This module of creatine metabolism is extended by linking to the other pathways such as urea cycle and creatinine synthesis for better understanding of its function as well as regulation of creatine level. The extended module is illustrated in Figure 3.6. Creatine is synthesized from arginine and glycine by action of enzymes glycine amidinotransferase and guanidinoacetate methyltransferase (Figure 3.6). Creatine is either spontaneously broken down to creatinine or converted to phosphocreatine by creatine kinase (Figure 3.6).

Module of the sucrose metabolism. In Figure 3.4c, a module of starch and sucrose

metabolism is illustrated. In this module, the elevated expression of *Pygb* was found for all three mice (Figure 3.5). The enzyme encoded by *Pygl* catalyzes the reaction to release glucose-1-phosphate from liver glycogen stores. The expression of *Pygl* was found to be downregulated 0.6- and 0.3-fold for C57Bl6 and PWD mice, while it was unchanged for A/J mice. A different degree of regulation of *Pygl* may deregulate the glycogen content in the liver. In the study of Miyazaki et al., (2007) a downregulation of glycogen content is observed in the mouse fatty liver. PWD mice show high steatosis (section 3.1) and due to a downregulation of *Pygl* the deregulation of the glycogen content may be expected which is inline with the study of Miyazaki et al., (2007).

Module of the carnitine shuttle system. Another module that was identified is related to the carnitine shuttle system (Figure 3.4d). The fluxes of this module are fully correlated to the module in Figure 3.4a. The module of carnitine shuttle (Figure 3.7d) comprises mitochondrial L-carnitine shuttle pathway along with a reaction of fatty acid oxidation which is catalyzed by acyl-coenzyme A dehydrogenase (Acadm).
In liver, the system of carnitine shuttle seems to be an important player in the production of energy in cells and the perturbation of the expression data of *Acadm* under DDC-treatment may indicate that the carnitine shuttle system is perturbed. The expression of *Acadm* was found to be upregulated 1.8- and 1.3-fold in PWD and C57Bl6 mice, while it was not changed for A/J mice. The different degree of regulation may alter the throughout flux though the module of fatty acid oxidation and carnitine shuttle (Figure 3.4d). In liver, fatty acids are converted to acyl-CoA that shuttles into the mitochondria through the carnitine shuttle system for β-oxidation and generation of acetyl-coenzyme A (acetyl-CoA) (Wakil and Abu-Elheiga, 2009). Therefore, the mechanism of L-carnitine and its shuttle pathway are elaborated in more detail for better understanding of the function of the module that is shown in Figure 3.4d.

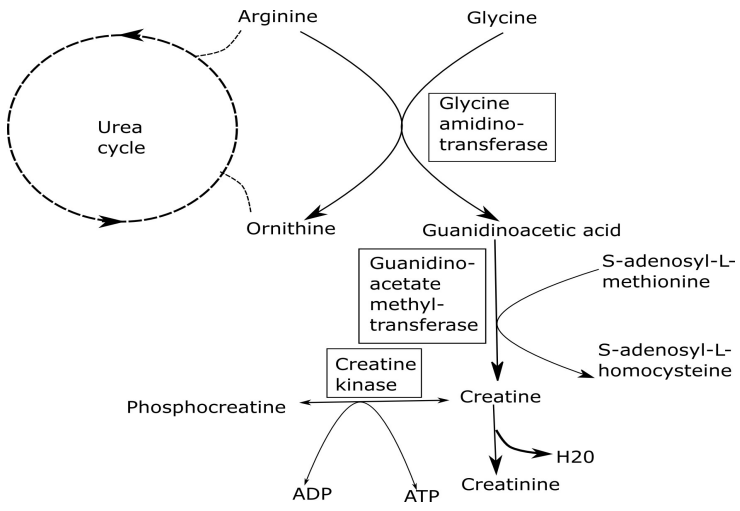

Figure 3.6. Creatine metabolism. Creatine is manufactured in the human body from arginine and glycine via urea cycle. In mammals, the guanidino residue of arginine is transferred to glycine by action of glycine amidinotransferase and forms guanidino acetate and ornithine. Guanidino acetate is further methylated by guanidinoacetate methyltransferase and converted to creatine. Creatine can be phosphorylated and yields phosphocreatine by action of creatine kinase. In addition to this, creatine is converted to creatinine for excretion in urine.

The carnitine shuttle system seems to be perturbed due to different degree of regulation of the Acadm. Thus, the carnitine shuttle system is described in more detail (Figure 3.7,(Michal and Schomburg, 2013):

1. Acyl-CoA thioester binds to carnitine and subsequently forms acyl-carnitine by carnitine-O- palmitoyltransferase I that is located on the outer mitochondrial membrane.
2. Acyl-carnitine enters the mitochondria matrix from cytosol through the shuttle by a carnitine-acyl-carnitine translocase
3. Acyl-carnitine is converted to carnitine by carnitine-O- palmitoyltransferase II that is located on the inner mitochondrial membrane. Then, carnitine returns to the cytosol.

In short, the transport of a medium-chain fatty acyl-CoA, 4,8-dimethylnonanoyl-CoA, from mitochondria to external space via carnitine shuttle system may be regulated at different levels among mice may be because of the different levels of expression of *Acadm*.

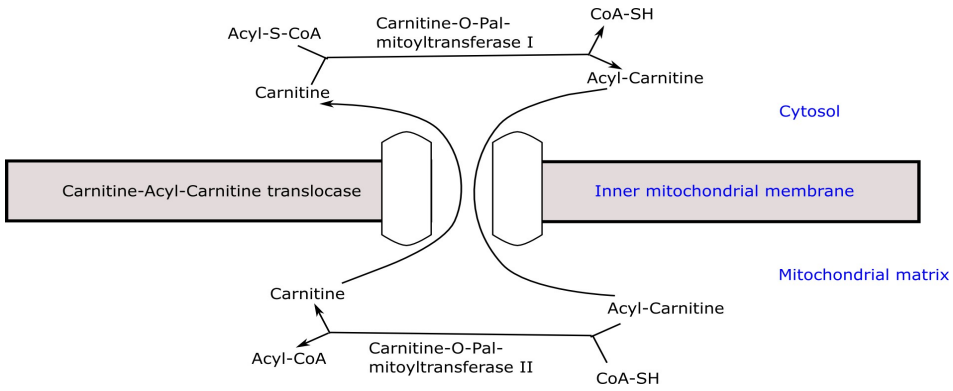

Figure 3.7. Carnitine shuttle. Acyl-CoA thioester (Acyl-S-CoA) and carnitine formed Acyl-carnitine by carnitine-O- palmitoyltransferase I. Acyl-carnitine is transported from cytosol to mitochondria matrix by carnitine-acyl-carnitine translocase. In mitochondria matrix acyl-carnitine forms carnitine which can be returned back to the cytosol. (Modified after Figure 6.1-8, p78, biochemical pathways (Michal 1998).

Module of the D-alanine metabolism. A module of D-alanine metabolism is shown in Figure 3.4e. In this module, the expression of *Slc16a7* which is a transporter of pyruvate, was upregulated for all three mice (Figure 3.5). The degree of upregulation of *Slc16a7* for A/J and C57Bl6 mice was higher than in the PWD mice. Different degree of regulation of expression data of *Slc16a7* may change the throughout flux through this module. A reaction of this module is involved in the conversion of pyruvate from D-alanine by action of *Dao* that encodes a peroxisomal enzyme D-amino acid oxidase and during this conversion hydrogen peroxide (H_2O_2) is produced. Due to different degree of regulation of expression data of *Slc16a7* one may expect enhanced production of hydrogen peroxide in A/J and C57Bl6 mice compared to the PWD mice. Wijeratne et al. (2005) reveals that the elevated concentration of hydrogen peroxide caused cell membrane leakage and DNA damage as well as oxidative stress. Oxidative stress is known to be an important player in the progression of disease ranging from steatosis to NASH (Koek et al., 2011). Thus, this module may be linked to steatohepatitis-like phenotypes to mouse strains A/J and C57Bl6.

Module of the cholesterol metabolism. Reactions of this module are catalyzed by *Dhcr24* and *Ebp* (Figure 3.4f). *Dhcr24* encodes an enzyme that catalyzes the reduction of delta-24 double bond of sterol intermediates during cholesterol biosynthesis. The degree of elevation of expression data of *Dhcr24* was higher in A/J and C57Bl6 mice compared to PWD mice. Thus, in A/J mice compared to PWD mice one may expect higher level of zymostenol accumulation. Accumulation of zymostenol is observed in a study of tamoxifen and AEBS ligand induced apoptosis (de Medina et al., 2009). Therefore, A/J and C57Bl6 mice are expected to be more prone to apoptosis compared to PWD mice, which is inline with the apoptosis phenotypes of mice (section 3.1, phenotype description).

Module of the pyrimidine catabolism. A module related to pyrimidine catabolism was identified (Fig 3.4g). This module may be involved in the hepatic steatosis due to a previous study which revealed hepatic steatosis is induced by disruption of uridine homeostasis through the overexpression of an enzyme (e.g. uridine phosphorylase 1) of the pyrimidine catabolism and salvage pathway (Le et al., 2013). Therefore, to understand the function of this module in more detail the interactions of this module are extended to the other branches of pyrimidine metabolism, such as salvage pathway and *de novo* synthesis. The extension of this module is illustrated in Figure 3.8. The expression of *Dypd* that encodes a pyrimidine catabolic enzyme was downregulated for all three mice, but by degree: in A/J and C57Bl6 mice is higher than in PWD mice (Figure 3.5). The downregulation of *Dypd* may change the throughout fluxes of this module as well as change the concentration of D-3-amino-isobutanoate that is thymine catabolite. A thymine catabolite, beta-aminoisobutyric acid, increases fatty acid oxidation in the liver and reduced the accumulation of body fat in mice fed a standard chow (Begriche et al., 2008). Based on this study one may expect deregulation of D-3-amino-isobutanoate that may be involved in the accumulation of fat and thus also linked to fatty liver of mice.

Module of the L-cysteine and L-methionine synthesis. Another module that was identified is the biosynthesis of L-cysteine from L-methionine (Figure 3.4h). In this module, the protein encoded by *Cbs* catalyzes the conversion of L-homocysteine to L-cystathionine. The expression of *Cbs* was found to be downregulated in A/J and C57Bl6 mice, while not to be changed for the PWD mice (Figure 3.5). *Cth* encodes an enzyme that catalyzes the conversion of L-cystathionine to L-cysteine. The expression of *Cth* was downregulated 0.57- and 0.72-fold in A/J and C57Bl6 mice, and was upregulated 1.54-fold in PWD mice. The different degree of regulation of *Cth* and *Cbs* may alter the throughout fluxes of this module as well the concentration of the metabolites (e.g. L-methionine) of the module. L-methionine is an essential amino acid required for protein synthesis. L-methionine is metabolized mainly by the liver where it participated, together with ATP, in the formation of *S*-adenosylmethionine (SAMe) (Martínez-Chantar et al., 2003).

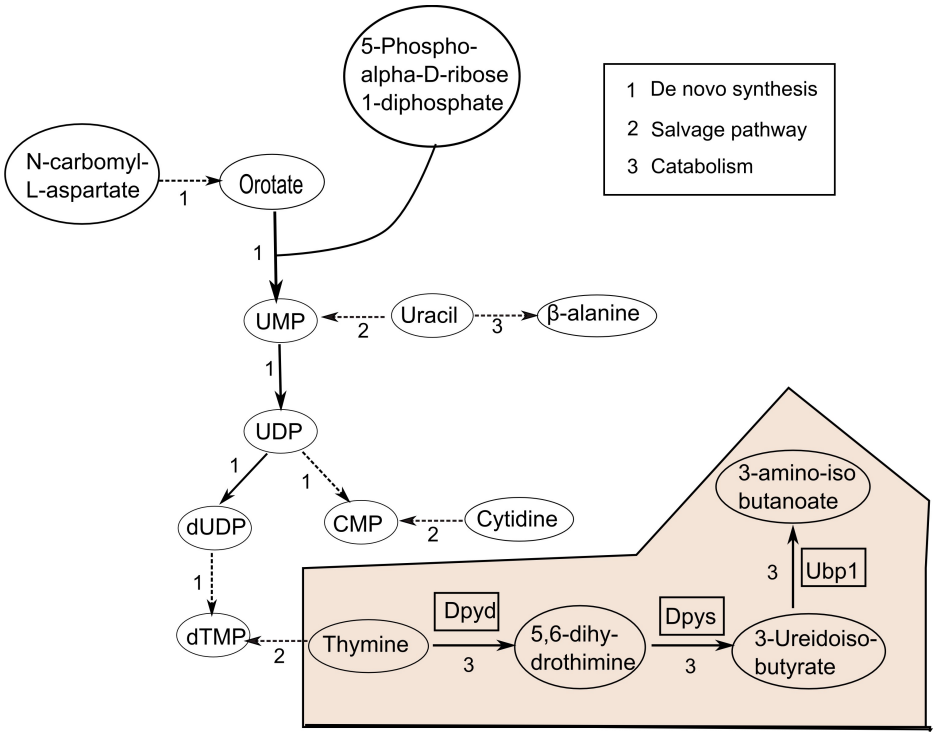

Figure 3.8. Pyrimidine metabolism. A rough sketch of pyrimidine metabolism is illustrated, which comprises three metabolic process of pyrimidines, *de novo* synthesis (1), salvage pathway (2) and catabolism (3). *De novo* synthesis of pyrimidine nucleotide (e.g uridine 5'-phosphate/UMP) starts with N-carbomyl-L-aspartate. L-aspartate and carbomyl phosphate are the source of N-carbomyl-L-aspartate. N-carbomyl-L-aspartate is converted to orotate in series of steps (dotted line). 5-phospho-alpha-D-ribose 1-diphosphate is added to orotate and then converted to UMP that subsequently converted to uridine 5'-diphosphate (UDP). UDP is either converted to cytidine-5'-monophosphate (CMP) or converted to deoxythymidine 5'-phosphate (dTMP) via 2'-deoxyuridine 5'-diphosphate (dUDP). Most of the foods comprise nucleic acid which can survive in the acid medium of the stomach and then nucleic acid can go through degradation process, which result in the release of uracil, thymine and cytidine. These free pyrimidine are reconverted to their corresponding nucleotides through salvage pathway. Animal cells degrade pyrimidine nucleotides to their component bases using dephosphorylation, deamination, and glycosidic bond cleavages. The resulting uracil and thymine are then broken down in the liver through reduction in the catabolism process. This pyrimidine catabolism process comprised the conversion of uracil to β-alanine along with the conversion of thymine to 3-amino-isobutanoate. Enzymes *Dpyd*, *Dpys* and *Ubp1* are involved in the series reactions between thymine to 3-amino-isobutanoate. These group of reactions were identified as a module of pyrimidine catabolism (Figure 3.4 g) in the module analysis that shown by shaded area.

The deregulation of concentration of SAMe may lead to fatty liver disease (e.g. NASH) and to the development of hepatocellular carcinoma (HCC) (Mato et al., 2013). Alteration of concentration of L-methionine may deregulate SAMe that may be linked to steatohepatitis-like phenotypes of mouse strains A/J and C57Bl6. SAMe is not part of this module; thus, one can extend this module for including SAMe and its interactions.

Module of the Pentose phosphate pathway (PPP). A module of pentose phosphate pathway (PPP) and the oxidative and the non-oxidative phase of PPP are illustrated in (Figure 3.4i & 3.9). In oxidative phase, NADPH is generated, and in non-oxidative phase, ribose 5-phosphate is generated via other sugars of glycolysis pathway, such as glyceraldehyde 3-phosphate and fructose 6-phosphate. Aforementioned module comprised the non-oxidative process of PPP in which the enzymes encoded by *Taldo1* and *Tkt* are involved in catalyzing the reactions of the module (Figure 3.4i). The expression of *Taldo1* was found to be (>2-fold) upregulated in A/J mice, and to be not changed in C57Bl6 and PWD mice (Figure 3.5). An elevation of the expression of *Tkt* (>2-fold) was observed in A/J and C57Bl6 mice, while it was not to be changed in PWD mice (Figure 3.5). Pertubation of gene expression data of *Taldo1* and *Tkt* may perturbed the throughout fluxes of this module and subsequently deregulate the metabolic concentrations. Experimentally measured concentration of ribose 5-phosphate and ribulose phosphate together is found to be upregulated (>1.5 fold) under DDC-treatment condition in all three mice.

Taken together, I introduced a novel enrichment method to identify fully coupled modules that may be related in disease conditions using metabolic networks and expression data. Applying this method I described several important modules of a metabolic network that may be involved in the liver disease. For example, modules of fatty acid, D-alanine and creatine metabolism are linked to oxidative stress that may cause the steatohepatitis-like phenotypes. Within a module related to carnitine shuttle system the transportation of 4,8-dimethylnonanoyl-CoA is found to be affected at different levels among the three mose strains. Furthermore, a module of cholesterol metabolism is linked to apoptosis that is a phenotype for the evaluation of the steatohepatitis. Moreover, a module of pyrimidine catabolism is linked to the steatosis or fatty liver as well as in a module of pentose phosphate pathway the experimentally measured metabolites and genes are found to be perturbed due to DDC-treatment. In addition, a module of cysteine and SAMe metabolism is identified and the deregulation of SAMe may develop steatohepatitis along with hepatocellular carcinoma. These modules may be used to develop small scale kinetic models for a better understanding the dynamics of disease progression.

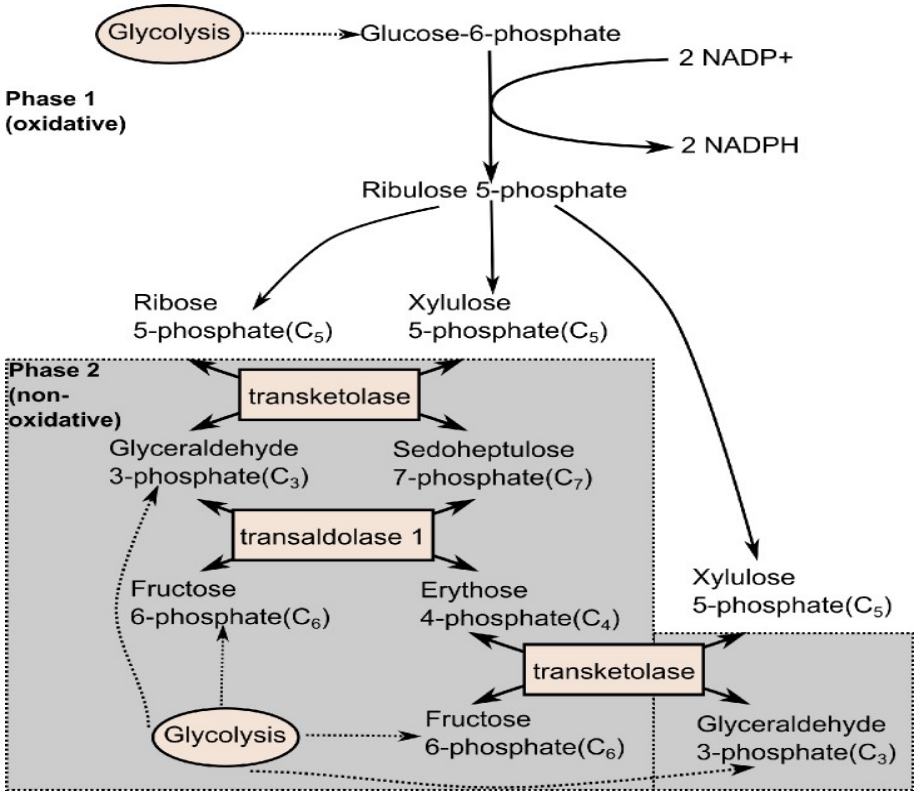

Figure 3.9. Pentose phosphate pathway/PPP. PPP is a process that generates NADPH and Ribose 5-phosphate (5 carbon sugar). It proceeds in two distinct phases. The first is oxidative phase which comprises generation of NADPH, and second is the non-oxidative synthesis of ribose 5-phosphate. During oxidative phase, glucose 6-phosphate, is formed in glycolysis pathway, is converted to ribulose 5-phosphate by production of NADPH which is used in reductive biosynthesis reactions (e.g fatty acid synthesis). And then, ribulose 5-phosphate is either converted to ribose 5-phosphate or converted to xylulose 5-phospahte. During non-oxidative phase, ribose 5-phosphate (5 carbon sugar) is formed from interconversion of glyceraldehyde 3-phosphate (3 carbon sugar) and fructose 6-phosphate (6 carbon sugar) which are produced in glycolysis pathway. Enzymes transketolase and transaldolase 1 are involved in non oxidative phase of PPP. Non-oxidative phase of PPP was identified in the module analysis (Figure 3.4 i).

3.2.2 Identification of elementary flux modes

The aim of this work is to identify pathways or modules related to liver diseases. To achieve this, one can identify modules which satisfy the following criteria: metabolites and genes of the modules are deregulated in the liver disease related molecular data. Such modules may be identified using elementary flux modes (EFMs) and liver-specific molecular data.

Elementary flux mode (EFM) analysis is a useful tool to identify pathways of a minimal set of enzymes that operate at steady state of the cellular metabolism. These pathways represent independent cellular physiological states. The GMM model (section 2.6) is used in this study to obtain a complete set of EFMs of the model. Such a task is computationally intensive. Computation and analysis of EFMs of the GMM model is difficult due to the exponential growth of the number of EFMs. de Figueiredo et al., (2009) have proposed a method to obtain k-shortest elementary flux modes from a large-scale network (section 2.2).

Often only differentially regulated metabolites between two conditions (e.g. control vs disease) are identified from a list of few hundreds of metabolites. One wants to know why these identified metabolites are differentially regulated, for example, because of the changes in the upstream regulatory systems which is accessible via transcriptomics and proteomics data. There are tools such as ConsensusPathDB which consider both differentially expressed genes and differentially regulated metabolites in a pathway analysis (Kamburov et al., 2011b). However, such a method does not take into account compartmentalized metabolic networks as well as cross-talks between metabolic pathways. Alternatively an elementary flux modes enrichment analysis may be used in which, firstly, shortest EFMs can be computed from compartmentalized metabolic networks. Secondly, most relevant EFMs related to disease can be enriched by using its molecular data of disease. This approach can have two advantages: 1) it uses a curated compartmentalized metabolic model and considers cross-talks of metabolic pathways, and 2) identified pathways or modules can be used to construct a small-scale dynamic model that allows a better understanding of the metabolic regulation in a disease context.

Here, I have applied an EFMs enrichment analysis on the metabolic profiles and gene-expression data of A/J, C57Bl6 and PWD mice from the DDC-treatment experiments and respective controls. A t-test (p-value<0.01) analysis was performed to get differentially regulated metabolites between DDC-treated vs control data. The assumption is, that perturbation of metabolic concentrations may occur due to perturbation of expression data of enzyme-coding genes. I applied a two-step analysis to identify such EFMs in which both metabolites and genes were perturbed. In a first step, I identified a set of shortest EFMs containing a differentially regulated metabolite. In a second step, most relevant EFMs that may be involved in liver disease were enriched using a set of shortest EFMs and gene expression data of DDC-treated as well as control conditions.

Several differentially regulated metabolites between DDC-treated vs control are present in the GMM model: putrescine, lanosterol, D-glucose, L-Asparagine, L-glutamate, L-citrulline,

L-argine, L-aspartate, L-glutamine, ornithine, desmosterol, fumarate, prostaglandin D2 and 5- and 15-HPETE. For each of these metabolites, those reactions were identified from the GMM model (section 2.1) for which the metabolite is a product. The identified reactions are members of a set S. An additional constraint, $z_r=1$, was included in the method of de Figueiredo et al. (section 2.2). This constraint makes sure that the reaction r is present in the shortest elementary mode. I have applied such a modified k-shortest EFM method and for each reaction r in the S, 50 shortest EFMs were computed. However, EFMs could not be identified within a reasonable time for those reactions that have products lanosterol, 5- and 15-HPETE, prostaglandin D2, and desmosterol. Thus, metabolites lanosterol, 5- and 15-HPETE, prostaglandin D2, and desmosterol are not considered in the EFM analysis. 1530 shortest EFMs were obtained through the remaining differentially regulated metabolites putrescine, D-glucose, L-asparagine, L-glutamate, L-citrulline, L-argine, L-aspartate, L-glutamine, ornithine, fumarate which are used in an EFM enrichment analysis.

Expression data was used to reduce the size of 1530 EFMs. For identification of differentially expressed EFM, a score is defined for which the expression data of the genes that participate in the EFM was used. This score is given below:

$$score = \sum_{r \in ER} \sum_{g\ participate \text{ in } r} \left| log2(ratio_g) \right| * \left| \log_{10}(pvalue_g) \right| \tag{3.1}$$

where, ER is the set of reactions of a given EFM that is catalyzed by an enzyme, *g* is the gene which encodes an enzyme and participates in the reaction *r*, ratio is obtained by the division of the mean value of expression data over DDC-treated conditions by mean value of expression data over control conditions, p-value is obtained by a t-test between expression data of DDC-treated *vs* control conditions.

The EFMs obtained after analysis EFMs were in descending order according to the score and two EFMs were identified in top 200 (Figure 3.10) that comprise three differentially regulated metabolites. L-arginine, ornithine and putrescine with in the first EFM (EFM1; Figure 3.10a) along with D-glucose, L-glutamine and L-asparagine with in the second EFM (EFM2; Figure 3.10b) were found to be deregulated between DDC-treated *vs* control data. Genes *Aoc3*, *Slc7a1*, *Gpx2*, *Gsr*, *Abp1*, *Slc7a6* and *Slc3a2* along with the metabolites L-arginine, ornithine and putrescine of the first EFM (Figure 3.11a) shown high |log2(ratio)|-value and |log10(p-value)|-value. In the second EFM, an increase of |log2(ratio)| and | log10(p-value)| of the metabolites D-glucose and L-Asparagine along with the genes *Slc5a1*, *Slc1a5* and *Slc5a9* were found (Figure 3.11b).

Most genes and metabolites showed a consistent trend when comparing DCC vs control mice. Only few candidates appeared as relevant when comparing different DDC-treated mouse strains among each other. However, some genes of EFM1 (Figure 3.10a) were found to be differentially regulated among mice (Figure 3.11b). The expression of *Aoc3* was found to be upregulated in C57Bl6 and PWD mice, and to be downregulated in A/J mice (Figure 3.11b). While an elevated expression is found for *Gsr, prdx1*,*Slc9a1* and *Aoc2* for all three mice, the extend of upregulation was higher in A/J and C57Bl6 mice comparing to PWD mice (Figure 3.11b). The expression of *Arg1* was found to be upregulated in PWD mice, and to be not changed in A/J and C57Bl6 mice (Figure 3.11b). A downregulation of *Slc6a12* was found in all three mice, and quantitative effect appeared as higher in A/J and C57Bl6 mice comparing to PWD mice. The expression of *Odc1* was found to be highly upregulated in A/J mice comparing to PWD mice, and to be downregulated in C57Bl6 mice. Experimentally measured metabolic concentration of glutamine was found to be downregulated in PWD mice, but not changed in A/J and PWD mice. An elevated concentration of putrescine was found for all three mice, but the degree of elevation is higher in A/J and C56Bl6 mice than in PWD mice. The downregulation of glucose concentration was found in all three mice. Glucose was found to be at higher degree of downregulation in PWD mice compare to A/J and C57Bl6 mice.

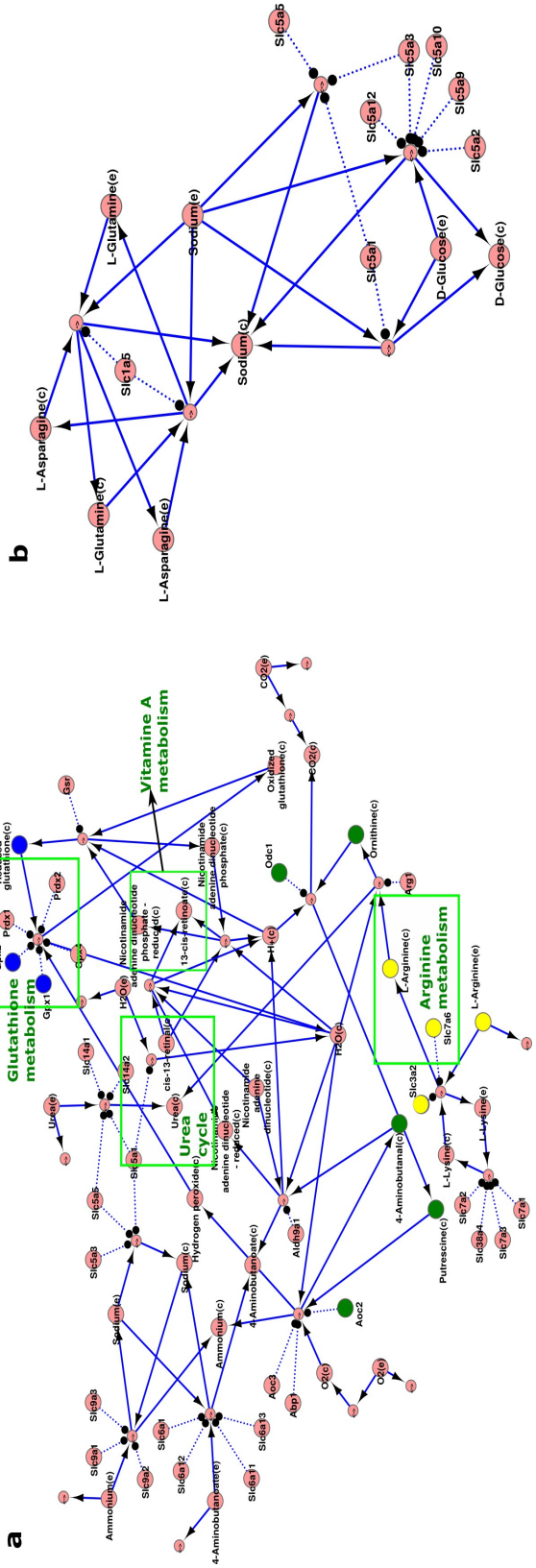
b
Slc5a5
Slc5a3
Slc5a10
Slc5a9
Slc5a12
Slc5a2
Slc5a1
L-Glutamine(e)
Sodium(e)
D-Glucose(e)
D-Glucose(c)
Sodium(c)
Slc1a5
L-Asparagine(c)
L-Glutamine(c)
L-Asparagine(e)
a
Vitamine A metabolism
Glutathione metabolism
Urea Cycle
Arginine metabolism
Gsr
Reduced glutathione(c)
Oxidized glutathione(c)
Pdx2
Pdx1
Gpx2
Gpx3
Gpx1
Nicotinamide adenine dinucleotide phosphate - reduced(c)
Nicotinamide adenine dinucleotide phosphate(c)
13-cis-retinoic(c)
H+(c)
H2O(e)
cis-13-retinal(c)
Slc16a1
Slc16a2
Urea(e)
Slc5a8
Slc5a3
Sodium(c)
Hydrogen peroxide(c)
Nicotinamide adenine dinucleotide - reduced(c)
Nicotinamide adenine dinucleotide(c)
Aldh3a1
4-Aminobutanal(c)
4-Aminobutanoate(c)
Ammonium(c)
Aoc3
Abp1
O2(c)
O2(e)
Aoc2
Odc1
Ornithine(c)
Arg1
L-Arginine(c)
L-Arginine(e)
Slc7a6
Slc3a2
L-Lysine(c)
L-Lysine(e)
Putrescine(c)
Slc7a2
Slc38a4
Slc7a3
Slc7a1
Slc9a3
Slc6a1
Slc6a3
Slc6a11
Slc9a1
Slc9a2
Ammonium(e)
4-Aminobutanoate(e)
CO2(e)
CO2(c)

Figure 3.10. Identified EFMs. a) EFM comprises L-arginine, ornithine and putrescine. This module comprises several reactions from several pathways, such as urea cycle, glutathione, arginine, and vitamine A metabolism (pathways are shown with green boxes). In this module, L-arginine (yellow) is transported from the extracellular space to the cytosol by solute carrier family (Slc3a2 , Slc7a6; yellow) and then converted to ornithine (green) and produces urea by arginase 1 (Arg1; green). Ornithine is converted to putrescine (green) by ornithine decarboxylase 1 (Odc1; green)) that can be further converted to 4-aminobutanal (green) by action of amine oxidase, copper containing 2 (Aoc2; green). 4-aminobutanal is secreted to extracellular region. This module comprised a reaction of glutathione metabolism by which reduced glutathione (blue) is converted to oxidized glutathione by glutathione peroxidase (Gpx1,Gpx2; blue).

b) EFM comprises D-glucose, L-glutamine and L-asparagine. This EFM participates in transport of D-glucose, L-glutamine and L-Asparagine by action of solute carrier family. Genes, such as *Slc5a1*, *Slc5a2*, *Slc5a3* encode Sodium/glucose cotransporter that is used to transport glucose from extracellular to cytosol. Solute carrier family 1 (neutral amino acid transporter), member 5 encoded by gene *Slc1a5* and is used to transport L-glutamine and L-Asparagine from the extracellular space to the cytosol.

Taken together, this analysis shows that in the EFM modules (Figure 3.11) some metabolites and genes are found to be deregulated over all mice (DDC vs control) as well as to be deregulated at different degree among mice. This indicates these EFM modules may explain aspects of the observed steatohepatitis-like phenotypes. Furthermore, these modules can be of interest for dynamic modeling to better understand their functions.

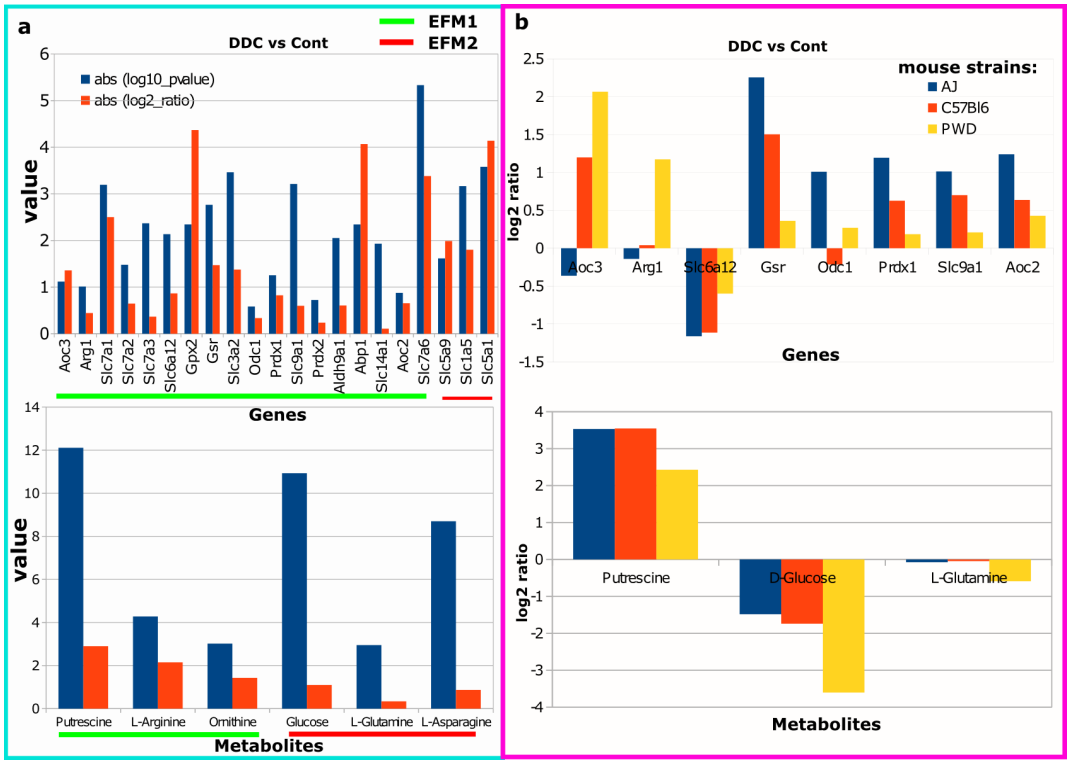

Figure 3.11 Genes and metabolites of EFMs. Genes and metabolites of the two EFM (EFM1 and EFM2 of Figure 3.10) are shown. **a)** Values of |log10(p-value)| and | log2(ratio)| of genes in upper panel and of metabolites in lower panel are shown. Green underline represents first EFM (Figure 3.10 a) and Red underline represents the second EFM (Figure 3.10 b). **b)** The upper panel indicates the log2 ratio of the genes of the first EFM (Figure 3.10a) that have different degree of regulation among all three mouse strains. The lower panel indicates log2 ratio of the metabolites D-glucose that belong to the first EFM (Figure 3.11 a) and putrescine and L-glutamine that belongs to the second EFM (Figure 3.11 b).

In conclusion, I proposed a novel method to identify EFMs which comprises both deregulated metabolites and genes between two condition using metabolic networks and gene expression data and metabolic concentration data. Applying this method on data of mouse liver samples I identified two EFMs in which the first EFM comprised deregulated metabolite L-arginine, ornithine, and putrescine while second the EFM comprised D-glucose, L-glutamine and L-asparagine. These EFMs may be used to construct kinetic models for better understanding of metabolites and enzymes regulations.

3.3 Flux balance analysis of metabolic models using gene expression data

After feeding 8 weeks a DDC-supplemented diet the three mouse strains A/J, C57Bl6 and PWD show different degrees (A/J, high; C57Bl6, low; PWD, unspecific) of steatohepatitis-like phenotypes (section 3.1.1). The DDC-supplemented diet might have changed the liver metabolism at different levels for all three mice, and an analysis of metabolic fluxes may provide an insight to understand the different degrees of liver metabolism and steatohepatitis-like phenotypes. Thus, metabolic flux analysis was performed for both control and DDC-treated mice. FBA allows an analysis of metabolic fluxes of large-scale biological networks, particularly genome-scale metabolic networks. An extension of FBA framework can be used to integrate gene expression data of all three mice for the control state as well as for the DDC-treated state into the GMM model (section 2.1; Navid and Almaas, 2012). This integration of expression data into the model may give a hint how the metabolic fluxes are regulated at different degrees among all three mice. This different degree of regulation can be traced through the GMM model and one can identify which pathways or modules of the GMM model have different degree of flux regulation that might be linked to mice phenotypes.

FBA estimates the intracellular fluxes of a genome-scale metabolic model by optimizing an objective function (e.g., biomass production) subject to constraints on fluxes. The pre-defined specific biological objective are less likely to be always valid for cells and different conditions in multicellular organisms. Thus, a pre-defined objective function may not be a true indicator for all kinds of disease or conditions. One would need to validate this pre-defined objective function before using it for any specific conditions. To circumvent this, there are some methods for which there is no need to define a pre-defined objective function (Lee et al., 2012). There are several methods such as, E-Flux and GIMME, for which it is needed to formulate a pre-defined objective function, but gene-expression can be integrated in their modeling approaches (Becker and Palsson, 2008; Colijn et al., 2009). The method E-Flux is an extension of FBA that incorporates gene expression data into FBA by constraining the maximum flux constraints as a function of measured gene expression (Colijn et al., 2009). Whereas GIMME incorporate gene expression data into their models by transforming gene expression levels to binary states. The disadvantage to GIMME method is that it requires a user-specified gene expression value as a cutoff to determine activity of the reaction. Thus, E-Flux (see section 2.3) is chosen in this study because it can use quantitative expression data without transforming to binary states. In addition, mGX-FBA is also chosen for analyzing the perturbation effect which may occur due to the inclusion of relative change of expression data in the optimization problem (section 2.4). Using both computational methods the E-Flux and mGX-FBA, expression data of all three mouse strains was integrated into the GMM model and obtained metabolic fluxes were compared. Due to limitation of

experimentally measured flux data these computationally obtained fluxes can not be validated. Thus, it is difficult to judge based on computationally obtained fluxes which methods are better or which methods have true or wrong prediction of fluxes. Of course at least these computationally obtained fluxes allow us to identify some modules of the GMM model that may indicate different degree of flux regulation that might be linked to mice phenotypes.

3.3.1 Identification of flux using mGX-FBA

For identification of different degrees of flux regulation among all three mice, gene expression data were integrated into the metabolic model. Based on gene expression data of all three mouse strains from control as well as DDC-treated conditions, the metabolic fluxes were estimated using the FBA formalism and the GMM model. The mGX-FBA method was used (see section 2.4) and this method uses relative change between two conditions (e.g., ratio between DDC-treatment and control) in an optimization problem. mGX-FBA performs the following steps for flux analysis:

1) The flux distribution v^{cont} is obtained by maximizing the flux though a pre-defined biomass reaction of the GMM model. This flux distribution is used as reference state (control mice).

2) v_i^{min} and v_i^{max} are calculated using flux variability analysis (FVA; Mahadevan and Schilling, 2003). $\bar{v}_i$ is the average of v_i^{min} and v_i^{max}.

3) First of all those genes of the GMM model were selected, for which mean expression value of three biological replicates were greater than 0.1 RPKM (Gan et al., 2010). Then for these selected genes, relative fold changes of expression data (DDC-treated *vs.* control) were calculated. A cutoff was used for moderate up- (>=1.5 fold) and moderate down-(<= 0.66 fold) regulated genes. Reversible reactions were not included in this analysis (section 2.4). Although those reactions are normally reversible, they were selected as irreversible if they carry flux in only one direction under the conditions imposed by environmental constraints. Therefore, a cutoff ($v_i^{min} \wedge v_i^{max} > 1e\text{-}06$ or $v_i^{min} \wedge v_i^{max} < -1e\text{-}06$) that represents non zero flux, was used on the result of FVA for obtaining irreversible reactions.

4) The aforementioned step 3 was applied to gene expression data of mouse strains A/J, C57Bl6 and PWD. As a result, constraints were applied to the 90, 86, and 84 reactions of the GMM model for A/J, C57Bl6 and PWD mice, respectively (section 2.4; steps 4 and 5).

Applying the aforementioned steps of mGX-FBA, steady state fluxes were calculated for A/J, C57Bl6 and PWD mice in DDC-treated state. Note that, all three control mice steady

state fluxes are identical because they were obtained by maximizing biomass function only subject to structural constraints of the metabolic network (step1).

The goal of this analysis was to identify some modules or pathways for which the different degree of flux regulation might be found among mice. Subsequently, these identified modules and their fluxes might provide an insight about differences in degree of mice phenotypes. To achieve this goal, first of all, the relative deviation of flux (section 2.4.1) was calculated for measuring the difference of flux distribution between the control and the DDC-treated state (steady state fluxes which obtained by the mGX-FBA method were used for this calculation). The GMM model is derived from Recon1 model (Duarte et al., 2007), and most of the reactions are the same in both models. The pathway map of Recon1 can be found (Schellenberger et al., 2010) and used for visualization of the GMM model. The complete set of metabolic pathways of the Recon1 is divided in three parts, carbohydrate, amino acid and lipid metabolism. Reactions of these three parts of Recon1 are visualized by coloring with relative flux deviation values using COBRA (Hyduke et al., 2011). The colored pathway maps indicate that all three mice have different degree of flux regulation for some pathways or modules which are part of carbohydrate and lipid metabolism (Figure 3.12 and 3.13).

A high increase of flux (DDC-treated *vs.* control) was observed for the pentose phosphate pathway of the PWD mice in comparison to A/J and C57Bl6 (phenotypes; Figure 3.1). The pentose phosphate pathway has two distinct phases. In the first oxidative phase NADPH is generated and in second non-oxidative phase 5-carbon sugar (ribose-5 phosphate) is synthesized from other than 5-carbon sugars, such as glucose-6-phosphate (6 carbon sugar) (Figure 3.12). NADPH is a reducing agent, which is involved in the protection against the toxicity of reactive oxygen species (ROS) or oxidative stress. Due to increased flux through the oxidative phase of pentose phosphate pathway (G6PDH2r and GND) one may expect a higher production of NADPH in PWD mice (Figure 3.12). Thus, PWD mice may be protected better against oxidative stress compared to C57Bl6 and A/J mice. This result is consistent with the pervious work on steatohepatitis: As we know from two-hit theory oxidative stress may play an important role for disease progression from steatosis to steatohepatitis (Day and James, 1998). The flux analysis of the oxidative phase of the pentose phosphate pathway may suggest that A/J and C57Bl6 mice are under oxidative stress than PWD mice, suggesting another link of ROS differences in the steatohepatitis-like phenotypes (section 3.1.1).

For PWD mice, a high decrease of flux was found through the reactions of the glycolysis pathway. This pathway is driven by the enzymes phosphofructokinase (PFK), fructose-bisphosphate aldolase (FBA), and glycerol-3-phosphate acyltransferase (GPAD) among others (Figure 3.12). The flux through GAPD was found to be downregulated in PWD mice, while it was not changed in A/J and C57Bl6. Furthermore, the flux through gluconeogenesis pathway (pyruvate to glucose) is observed, which comprises three steps: 1) oxaloacetate (oaa) to phosphoenolpyruvate (pep) catalyzed by phosphoenolpyruvate carboxykinase (PEPCK), 2) D-fructose 1,6-bisphosphate (fd6) to D-glucose 6-phosphate (g6p) catalyzed by fructose bisphosphatase (FBP), 3) g6p to glucose (glc-D) catalyzed by glucose 6-phosphate phosphatase (G6PPer). An increased flux was obtained through the PEPCK in all three mice, through the FBP in PWD and C57BL6 mice, and through the G6PPer in AJ and C57BL6 mice (Figure 3.12). Fluxes of glycolysis and gluconeogenesis pathways are perturbed at different degrees among mice. The liver plays an important role in regulation of glucose level. During fast the liver accomplishes via gluconeogenesis to maintain the glucose level. The perturbation of fluxes through glycolysis and gluconeogenesis may affect the glucose level. Thus, one may expect perturbation in glucose concentration for all three mice. Experimental the glucose concentration was found to be decreased in A/J, C57BL6 and PWD during DDC-treatment which is in inline with the expectation (Figure 3.12).

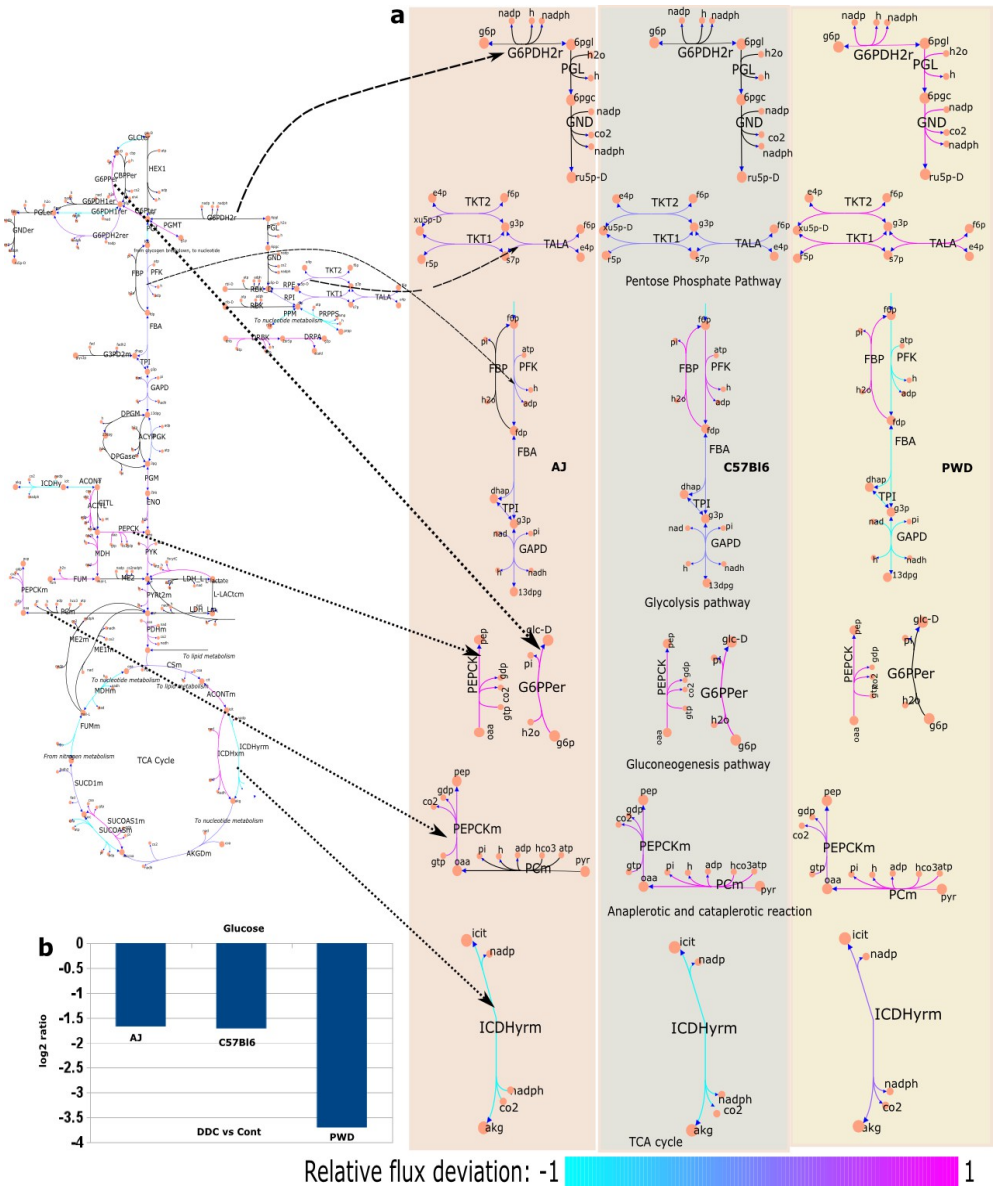

Figure 3.12. Results of flux analysis performed with mGX-FBA. Glucose metabolism. **a)** Black colored reaction arrows indicate fluxes that are very low in both control and DDC-treated conditions. Red and black colored reaction arrows indicate up and down regulation of fluxes under DDC-treatment conditions. The reactions of glycolysis, gluconeogenesis, pentose phosphate pathway, TCA cycle, anaplerotic and cataplerotic are colored with the relative flux deviation value (DDC-treated vs control). Full form of abbreviations can be found in BiGG (Schellenberger et al., 2010). **b)** Fold change of glucose concentration.

Anaplerotic and cataplerotic reactions are very useful to maintain the function of the citric acid (TCA) cycle. Due to insufficient oxalacetate the first reaction of the TCA cycle, citrate synthase, that catalyzes the condensation of oxalacetate with acetyl-CoA cannot occur. Thus, anaplerotic reaction e.g., pyruvate carboxylase, which synthesizes oxaloacetate from pyruvate in the mitochondrial matrix, is required to ensure the TCA cycle works. If intermediates (e.g., oxaloacetate) accumulate in the mitochondria, then it is important to remove them. Cataplerotic reactions are involves in the disposal of TCA cycle intermediates. An example of a cataplerotic reaction catalyzed by PEPCK is the generation of phosphoenolpyruvate from oxalacetate that is common route for gluconeogenesis. The flux through this reaction was found to be increased for all three mice under DDC-treatment (Figure 3.12). For the flux through the anaplerotic reaction mediated by Pcm an elevation was found in C57Bl6 and PWD mice, but not in A/J mice (Figure 3.12). This may suggest less production of oxaloacetate in A/J which might affect the TCA cycle. The flux through one reaction mediated by ICDHyrm (substrate, isocitrate and NADP; product, 2-oxoglutarate, NADPH, CO2) of the TCA cycle was found to be upregulated for PWD mice, and found to be downregulated for A/J and C57Bl6 mice. This may suggest the higher production of NADPH for PWD mice compared to A/J and C57Bl6 mice. This further supports the idea that PWD mice might show a different response to ROS comparing to the other mouse strains.

In all three mice the flux analysis also indicates different levels of flux regulation within the fatty acid and the cholesterol metabolism. The fatty acid metabolism module is shown in Figure 3.13a. The relative flux deviation through the reaction CSNATr is increased (red) in C57Bl6 and PWD mice. Thus, one can expect accumulation of carnitine (antioxidant) in C57Bl6 and PWD mice and it may be a reason that mice C57Bl6 and PWD mice is less vulnerable to oxidative stress in compare to A/J mice.

The relative flux deviation through the cholesterol biosynthesis metabolism is shown in Figure 3.13b. It, is slightly decreased for AJ mice, and is increased for C57Bl6 and PWD mice. The concentration of desmosterol (dsmsterol) that is a downstream of cholesterol biosynthesis pathway is observed to be upregulated in C57Bl6 and PWD mice and slightly downregulated in AJ mice. That is inline with relative flux deviation of cholesterol metabolism (Figure 3.13b & c).

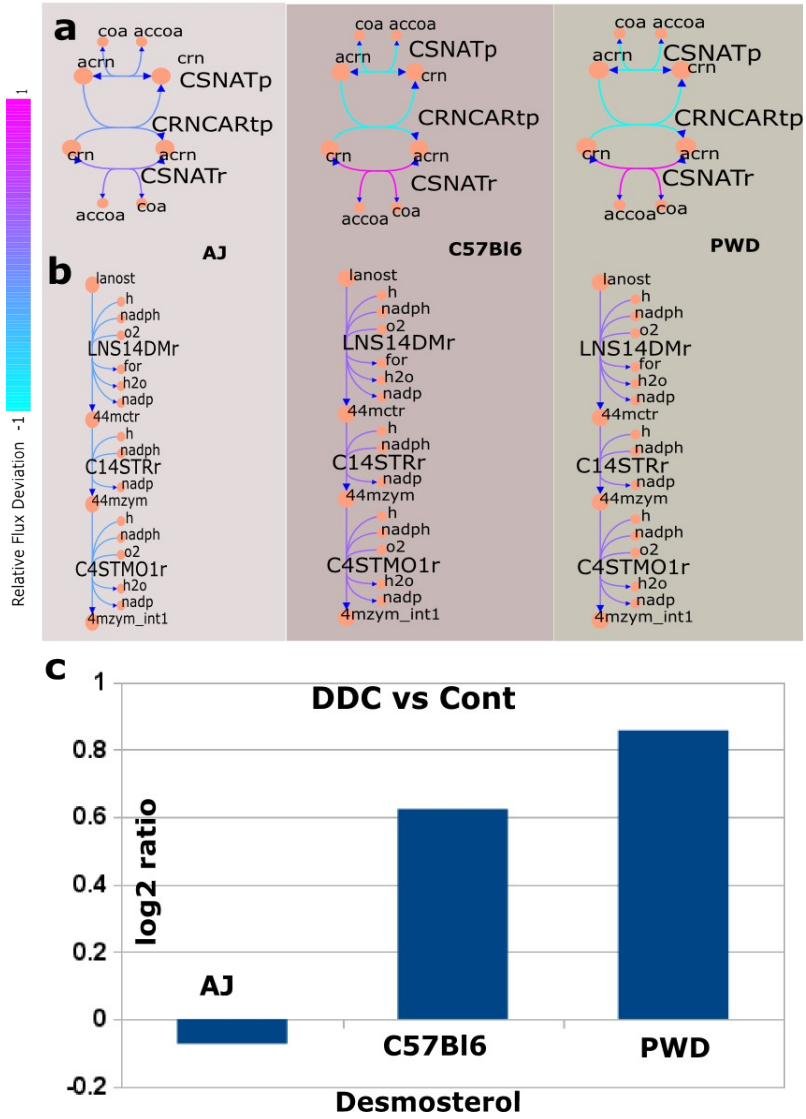

Figure 3.13. Result of flux analysis of lipid metabolism. a) A module of fatty acid metabolism. An increase (red) in the relative flux deviation of the reaction CSNATr (substrate, acetylcarnitine (acrn); product, carnitine (crn); enzyme, carnitine-acetyltransferase) is found for C57Bl6 and PWD mice. **b)** A module of cholesterol biosynthesis. The relative flux deviation of cholesterol biosynthesis is found to be decreased (slightly blue) for A/J, but increased (slightly red) for C57Bl6 and PWD mice. **c)** Fold change of desmosterol concentration.

Overall, this analysis suggest that fluxes through several independent metabolic modules might be different in high *vs.* unspecific responder to DDC. Interestingly, metabolites associated with these reactions are functionally related to ROS response, inline with previous results presented in this work.

3.3.2 Identification of flux using E-Flux

One limitation of mGX-FBA is that it uses the same steady state flux for all three controla. To overcome this, metabolic fluxes were also analyzed by using E-Flux that is combination of expression and flux (section 2.3). The E-Flux method requires an objective function and I used an already defined biomass function of the GMM model (section 2.1). The COBRA tool (Hyduke et al., 2011) was used to identify exchange reactions, i.e., reaction that link the model to the environment, in the GMM model that resulted in 455 reactions. The lower and upper bound for each exchange reaction is set to -1 and +1 (Colijn et al., 2009). Conceptually, the E-Flux method can be understood as setting the width of "pipes" around particular reactions as a function of expression state. Therefore, the low expression indicated a thin pipe around the reaction or in other words the flux through the reaction is limited.

Let a vector $\boldsymbol{v} = (v_1, v_2,...,v_r,...., v_R)$, where $v_1, v_2,...,v_r,...., v_R$ are the components of v and R is the number of reactions of the GMM model. A component of v, v_r, is derived from expression data, where r is a reaction of the model. For example, if a single gene encodes an enzyme that catalyzes a reaction r then its expression value is equal to v_r. If a reaction is catalyzed by isozymes that are encoded by more than one gene then the maximum expression of these genes is used for obtaining v_r. Then, $M=max(v)$ is calculated and v is normalized between 0 and 1 by dividing all elements by M. Note that, for those reaction for which either expression data is not available or enzymes information is not associated (e.g., decay or some transport reactions), the v_r is used as 1 that allows the maximum flux. For each reaction r of the GMM model, the upper bound is set to v_r as well as the lower bound is set to 0 or $-v_r$ depending whether r is irreversible or reversible.

Expression data of three biological replicates of A/J, C57Bl6 and PWD mice were integrated to the GMM model by using the E-Flux method and the flux through biomass reaction was maximized. Maximum flux of the biomass reaction is shown in Figure 3.14. A clear difference of biomass production rate between DDC-treated *vs.* control is observed for A/J and C57Bl6 mice, but it is difficult to observe for PWD mice (Figure 3.14). The difference (DDC-treated *vs.* control) of biomass production rate of mouse strains A/J and PWD are inline with their steatohepatitis-like phenotypes (section 3.1.1). The difference of biomass production rate is inline with the high steatohepatitis-like phenotypes of the C57Bl6 mouse strain, but the strain shows low steatohepatitis-like phenotypes.

To further replicate reasons for different degrees of flux regulations among mouse strains using an analysis was performed analogy to mGX-FBA described above. First, the relative flux deviation (DDC *vs.* control) was calculated. Subsequently relative flux deviation values were used to color the reactions of the metabolic maps of BiGG for all three mice. Furthermore, the colored metabolic maps of three mice were compared by observing the color differences. The relative flux deviation through a module of aminosugar metabolism and the glycogen synthesis and degradation pathway were found to be decreased in A/J mice, but it was increased in C57Bl6 and PWD (Figure 3.15a & b). This observation indicates that

these pathways might show a strain specific response and would be the candidates for further experimental investigation. The relative flux through glyconeogenesis pathway was found to be perturbed in all three mice, particularly the flux through FBP was found to be upregulated in C57Bl6 and PWD, but downregulated in A/J (Figure 3.15c). The pertubation of the flux through the gluconeogenesis may perturb the glucose concentration which is inline with the experimentally measured glucose concentration under DDC-treatment.

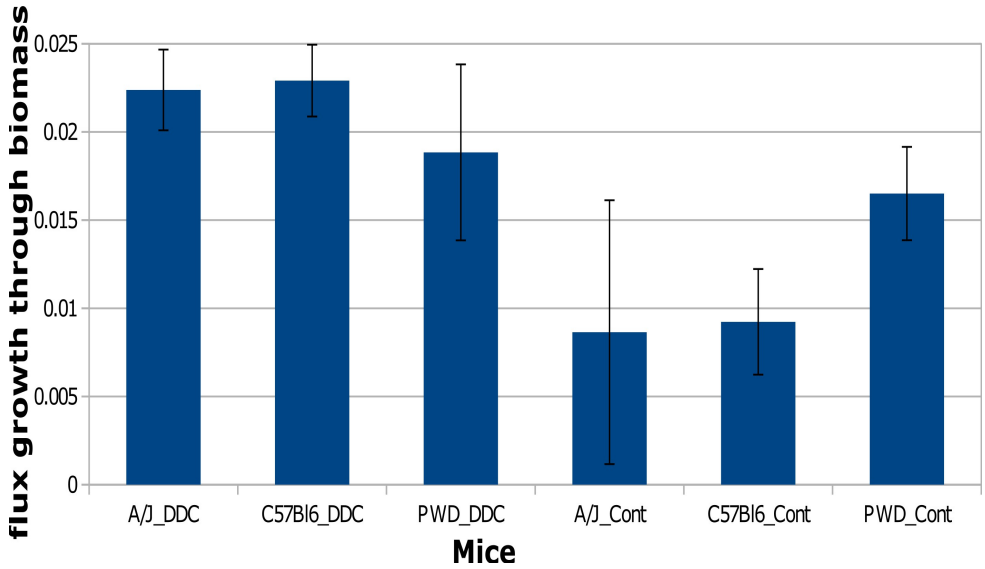

Figure 3.14. Biomass production rate. Maximum biomass production rate is higher in DDC-treated state than control for A/J and C57Bl6 mice.

The colored map of lipid metabolism is observed for A/J, C57Bl6 and PWD mice. The relative flux deviation through the reactions of desmosterol (dsmsterol) production was found to be increased in C57Bl6 and PWD mice, but not changed for A/J (Figure 3.16a). Based on this flux regulation, one may expect the concentration of desmosterol to be deregulated in C57Bl6 and PWD during for DDC-treatment. Experimentally measured concentration of desmosterol was found to be upregulated in C57Bl6 and PWD mice which is inline with the observation. Furthermore, the relative flux deviation through the bile acid synthesis pathway was found to be decreased in A/J mice, but to be changed in C57Bl6 due to DDC-treatment, while it was increased in PWD mice (Figure 3.16b). An increase of the flux through the bile acid synthesis pathway in PWD mice may increase the concentrations of metabolites of this pathway. Indeed, a downstream metabolite of the bile acid synthesis pathway, taurocholic acid/tchola measured experimentally to be upregulated (4-fold) in PWD mice which is inline with the expectation. However, based on the analysis were expected to show a downregulated flux flux through the bile acid synthesis pathways, but the concentration of taurocholic acid was found to be upregulated experimentally (1.8-fold). It may suggest A/J has strain-specific effect and to know the reason in detail one require experimental flux data. Furthermore, Tanaka et al., (2012) observed the disruption of bile

acid homeostasis in mice with NASH and also in my study the change of the flux through this pathway is observed.

In a module of fatty acid activation, reactions FACOAL1812 and FACOAL1813 are involved in the production of vaccenic and elaidic acid, respectively. Vaccenic acid occurs in dairy products such as milk and butter. Elaidic acid is trans isomer of oleic acid that occurs in various animals, vegetable fats and oils. The relative flux deviation though the reactions FACOAL1812 that and FACOAL1813 was found to be decreased in A/J mice, but it was increased in C57Bl6 and PWD (Figure 3.16c).

The high flux may lead to accumulation of fatty acid such as vaccenic and elaidic acid in PWD and C57Bl6 mice. The accumulation may form lipid droplets that is a sign of steatosis which is inline with the high steatosis of PWD mice (section 3.1.1). For all three mose strains, the metabolic pathways cholesterol, bile acid, and fatty acid have different degree of metabolic response to a DDC-supplemented diet. This finding suggests that these pathways might be good candidates for further experimental investigation.

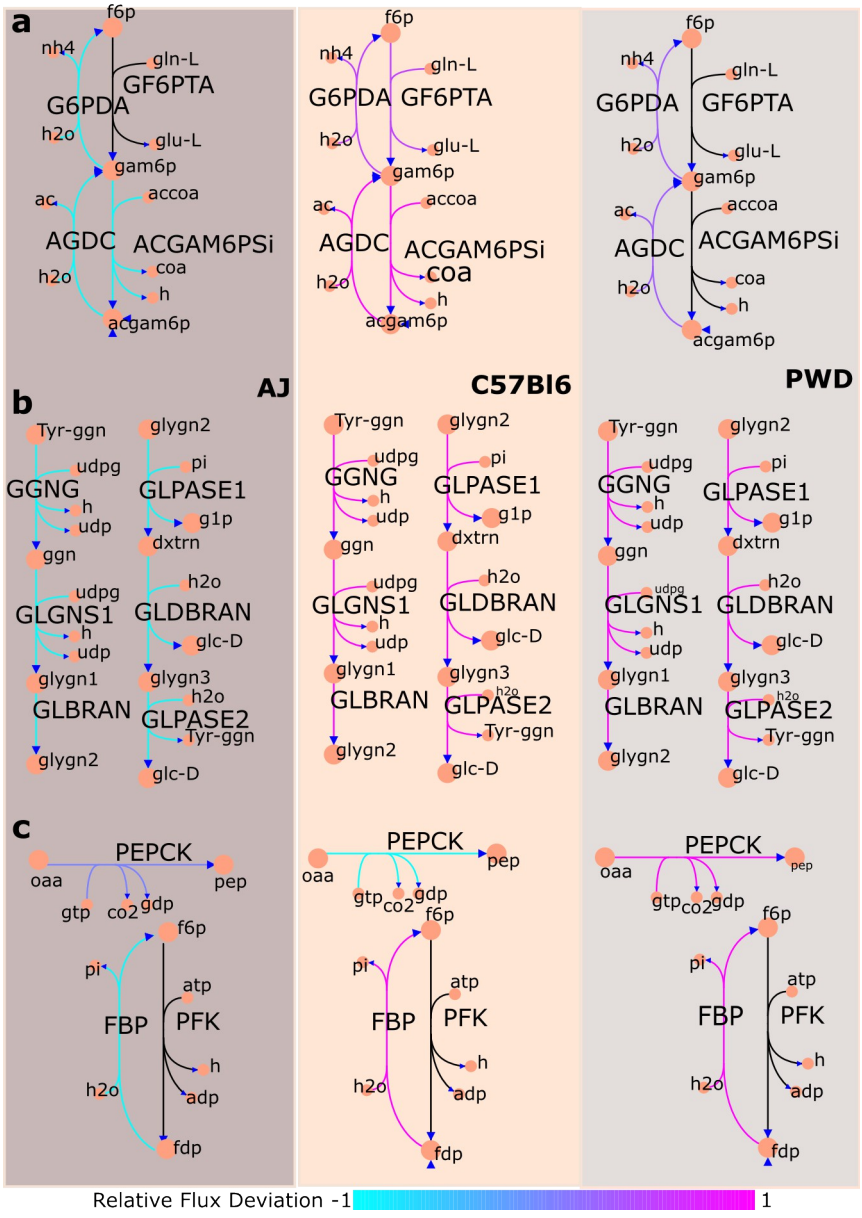

Figure 3.15. Glucose metabolism using E-Flux. a) Aminosugars module. **b)** Glycogen synthesis & degradation. **c)** Glycogenesis

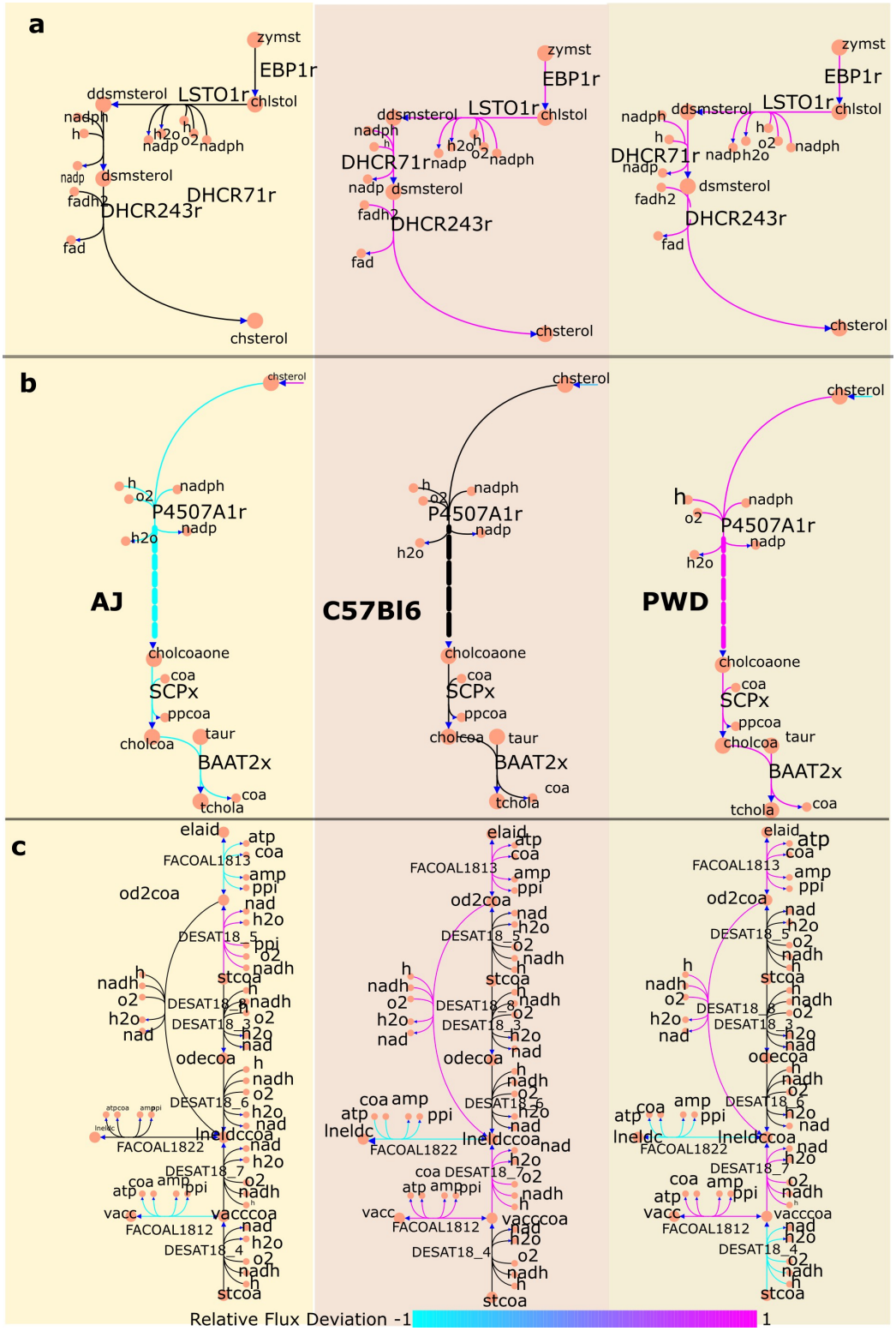

Figure 3.16. Lipid metabolism using E-Flux. a) Cholesterol module. **b)** Bile acid synthesis **c)** Fatty acid activation.

3.3.3 Comparison of the E-Flux and mGX-FBA method

The integration of different sets of gene expression data into the metabolic model may give different flux distributions. So, I expected the integration of expression data of all three mouse strains may give different degree of flux regulation which may be useful for understanding their phenotypes. Thus, E-Flux and mGX-FBA are used to identify different degrees of flux regulation among mice. As we have seen in previous sections (3.1.1 and 3.1.2) both methods were able to identify some modules or pathways that have different degree of regulation among mice. In this section results of both methods are compared.

Overlapped results of the E-Flux and mGX-FBA method. A high decrease of flux through a reaction of glycolysis (FBA) was found in PWD mice using both methods. The flux through a reaction of gluconeogenesis (FBP) was found to be upregulated in C57Bl6 and PWD mice in the results of both methods. Due to deregulation of fluxes of glycolysis and gluconeogenesis one may expect deregulation of the glucose concentration. Experimentally measured the glucose concentration was found to deregulated which is inline with the expectation. (Figure 3.12b).

The flux through the cholesterol biosynthesis was found to be upregulated in PWD mice, but it was downregulated in A/J. However, the upstream part of cholesterol biosynthesis were identified by the mGX-FBA method to be delegulated, while the downstream part of cholesterol biosynthesis is identified by the E-Flux method (Figure 3.12b & 3.15b). Albeit these discrepancies, deregulation cholesterol biosynthesis fluxes may in the end alter the concentrations of metabolites genereted by this pathway. Indeed, the concentration of desmosterol was found to be deregulated which is inline with the expectation.

In short, *in silico* fluxes of the pathway glycolysis, gluconeogenesis, and cholesterol biosynthesis were found to be altered and also the concentrations of glucose and desmosterol were deregulated. This finding suggest that these pathways may be good candidates for further experimental investigation.

Differences in results of the E-Flux and mGX-FBA method. The flux through oxidative and non-oxidative phase of the pentose phosphate pathway was found to be increased for PWD mice in the mGX-FBA method, which is contrast of the E-Flux method. The degree of flux regulation of the pentose phosphate pathways is linked to oxidative stress and mice phenotypes when mGX-FBA is used, while it is not linked to mice phenotypes when E-Flux is used. For this pathway, mGX-FBA is better than E-Flux in prediction of mice phenotypes. However, it is difficult to judge the prediction of both methods because lack of experimental flux data.

In contrast of using mGX-FBA, the bile acid synthesis pathways is found to be deregulated at different degree by using E-Flux (Figure 3.16b). The concentration of taurocholic acid/tchola that is a downstream metabolite of the bile acid synthesis pathways can be expected to be upregulated for PWD mice due to an increase of flux through bile acid

synthesis pathway. Experimentally measured concentration of taurocholic acid was upregulated (4-fold) in PWD mice, which is inline with the expectation. However, A/J mice was not inline with the expectation as an increase (1.8 fold) of concentration of taurocholic acid was found while the flux through the bile acid synthesis pathways was downregulated.

The alteration of throughout fluxes of the GMM model using the E-Flux and mGX-FBA method. The aim of this work was to study changes in fluxes of metabolic networks in two conditions. To address this, gene expression data of the mouse strains A/J, C57Bl6 and PWD in DDC-treated as well as control conditions were integrated into the GMM model using the mGX-FBA and E-Flux methods and fluxes were obtained (section 3.3.1 and 3.3.2). To study the difference of throughout fluxes of the GMM model between DDC-treated and control conditions, the flux through each and every reactions of the model were compared using relative flux deviation ($|d_r|$) value (section 2.4.1). To know a marked difference in the flux value of a reaction r between two conditions a moderate cutoff ($|d_r| > 0.5$) was set.

Aforementioned cutoff was applied on the results of fluxes of the mGX-FBA and E-Flux methods and the obtained results of both methods were compared. In the mGX-FBA method out of 3727 reactions of the GMM model a marked difference (DDC-treated vs control) in the flux value were obtained for 517, 773 and 596 in mouse strains A/J, C57Bl6 and PWD, respectively, while 1089, 878 and 1077 were obtained for A/J, C57Bl6 and PWD in the E-Flux method (Figure 3.17a). This result suggest that E-Flux method alters fluxes through a higher number of reactions of the GMM compared to the mGX-FBA method. Out of 3727 reactions, 259, 255 and 266 were common in the both methods for A/J, C57Bl6 and PWD mice, respectively (Figure 3.17a).

3.3.4 Comparison of flux regulation with- and without-integration of gene-expression data

The aim of this study is to analyse whether the integration of gene expression data into the model (section 3.3.2 and 3.3.3) is required for understanding the flux regulation or it is sufficient to judge the flux regulation based on only regulation of gene expression data. To address this, the bile acid synthesis and cholesterol biosynthesis pathways are chosen due to the metabolic concentrations of these pathways were measured. Thus, these pathways can be used as a validation such as the regulation of the metabolic concentrations may be estimated from the regulation of fluxes.

The integration of gene expression data into the model has been already done in section 3.3.2 and 3.3.3. On the one hand, for an example of the integration of gene expression data the different degrees of flux regulation was observed for bile acid synthesis pathway among all three mouse strains (Figure 3.16b). The flux regulation through the bile acid synthesis pathway is inline with the regulation of the concentration of taurocholic acid for PWD mice, while it was not inline with the A/J (section 3.3.3).

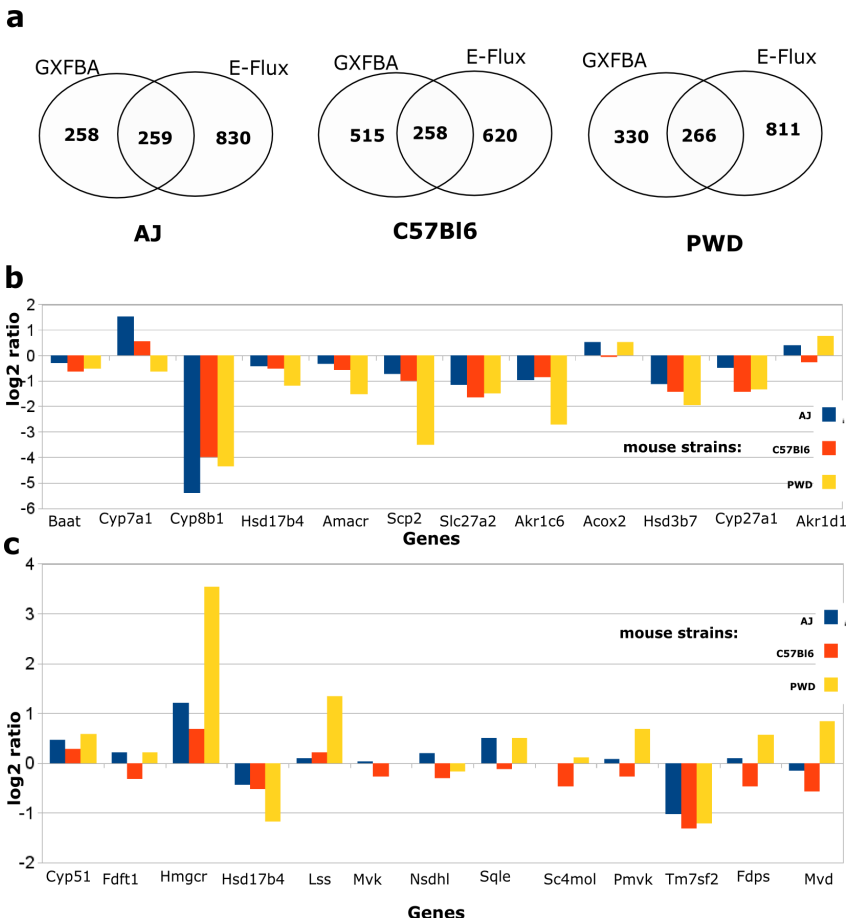

Figure 3.17. Comparison of E-Flux and mGX-FBA. a) Results of differences in reaction fluxes of the mGX-FBA and E-Flux method. **b)** Fold change (DDC-treated vs control) of genes of bile acid synthesis pathway**. c)** Fold change (DDC-treated vs control) of genes of cholesterol biosynthesis synthesis pathway.

On the other hand, the regulation of gene expression data of the bile acid synthesis pathway were observed in Figure 3.17b. Some genes of this pathway were upregulated while some were downregulated (Figure 3.17b). For example, *Cyp7a1* was found to be upregulated in A/J, while *Cyp8b1* was downregulated. In PWD, *Akr1d1* was upregulated and *Scp2* along with *Cyp81b* were downregulated. Thus, the judgement of regulation of the flux or the concentration of bile acid synthesis pathway based on only gene expression data is

problematic.

On the one hand, for another example of the integration of gene expression data the flux through cholesterol biosynthesis pathway was observed to be upregulated in C57Bl6 and PWD mice, while it was slightly downregulated in A/J (Figure 3.17b). The regulation of the concentration of desmosterol is inline with the regulation of the flux through this pathway for all mouse strains (section 3.3.2). On the other hand, gene expression data of the cholesterol biosynthesis were observed in Figure 3.17c. It is difficult to judge the regulation of the flux or the concentration of this pathway due to some genes of this pathway were upregulated and some genes were downregulated in all three mouse strains. Thus, integration of gene expression data into the metabolic model can provide a better understanding of flux regulation than the estimation of fluxes solely based on gene expression data.

Taken together, to understand different degrees of flux regulation among mouse strains I performed an integrated analysis of gene expression data of mouse liver samples of a metabolic network using the methods mGX-FBA and E-Flux. Both methods identified different degrees of flux regulation among mouse strains via glucose and cholesterol biosynthesis pathways. Interestingly, the measured experimental concentration of glucose and desmosterol are found to be deregulated at different degrees among mice strains due to DDC-treatment. E-Flux method shows overall flux alteration of a larger number of reactions of the GMM than the mGX-FBA method. The difference between DDC-treated *vs.* control of the flux through the biomass reaction is in line with mouse steatohepatitis-like phenotype (Figure 3.14). Furthermore, the judgement of flux regulation based only on gene expression is difficult (Figure 3.17b & c); thus aforementioned methods may be an option to explore *in silico* based flux regulation.

3.4 Identification of potential drug targets

There are only a few medicines that can cure, prevent or slow the progression of liver disease such as, steatosis, steatohepatitis and hepatocellular carcinoma (HCC). Thus, the identification of new potential drug target candidates is a very important task in bringing innovative new medicines to patients. Systems biology approaches are suggested as a promising tool for improving decision making in pharmaceutical development (Butcher et al., 2004). High throughput gene-expression and genome-scale models have been used for identifying potential drug targets in cancer (Folger et al., 2011). One aim of this work is to identify potential drug targets in liver disease such as steatosis, steatohepatitis and hepatocellular carcinoma (HCC) through metabolic networks. This aim may be accomplished by applying a constraint based modeling approach to metabolic networks. A similar approach for identification of drug targets in cancer was proposed by Folger et al. (2011).

I have applied a constraint based modeling approach to the GMM model to identify *in silico* drug targets in liver disease. In this study, a drug target is a gene which encodes an enzyme and the activity of this enzyme might be inhibited by drugs. *In silico* drug targets were identified for three different types of the liver diseases: steatosis, steatohepatitis, and HCC.

Steatosis can be caused by accumulation of lipids and subsequent formation of lipid droplets (Anderson and Borlak, 2008). For steatosis, I formulated a biomass reaction and the flux through this reaction was used as a readout of the disease (section 2.6).

Steatohepatitis may be caused by accumulation of lipids, inflammation and oxidative stress. A second biomass reaction was formulated for the steatohepatitis and the flux through the reaction was used as readout of the disease (section 2.6).

HCC is one of the most common malignant tumors and accounts for most liver cancers (Li, 2013). For readout of HCC phenotypes, a third biomass reaction was formulated (section 2.6).

Before using the biomass reaction for the identification of drug targets in steatosis and steatohepatitis, I tested whether the flux through this reaction can be used as a readout of disease phenotypes or not. To achieve this, I integrated gene expression data to the GMM model and compared the flux through the biomass reaction with the mouse phenotypes. The comparison was performed using gene expression data of the mouse strains A/J, C57Bl6 and PWD. In this comparison analysis, first of all, the flux through the biomass reactions for steatosis and steatohepatitis, was calculated using the combination of GMM model and expression data. To address this, the following steps were used:

1) The flux through the biomass reactions were maximized by using flux balance analysis and the obtained flux distributions (v^{Ref}) were represented as a reference state.

2) In this study, "reaction expression" is termed as the expression of an enzyme or a group of similar enzymes catalyzing a reaction that is reflected by the enzyme or enzyme-encoding genes. The reaction expression of each enzymatic reaction of the GMM model was computed either as maximum of the expressions of enzyme-encoding genes or as direct expression value if there is a single enzyme-encoding gene. Separately, the reaction expressions were computed for 3 biological replicates (BR) of A/J and C57Bl6 mice in control as well as in DDC-treated conditions, and for 3 BR of PWD mice in control along with 2 BR in DDC-treated conditions.

3) For each reaction of the GMM model, the coefficients of variation (CV) were computed and the computation was performed over the reaction expressions of the aforementioned biological replicates of the mouse strains A/J, C57BL6 and PWD in both control and DDC-treated conditions. A set of reactions, S^R, was identified in which each reaction has moderate variation (CV>0.65) of reaction expression over the biological replicates. To incorporate particular tissue specific data to the GMM model might be problematic. For example, assume that some genes have very low expression or are not expressed at all in a particular tissue. If the reactions corresponding to these low expressed genes are completely blocked in the GMM model, then it may affect some important functions of the GMM model. Thus, to avoid complete elimination of flux the reaction expression of each reaction of S^R is transformed between 0.2 and 1 by using the following transformation:

$$0.2+\frac{(1-0.2)\times(rexp_i-min(rexp))}{(max(rexp)-min(rexp))} \qquad (3.2)$$

where $rexp_i$ is reaction expression of reaction i of the set S^R, and *min(rexp)* and *max(rexp)* are minimum and maximum reaction expressions over all biological replicates, respectively.

4) Using different levels of gene expression data among mice can change the bound constraints such as lower and upper bounds of reaction fluxes of the GMM model. Using flux balance analysis, to obtain different degrees of flux regulation among mouse strains I have applied gene expression of mouse strains for constraining the reaction of the GMM model. For each reaction r of the S^R, expression was transformed using formula (3.2). Then transformed expression (*texp*) was used to set $LB(r)=v^{Ref}(r)\times texp$ of *r* if $v^{Ref}(r)$ is negative and $UB(r)=v^{Ref}(r)\times texp$ of *r* if $v^{Ref}(r)$ is positive, where *LB* and *UB* are lower and upper bound, respectively.

5) The biomass reaction was maximized using gene expression data of aforementioned biological replicates of A/J, C57Bl6 and PWD mice of control and DDC-treated conditions. The flux through the biomass reaction was normalized by dividing the flux through the biomass reaction that was obtained by step1. This normalized score

is called "objective score" and the reference state has an objective score equal to 1.

For steatosis and steatohepatitis, the objective score of the control mouse strains were lower than the DDC-treated mouse strains (Figure 3.18a). The objective score of steatosis readout in DDC-treated condition was found to be higher in PWD mice compared to A/J and C57Bl6 mice (Figure 3.18b). The objective score for steatosis is inline with the steatosis phenotypes of PWD mice (section 3.1.1). For a readout of steatohepatitis in DDC-treated condition, the objective score in A/J mice was slightly higher compared to PWD mice (Figure 3.18b; section 3.1.1). The objective score for steatohepatitis is inline with the steatohepatitis-like phenotypes of A/J.

3.4.1 Selective drug target

The number of liver drugs that are currently available is not sufficient (Muriel and Rivera-Espinoza, 2008). Thus, the identification of a range of safe and efficient therapeutic agents as well as new drug targets is a very important task. Therefore, an analysis was performed to identify novel potential drug targets in liver disease applying constraint-based modeling approach to the GMM model. To measure a readout of each of steatosis, steatohepatitis and HCC separate biomass reactions were constructed (section 2.6). For each of steatosis, steatohepatitis and HCC, I performed an *in silico* drug targets analysis which proceeded through the following steps:

1) The flux through each objective reaction of the GMM model is maximized and the steady state flux is obtained for each reaction of the GMM model.
2) The flux through the ADP/ATP transporter reaction/ATPtm of the GMM model is maximized to obtain the maximum flux through ATP maintenance (ATPtm).
3) The knockout of each gene of the GMM model was performed by reducing the flux through reactions associated with the gene. The flux through the reactions was reduced by 90 % of the obtained flux in step 1.
4) For each gene G of the GMM model, a score is calculated by the formula:

$$\frac{RF(ATPtm)}{OF(ATPtm)} \times \left(1 - \frac{RF(BR)}{OF(BR)}\right) \tag{3.3}$$

where, $RF(ATPtm)$ and $RF(BR)$ represent the flux through the ATPtm and the biomass reaction after reducing 90% flux through the reactions associated with G, $OF(ATPtm)$ and OF(BR) represent original flux (step 1). Score 1 and 0 represent good and bad selective drug targets. After gene knock-down from the GMM model, if the flux through the reaction ATPtm is reduced markedly then the gene is not consider as a good drug target. Therefore, the term RF(ATPtm)/OF(ATPm) is used in the score calculation.

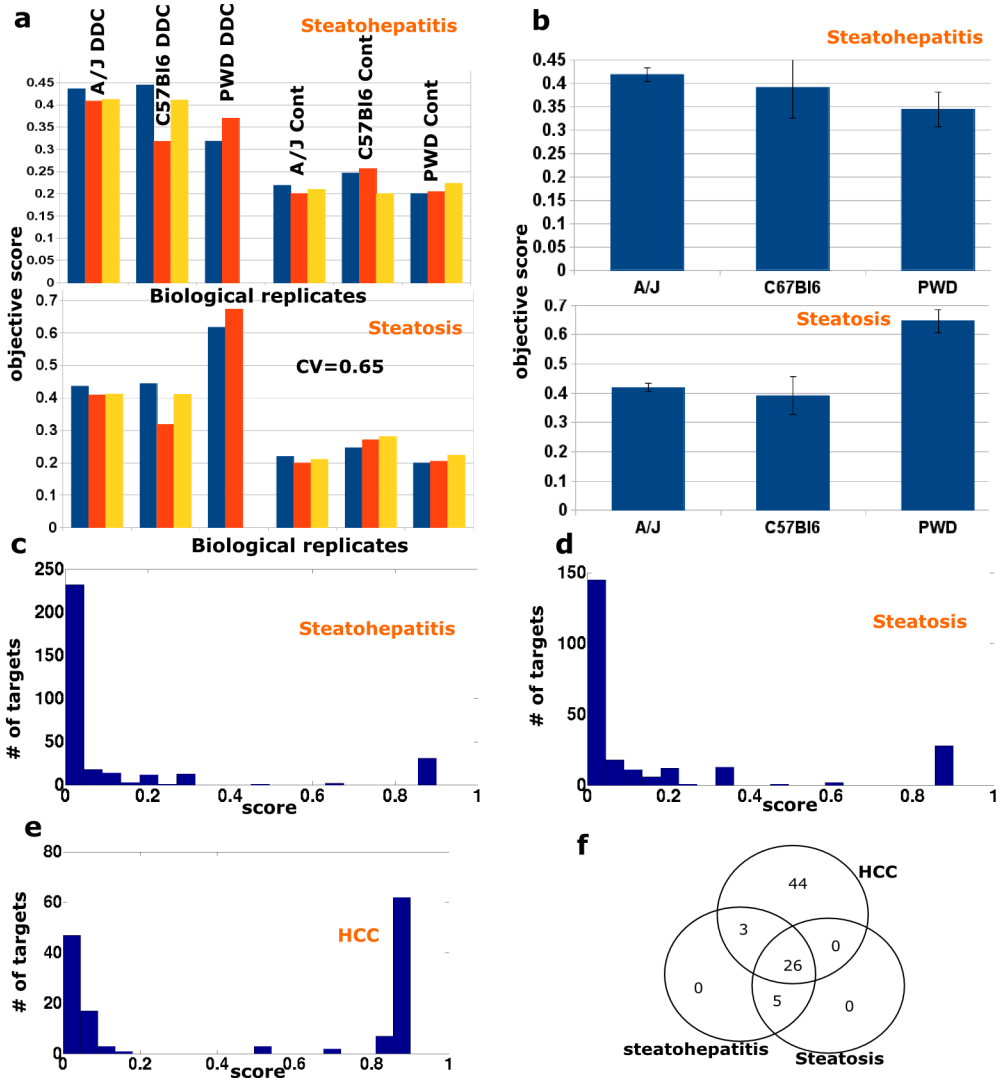

Figure 3.18. **Identification of drug targets. a)** Objective score used as a readout of steatohepatitis and steatosis. Upper pannel and lower pannel represents a readout of the steatohepatitis and the steatosis, respectively. **b)** Upper pannel and lower pannel represents the average (over biological replicates) of objective score of the steatohepatitis and the steatosis in DDC-treated mice. PWD mice show high degree of steatosis and A/J mice show high degree of steatohepatitis. **c)** Frequency of drug targets in steatohepatitis. High score indicates potential drug target. **d)** Histogram of drug targets in the steatosis **e)** Histogram of drug targets in the HCC. **f)** Venn diagram of drug targets of the steatosis, the steatohepatitis and the HCC.

For each of three liver diseases steatohepatitis, steatosis and HCC, aforementioned steps were used to calculate the score by knocking out one by one gene of the GMM model. After knocking out a gene of the GMM model, if the score was greater than 0.4 then the gene was

identified as an *in silico* drug target. Histograms of the scores are depicted for steatohepatitis (Figure 3.18c), steatosis (Figure 3.18d) and HCC (Figure 3.18e). The number of drug targets in aforementioned three disease is shown in a Venn diagram (Figure 3.18f and Table 3.1). In total 78 drug targets were obtained for aforementioned three diseases, 26 were common in all three, 5 were common in steatohepatitis and HCC, and 3 were common in steatohepatitis and steatosis. Out of 78, 17 are annotated as a drug targets in DrugBank (Knox et al., 2011) in which 7 are FDA-approved anticancer drugs (Table 3.2) that may be used in the treatment of HCC, and the remaining 61 may act as potential drug targets in liver disease. Furthermore, an analysis was performed to identify metabolic pathways corresponding to the drug targets of the aforementioned three liver diseases. The analysis identified as the top two ranked pathways, cholesterol and triacylglycerol/triglyceride metabolism that comprised 13 and 9 drug targets, respectively as well as other pathways which are shown in Table 3.3.

In the aforementioned comparison analysis of potential drug targets in liver disease, 26 drug targets were common in the steatosis, the steatohepatitis, and the HCC. From 26 drug targets, 13, 6 and 3 were identified from cholesterol metabolism, triacylglycerol synthesis and inositol phosphate metabolism, respectively. From 13 identified drug targets related to cholesterol metabolism were enriched for terpenoid backbone biosynthesis and steroid synthesis of KEGG pathways (Figure 3.19 & 3.20). Cholesterol biosynthesis begins with acetyl-CoA that is converted to 3-Hydroxy-3-methyl-glutaryl-CoA by 3-hydroxy-3-methylglutaryl-CoA synthase 2 (Hmgc2) (Figure 3.19). 3-Hydroxy-3-methyl-glutaryl-CoA is converted to mevalonate by 3-hydroxy-3-methylglutaryl-CoA reductase (Hmgcr) and subsequently mevalonate undergoes a series of steps and finally is converted to farnesyl diphosphate that is a precursor of steroid biosynthesis. In steroid biosynthesis, farnesyl diphosphate is converted to squalene by farnesyl diphosphate farnesyl transferase 1 (Fdft1) (Figure 3.20). Squalene is further converted to cholesterol in a series of biochemical reactions (Figure 3.20).

Hmgcr, a key enzyme in cholesterol biosynthesis, is identified as a drug target in my *in silico* drug targets analysis. Hmgcr has been already used as a target of a class of statins (e.g. atorvastatin, pravastatin, and rosuvastatin) drugs which are used to lower cholesterol levels by inhibiting the Hmgcr (Ekstedt et al., 2007). Clinical studies reported positive results regarding the use of statins in the treatment of NASH (Anderson and Borlak, 2008). In addition to Hmgcr from cholesterol biosynthesis, several other potential *in silico* drug targets were found and are highlighted in red in Figure 3.19 and 3.20. These drug targets may one day be used as drug targets for steatosis and the steatohepatitis. However, I did not find drugs for these targets after a review of the literature. These targets seem to be novel and experimental investigation would be needed for validation.

Table 3.1. Identified drug targets for steatohepatitis, steatosis and HCC. A set comparison analysis of drug targets is shown.

Only in HCC	**Only in HCC**	**Common in steatosis, steatohepatitsh and HCC**	**(Steaohepatitis and HCC-Steatosis)**	**(Steatosis and Steatohepatitis -HCC)**
Atic	Adssl1	Fdps	Kdsr	Dgat1
Psph	Ahcy	Isyna1	Sptlc2	Acsl1
Slc7a3	Cth	Cyp51	Sptlc1	Dgat2
Psat1	Gart	Pmvk		Acaca
Tyms	Pfas	Fdft1		Acacb
Ppat	Cad	Tm7sf2		
Pah	Ahcyl1	Atp10a		
Prps2	Umps	Atp8a1		
Slc6a14	Dhfr	Agpat6		
Slc38a4	Mat2b	Agpat5		
Slc38a1	Slc7a2	Agpat4		
Slc38a2	Mat1a	Agpat3		
Dhodh	Impdh2	Agpat2		
Pcbd1	Slc7a1	Agpat1		
Impdh1	Cbs	Gpam		
Paics	Glul	Idi1		
Adss	Adsl	Mvd		
Sgms1	Txnrd1	Hmgcr		
Gmps		Sc4mol		
Lgsn		Mvk		
Prps1		Hsd17b4		
Asns		Cdipt		
Phgdh		Impa1		
Pgs1		Impa2		
Crls1		Lss		
Mat2a		Sqle		

Table 3.2. Drug targets and approved drugs. Genes in the table may be used as drug targets in the HCC. These genes are targeted with some FDA approved anticancer drugs.

Drug target gene	FDA-approved drugs
Acaca	L-Glutamic Acid
Acacb	Alitretinoin
Agpat3	Candesartan
Agpat1	NADH
Adsl	Arginine
Atic	NADH
Ahcy	NADH

Acetyl-CoA carboxylase (*Acaca* and *Acacb*) was found to be an *in silico* potential drug target in the steatohepatitis and steatosis. Acetyl-CoA carboxylase catalyzes the reaction that provides malonyl-CoA from Acetyl-CoA. Malonyl-CoA is a precursor of biosynthesis of fatty acid. Metformin, is a widely used drug in the treatment of type 2 diabetes, and suppresses hepatic steatosis (Zhou et al., 2001). Metformin activates AMP-activated protein kinase (AMPK) in hepatocytes that suppresses the activity of acetyl-CoA carboxylase activity (Zhou et al., 2001). My *in silico* study of drug targets Acaca and Acacb is inline with the previous study of Zhou et al., 2001.

In silico drug targets of glycerolipid metabolism are shown in Figure 3.21. Glycerol-3-phosphate acyltransferase that is encoded by gene Agpat6, is involved in the first committed step of triacylglycerol biosynthesis (Figure 3.21). Liver-directed overexpression of mitochondrial glycerol-3-phosphate acyltransferase results in hepatic steatosis and increased triacylglycerol secretion (Lindén et al., 2006). Furthermore, pharmacological glycerol-3-phosphate acyltransferase inhibition decreases lipid accumulation in white adipose, brown adipose, and liver tissues (Kuhajda et al., 2011). Glycerol-3-phosphate acyltransferase (Agpat6) may be act as a drug target for steatosis based on the study of Kuhajda et al., 2011 and also Agpat6 is found as an *in silico* drug target in this analysis.

Table 3.3. Involved drug targets in metabolic pathways.

Pathways	**Drug targets for liver diesease**
Cholesterol Metabolism	Cyp51, Fdft1, Hmgcr, Hsd17b4, Lss, Mvk, Sqle, Sc4mol, Pmvk, Tm7sf2, Fdps, Mvd, Idi1
Triacylglycerol Synthesis	Dgat1, Gpam, Agpat3, Agpat5, Agpat1, Agpat2, Dgat2, Agpat4, Agpat6
Transport, Extracellular	Atp8a1, Slc7a1, Slc7a2, Slc7a3, Slc6a14, Slc38a2, Slc38a4, Slc38a1
Nucleotides	Adsl, Adssl1, Adss, Tyms, Impdh1, Impdh2, Txnrd1, Gmps
Methionine Metabolism	Mat1a, Cbs, Mat2b, Ahcyl1, Mat2a, Ahcy
IMP Biosynthesis	Gart, Paics, Atic, Ppat, Pfas
Sphingolipid Metabolism	Sptlc2, Kdsr, Sgms1, Sptlc1
Inositol Phosphate Metabolism	Impa1, Isyna1, Impa2
Glycerophospholipid Metabolism	Cdipt, Crls1, Pgs1
Glycine, Serine, and Threonine Metabolism	Psph, Psat1, Phgdh
Pyrimidine Biosynthesis	Umps, Dhodh, Cad
Glutamate metabolism	Glul, Lgsn
Fatty Acid Metabolism	Acacb, Acaca
Pentose Phosphate Pathway	Prps1, Prps2
Tyr, Phe, Trp Biosynthesis	Pah
Alanine and Aspartate Metabolism	Asns
Transport, Mitochondrial	Atp10a
Folate Metabolism	Dhfr
Fatty acid activation	Acsl1
Cysteine Metabolism	Cth
Tetrahydrobiopterin	Pcbd1

Isyna1, is a key enzyme, which is involved in the myo-inositol biosynthesis pathway (Figure 3.22). Isyna1 catalyzes the rate-limiting conversion of glucose 6-phosphate to myoinositol 1-phosphate (Figure 3.22). Carbamazepine may inhibit myo-inositol-1-phosphate synthase (Isyna1) and reduce inositol (Lin et al., 2013). Carbamazepine decreases the hepatic load of mutant α1-antitrypsin proteins and hepatic fibrosis in mouse model (Hidvegi et al., 2010). Isyna1 is found as a potential *in silico* drug target. Taken together, carbamazepine might be a novel drug candidate to be tested for the treatment of NAFLD which reduces the activity of Isyna1, but of course it would require experimental validation.

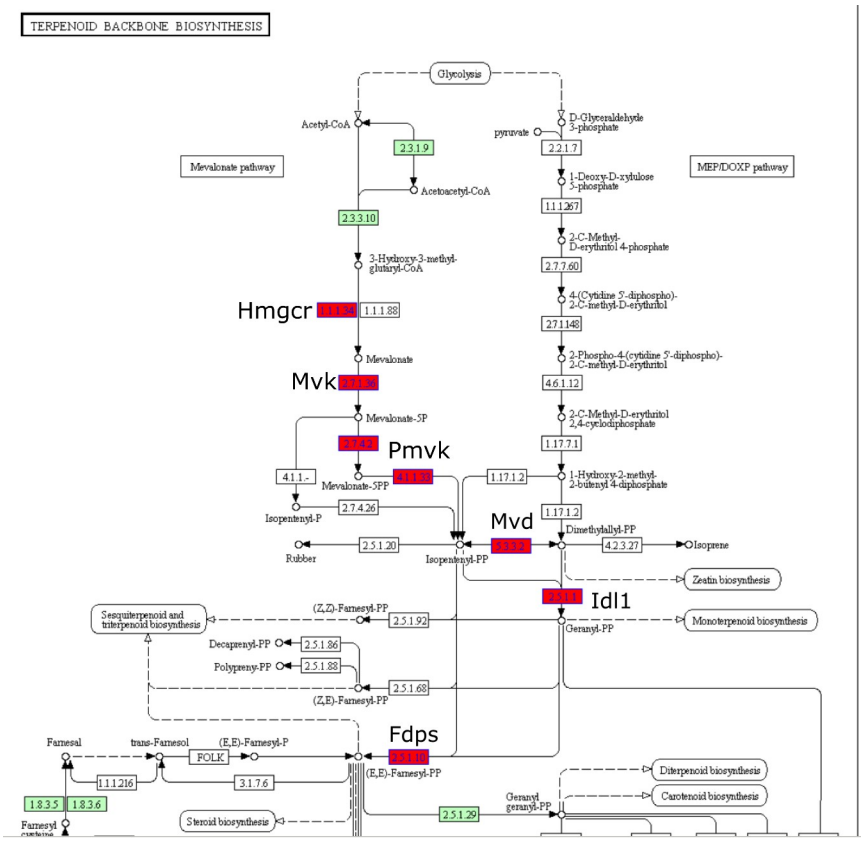

Figure 3.19.Terpenoid backbone biosynthesis.

Drug targets are shown in red color. Acetyl CoA is converted to farnesyl diphosphate that is a precursor of cholesterol biosynthesis. Hmgcr is involved in a rate-limiting step of cholesterol biosynthesis.

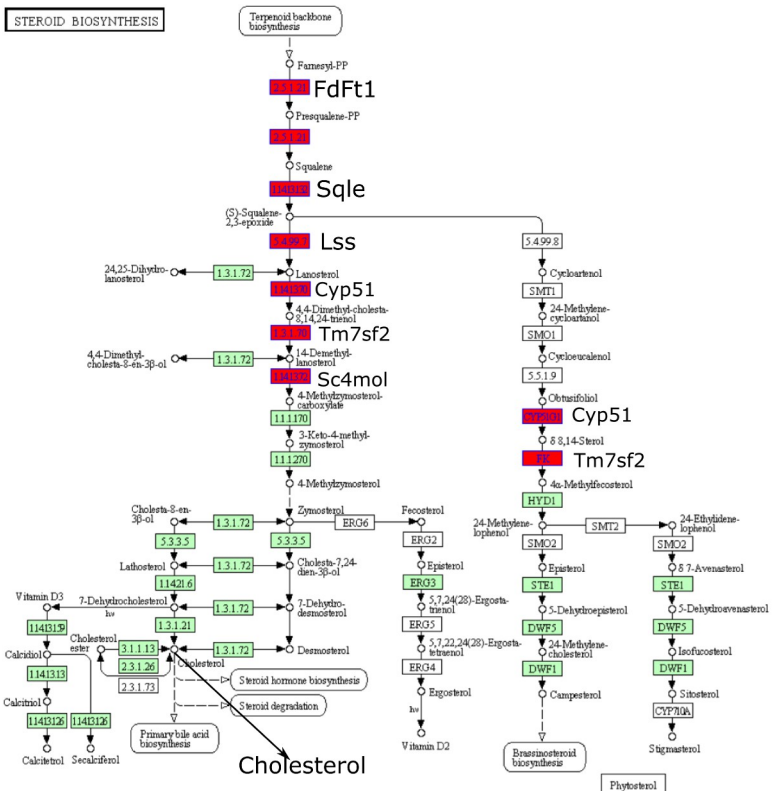

Figure 3.20. Steroid biosynthesis.

Farnesyl diphosphate is converted to cholesterol via a series of biochemical reactions. Red nodes indicate drug targets that are identified in *in silico* analysis.

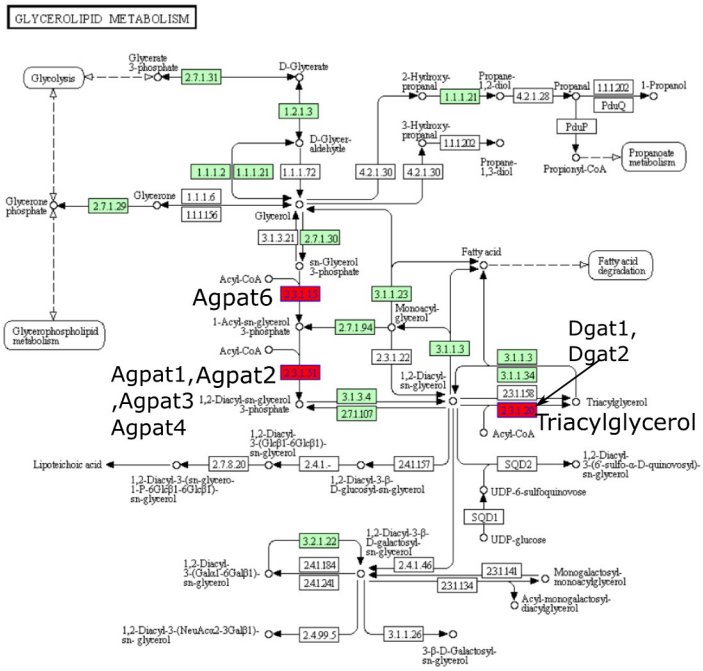

Fig 3.21. Triacylglycerol biosynthesis.

Triacylglycerol biosynthesis pathway begins with acyl-CoA and which then is converted to triacylglycerol through the series of biochemical reactions. Agpat6 catalyzes a rate-limiting step of triacylglycerol biosynthesis.

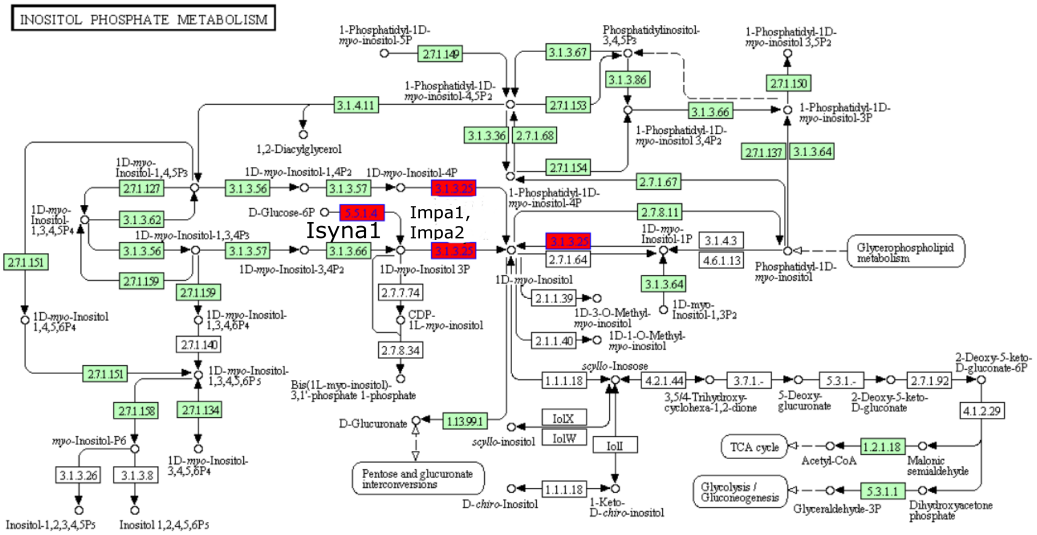

Fig 3.22. Inositol phosphate metabolism.

In this pathway, D-glucose-6 phosphate that is a product of glycolysis, is converted to myo-inositol. Isyna1 is involved in a rate-limiting step of myo-inositol biosynthesis. The identified possible candidates for drug targets are shown in red color.

In summary, I formulated biomass reactions for steatohepatitis and steatosis. The flux through the biomass reactions was able to mimic the steatosis and steatohepatitis phenotypes of mouse strains after integrating gene expression data of mouse strains into the GMM (Figure 3.18b). Using these biomass reactions as well as a biomass reaction of HCC I identified 78 potential candidates for drugs in liver diseases: steatohepatitis, steatosis and HCC, in which 7 candidates are already found to be FDA-approved anticancer drugs. 26 potential drug candidates are common in aforementioned liver diseases and reside in metabolism of cholesterol and inositol phosphate as well as triacylglycerol synthesis. *Hmgcr* was identified as a potential drug candidate in my in silico study and also was used as a target of statins drugs. In addition, based on a review of literature I propose some novel potential drug candidates, such as, *Agpat6*, *Isyna1, Acaca* and *Acacb,* but of course it would require experimental validations.

3.4.2 Reduction of flux through reactions associated with SREBP target genes

Sterol regulatory element binding proteins (SREBPs) are transcription factors that regulate cholesterol, fatty acid, triacylglycerol and phospholipid synthesis (Eberlé et al., 2004). There are three isoforms of SREBP, SREBP-1a, SREBP-1c and SREBP2 in mouse. SREBP-1c has been shown to play a key role in hepatic steatosis (Kim et al., 2010). The SREBP activation is essential for development of diabetic fatty liver (Moon et al., 2012). Experimentally measured protein level of SREBP-1 and SREBP-2 were found to be upregulated (1.70, 1.31)-fold, (1.51, 1.34)-fold and (1.84, 1.6)-fold for the A/J, C57Bl6 and PWD mice, respectively. The upregulation of SREBP proteins may lend steatosis and it leads to formation of lipid droplets.

The goal here is to find drug targets for steatosis by knocking out the target genes of SREBP. To address this, first of all, the target genes of SREBP were obtained from Reactome (Croft et al., 2011). Secondly, a biomass reaction of steatosis was introduced in the GMM model (section 2.6) and the flux through the biomass reaction was used as a readout of the formation of lipid droplets. The flux through the objective reaction was maximized and the obtained flux distribution of the GMM model is called by v^{opt}. Furthermore, an *in silico* knock out simulation for each target gene of SREBP was done by imposing a constraint on the set of target gene associated reactions. If r is a reaction in set of target gene associated reactions then the upper and the lower bound of the reaction r is equal to $v^{opt}(r)\times 0.5$ if $v^{opt}(r)$ is positive and $-v^{opt}(r)\times 0.5$ if $v^{opt}(r)$ is negative. Rates of reduction of lipid droplets after knockout gene targets, are shown in Figure 3.23.

To observe whether the expression patterns of SREBP target genes are changed due to DDC-treatment or not. Ratios of expression data between DDC-treated *vs.* control of target genes are listed in Table 3.4. In my *in silico* analysis, the genes *Cyp51*, *Hmgcr*, *Lss*, *Sqle*, *Pmvk*, *Fdps*, *Mvd* have higher upregulation in PWD than in C57Bl6 mice (Table 3.3) and knockout of these genes reduces the formation of lipid droplets (Fig 3.16). Thus, these genes may become drug targets of steatosis. Histological improvement of NAFLD and the reduction in the hepatic steatosis are found using the treatment with statins which inhibit an enzyme HMG-CoA reductase (Hmgcr) of cholesterol biosynthesis (Ekstedt et al., 2007; Ji et al., 2011; Onofrei et al., 2008; Rallidis et al., 2004).

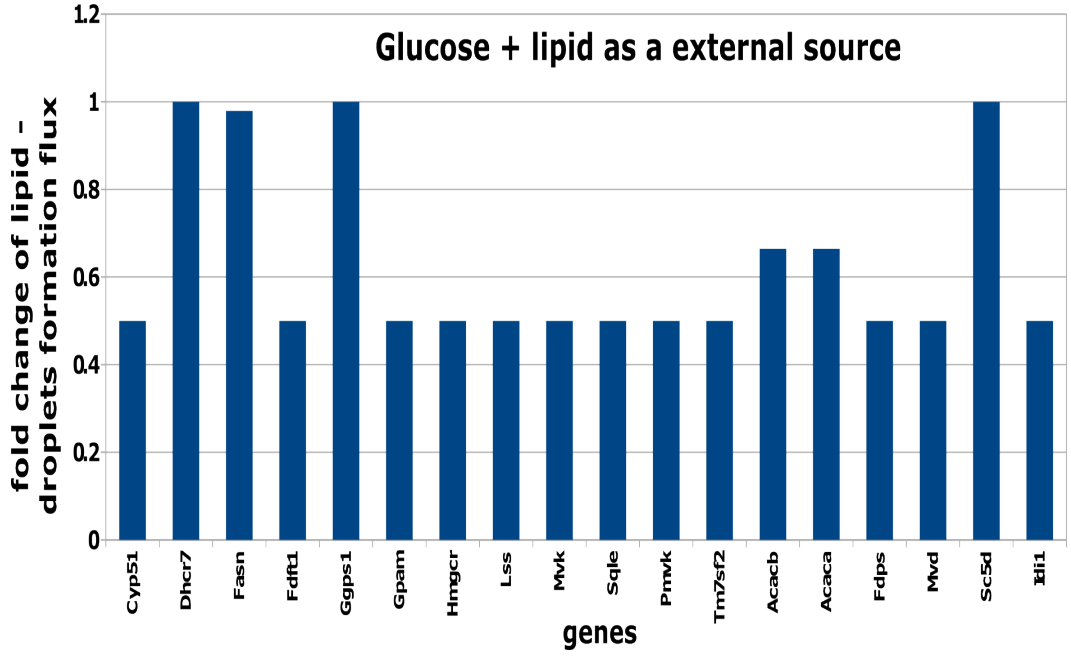

Figure 3.23. **Knock out of SREBP targets.**

Glucose and lipid are used as external source of the GMM model for the formation of lipid droplets. Each SREBP target gene was knocked-out and the change (before vs after knockout) of flux through the reaction of lipid droplets are shown.

Table 3.4. Target genes of SREBP. Ratios between DDC-treated vs control of expression data are calculated.

Genes	A/J ratio	C57Bl6 ratio	PWD ratio
Cyp51	1.38	1.21	1.51
Dhcr7	1.18	0.83	0.65
Fasn	0.82	0.43	0.63
Fdft1	1.16	0.8	1.16
Ggps1	1.53	1.08	1.18
Gpam	0.59	0.7	0.37
Hmgcr	2.31	1.6	11.6
Lss	1.07	1.15	2.51
Mvk	1.02	0.83	1
Sqle	1.42	0.92	1.4
Pmvk	1.05	0.83	1.61
Tm7sf2	0.49	0.4	0.43
Acacb	0.78	0.74	0.34
Acaca	1.87	1.4	0.99
Fdps	1.07	0.72	1.48
Mvd	0.9	0.67	1.78
Sc5d	0.45	0.71	0.36
Idi1	0.96	0.84	1.8

3.5 Qualitative and quantitative modeling

Based on over representation analysis of gene expression and metabolic profiles the arachidonic acid (AA) metabolic pathway was found to be deregulated for DDC-treated mice (section 3.1.1). Metabolites can be seen as end points of perturbations occurring at the gene level and the gene expression data might explain the metabolite data (Hoppe, 2012). Hence, to predict the change in metabolic concentrations gene expression data can be used to integrate into the metabolic models for *in silico* studies. Change in metabolite concentrations can be done; for example, how many metabolites are up- or downregulated under two conditions. To address this, two modeling approaches Monte Carlo ODE- and Petri net-based can be used (Ruths et al., 2008; Wierling et al., 2012). In Monte Carlo ODE-based one can integrate quantitative data (e.g., gene expression data), but cannot predict quantitatively. Hence, the prediction is termed in this study as "semi-quantitative". An alternative to Monte Carlo ODE-based, a Petri net-based method proposed by Ruths et al., (2008) can be used for a qualitative analysis such as, how many predictions are in agreement or disagreement. One can compare the results of the simulation of Monte Carlo ODE- and Petri net-based methods. The aim of this study is to understand the perturbation of metabolites of AA pathway that is identified as a perturbed candidate due to DDC-treatment. Both simulation methods were used and the results were compared with the experimental data.

There are only a few kinetic models of the AA metabolism available in the literature. For instance, Yang et al. (2007) proposed a kinetic model of AA metabolism in human. However, this model does not include some important metabolites, such as leukotriene D4, leukotriene C4, prostaglandin I2 and prostaglandin F2α (Yang et al., 2007). Another example is the proposed model of Gupta et al., (2009) in which leukotriene D4, leukotriene C4 are not included. These proposed models are not large enough for proper description of the experimental data. Hence, we compiled for this study a more comprehensive *in silico* model of the AA metabolism based on information from KEGG (Kanehisa and Goto, 2000), a biochemical text book (Michal and Schomburg, 2013) and the thesis of Schulz, (2012) (Figure 3.24).

3.5.1 Semi-quantitative analysis using Monte Carlo ODE-based approach

An appropriate distribution for the kinetic parameters was derived from 871 kcat values of mouse enzymes extracted from BRENDA (Scheer et al., 2011). The kcat values follow a log-normal distribution with mean 0.35 and standard deviation (SD) 3.5 (Figure 3.25). Here, I applied a Monte Carlo-based approach to study the biochemical network of AA metabolism, in which reactions are modeled by mass action kinetics and kinetic parameters were sampled from the aforementioned log-normal distribution (mean=0.35; SD=3.5; Wierling et al., 2012). This allows to study the biological system using a broader parameter space instead of using parameters, which were measured only under specific, e.g., *in vitro* conditions.

Enzyme amounts are assumed to depend linearly on the respective RNA-Seq-based gene expression values that were used to model enzyme concentrations.

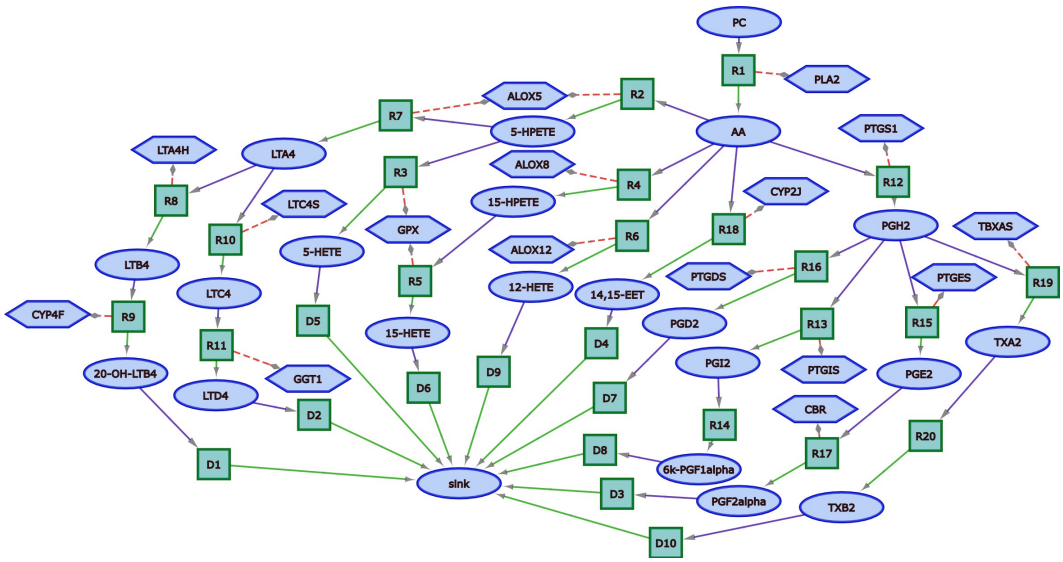

Figure 3.24. Model of arachidonic acid metabolism. Reactions, metabolites and enzymes are represented with rectangular, ellipse and hexagonal boxes, respectively. Dotted lines link enzymes with their respective reactions and solid lines indicate the mass flow through the system. Full names of components of the arachidonic acid metabolism model are abbreviated according to the following notations: prostaglandin I2, PGI2; leukotriene A4, LTA4; 5(S)-Hydroperoxy-6-trans-8,11,14-cis-eicosatetraenoic acid, 5-HPETE; 5-Hydroxyeicosatetraenoate, 5-HETE; leukotriene B4, LTB4; Thromboxane B2, TXB2; 15(S)-Hydroperoxy-6-trans-8,11,14-cis-eicosatetraenoic acid, 15-HPETE; 14,15-EET, 14,15-EET; 12-Hydroxyeicosatetraenoate, 15-HETE; 12-Hydroxyeicosatetraenoate, 12-HETE; 6-Keto-prostaglandin F1alpha, 6k-PGF1alpha; Prostaglandin E2, PGE2; 20-OH-Leukotriene B4, 20-OH-LTB4; Prostaglandin F2alpha, PGF2alpha; leukotriene C4, LTC4; arachidonic acid, AA; prostaglandin D2, PGD2; prostaglandin H2, PGH2; thromboxane A2, TXA2; leukotriene D4, LTD4; leukotriene A4 hydrolase, LTA4H; glutathione peroxidase 3, GPX; arachidonate 5-lipoxygenase, ALOX5; phosphatidylcholine, PC; phospholipase A2, PLA2; prostaglandin I2 (prostacyclin) synthase, PTGIS; leukotriene C4 synthase, LTC4S; arachidonate 12-lipoxygenase, ALOX12; sink arachidonic_acid, Sink_AA; leukotriene-B4 20-monooxygenase, CYP4F; cytochrome P450, CYP2J; thromboxane A synthase 1, TBXAS; prostaglandin D2 synthase, PTGDS; prostaglandin-endoperoxide synthase 1, PTGS1; prostaglandin E synthase , PTGES; arachidonate 15-lipoxygenase, ALOX15; geranylgeranyltransferase type I, CGT1; carbonyl reductase 3 , CBR;

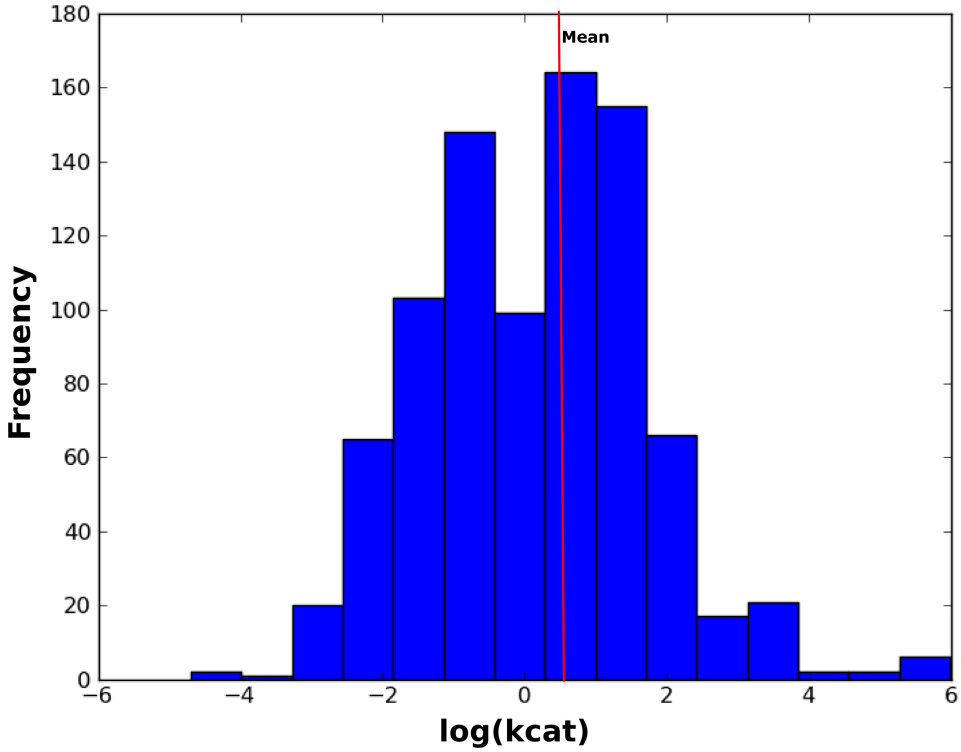

Figure 3.25. Distribution of kcat values. Distribution of mouse-specific kcat values from BRENDA database following a log-normal distribution (mean 0.35, standard deviation 3.5).

The AA release is catalyzed by phospholipase A2 (PLA2). In endothelial cell-pericyte co-cultures of rat cells it was shown that PLA2 is induced through the activation of the MAPK/ERK cascade (Anfuso et al., 2007). PLA2 is involved in inflammation driven liver fibrosis as a key feature of progressive NASH and PLA2 deficiency prevented activation of hepatic stellate cells and infiltration of F4/80-positive macrophages (Ishihara et al., 2012). Within our experiments the protein amount of phosphorylated ERK (pERK) was measured for the control and the DDC-treated mice using the reverse phase protein array (RPPA) technology. The ratios of pERK from DDC-treated vs. control samples were 1.90, 2.09 and 1.98 for A/J, C57Bl6 and PWD mouse liver samples, respectively. The effect of pERK on PLA2 activity was taken into account by multiplying the RNA-Seq expression value of PLA2 for the DDC-treatment simulation with the pERK ratio. I chose the gene whose expression data have lowest p-value (DDC-treated vs control) if the enzymes of a reaction is encoded by more than one gene.

An example of mapping the RNA-Seq expression to enzyme concentration is described below. Assume, a mass action kinetic equation,

$$v = k_{cat} \times E \times S \tag{3.4}$$

where v is the flux, k_{cat} is the enzymatic turnover number, E is the enzyme concentration, and S is the substrate concentration. Tissue specific RNA-Seq expression of gene are available and it can be used as a rough estimation for enzyme concentration. Assumption is that RNA-Seq expressions are linearly dependent on enzyme concentrations. Thus, the aforementioned equation is transformed in equation (Stavrum et al., 2013),

$$v = k_{cat} \times u \times rseq \times S \tag{3.5}$$

where *rseq* is the experimentally measured RNA-Seq expression value of gene that encode enzyme (unitless), and u is a unit used to convert RNA-Seq expression to enzyme concentrations. Unit u must be determined to obtain meaningful absolute fluxes and metabolite steady state concentrations of the biological model. However, the relative change of steady state concentrations of substrates and products of enzymatic reactions is independent of u (Stavrum et al., 2013). I did a qualitative analysis of relative change of steady state metabolite concentration in the AA model, and I set an arbitrary value 1nM to u.

Initial concentrations of enzymes of the AA model are estimated based on RNA-Seq expression data for three mouse strains A/J, C57Bl6 and PWD in control as well as DDC-treated conditions (Table 3.5). Initial concentrations of every metabolites of the AA model is set to 1nM. Mass action kinetics is used for every reaction in the AA model.

Table 3.5. Initial concentration of enzymes of AA model. Enzymes are approximated by RNA-seq expression of genes (RPKM values).

Genes	AJ control	AJ DDC	B6 control	B6 DDC	PWD control	PWD DDC
Lta4h	6.29	8.64	10.27	11.96	11.08	6.91
Gpx3	8.47	55.18	8.38	98.03	8.02	155.78
Alox5	0.08	0.50	0.22	1.04	0.06	0.24
Pla2g12a	8.25	20.87	8.43	35.14	7.14	22.99
Ptgis	1.36	3.57	1.08	9.01	0.55	0.78
Ltc4s	1.24	2.03	4.40	5.16	0.77	2.35
Alox12	0.04	0.16	0.03	0.51	0.04	0.30
Cyp4f18	0.21	6.60	0.67	9.22	0.41	6.62
Cyp2b10	0.46	55.49	0.64	45.23	0.01	83.76
Tbxas1	0.84	4.84	1.00	5.65	0.23	2.99
Hpgds	0.19	1.71	0.06	1.93	0.01	0.26
Ptgs1	3.78	8.04	4.95	6.99	1.02	2.15
Ptges3	4.71	4.95	12.19	5.71	10.84	8.11
Alox8	0.17	0.03	0.03	0.25	0.70	0.06
Ggt1	0.10	2.15	0.02	0.77	0.01	1.13
Cbr3	0.73	61.73	0.71	47.95	0.16	0.80

3.5.2 Qualitative analysis using Petri net-based simulator

One goal of this analysis is to understand the perturbation in metabolic concentrations of the AA pathway that can occur due to change in gene expression data. To address this, an alternative to Monte Carlo ODE-based method, a Petri net-based method was used to know which method provides a better prediction. The signaling Petri net-based simulator (SPN) is a simulator that uses Petri net modeling paradigm (Ruths et al., 2008) (section 2.8). The SPN simulation is used to characterize the dynamics of signaling networks. Here, I modified the firing scheme of the SPN to understand the metabolic pathways. An original python source code (http://www2.informatik.hu-berlin.de/top/pnk/#python) of Petri net simulator was modified to implement the Petri net-based simulator of metabolic pathways and this simulator is abbreviated as MPN in this study. The MPN was used to simulate the AA model. Furthermore, the prediction of changes in metabolic concentration was tested using MPN for the AA model.

Petri net representation of an enzymatic reaction (substrate, AA; product, 5-HPETE; and enzyme, Alox5) of the AA model is illustrated in Figure 3.26. This reaction is represented as the transition t. Assume that before execution/fire of the transition t, the places, AA, 5-HPETE and Alox5 contains 20, 15 and 15 tokens, respectively. When t fires tokens of AA can reduce, tokens of 5-HPETE can increase, and tokens of Alox5 remain unchanged.

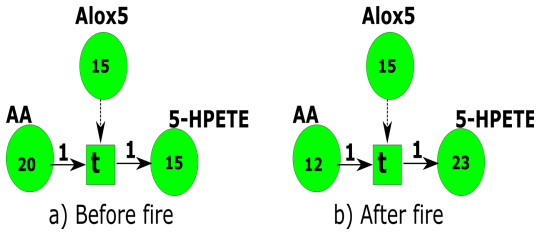

Figure 3.26. A Petri net model of a reaction (substrate AA; product 5-HPETE; and enzyme Alox5). Circular and rectangular node represents place and transition. Marking state **a)** before fire of the transition t **b)** after fire of t.

In the AA model, each place p may act as a substrate, a product, or an enzyme . Stoichiometric coefficient of substrates and products of a reaction are assigned as a weights to the edges from a substrate to a transition and from a transition to a product, respectively. In MPN, when a transition t executes/fires the tokens of places are updated according to the following steps:

Step 1: For each substrate, *s*, of t compute $\lfloor \frac{m(s)}{c} \rfloor$ and put in the set SS where, $m(s)$ and c are the number of tokens and stoichiometric coefficient of *s*, respectively, and SS is the set which contains number of tokens as an element. In case of Fig 3.26 a, AA is the only substrate of t and the execution of step 1 returned 20 tokens ($\lfloor \frac{20}{1} \rfloor = 20$). The set SS contains 20 tokens as an element.

Step 2: Calculate $speed = random(0, min(min(SS), m(e)))$ where, $m(e)$ is number of tokens of the enzyme *e* and $min(SS)$ is the minimum of the set SS. In case of Figure 3.26 the transition t has an enzyme Alox5 and a substrate AA and the evaluation $min(min(20),15)$ results in 15 tokens. Now, 15 tokens are used to execute the aforementioned formula of the speed. Then assume that random(0, 15) returns 8. Thus, 8 tokens are assigned to speed or firing rate of the transition t.

Step 3: Update tokens after firing t. For each substrate and product of t, $c * speed$ tokens are subtracted and added to the corresponding substrate and product. Enzyme of t remains unchanged. After firing t, 8 tokens are subtracted from the tokens of AA and added to tokens of 5-HPETE (Fig 3.25 a). Finally, after firing t a state illustrated in Fig 3.26 b is obtained. Tokens of AA, 5-HPETE and Alox5 are 12, 23, and 15, respectively.

Defining the initial state of MPN

The aim of this work is to understand perturbation of metabolic concentrations using the

RNA-Seq expression data. To accomplish this, a Petri net-based approach (MPN) is used to understand changes in metabolic concentrations. Setup of the initial state is required to run the MPN. For using the MPN, experimentally measured expression data is required to be transformed in tokens. Using expression data the initial state of enzymes of the AA model is shown in Table 3.5. This table is used to transform expression data into tokens. For transformation, first of all, row-wise expression values of each gene in Table 3.5 are used to calculate mean and standard deviation (SD). A cutoff is set to mean+0.5*SD that is used for cutoff1 and another cutoff is set to mean-0.5*SD that is used cutoff2. Expression data was transformed into tokens using the following steps:

1) 40 tokens are chosen for the high expression (if expression is greater than the cutoff1) (Ruths et al., 2008).

2) 5 tokens are chosen for the low expression (if the expression is less than cutoff2).

3) 20 tokens are chosen if the expression is in between cutoff1 and cutoff2.

Table 3.6 is obtained by using aforementioned steps to Table 3.5 and subsequently Table 3.6 is used to initialize the AA model for MPN simulation.

The source, phosphatidylcholines (PC), and the sink_AA are fixed components in the AA model and they are set to 20 tokens. Except source and sink, the remaining metabolites of the AA model is set to 0 tokens.

The MPN is simulated by two-time scale approach same as the SPN (section 2.8).

Table 3.6. Tokens of Petri net model of AA metabolism.

enzyme	AJ control	AJ DDC	B6 control	B6 DDC	PWD control	PWD DDC
LTA4H	5	20	40	40	40	5
GPX	5	20	5	40	5	40
ALOX5	5	20	20	40	5	20
PLA2	5	20	5	40	5	40
PTGIS	20	20	5	40	5	5
LTC4S	5	20	40	40	5	20
ALOX12	5	20	5	40	5	40
CYP4F	5	40	5	40	5	40
CYP2J	5	40	5	20	5	40
TBXAS1	5	40	5	40	5	20
PTGDS	5	40	5	40	5	5
PTGS1	20	40	20	40	5	5
PTGES	5	5	40	5	40	20
ALOX8	20	5	5	20	40	5
GGT1	5	40	5	20	5	40
CBR	5	40	5	40	5	5

Comparison of relative change in metabolic concentrations (Monte Carlo ODE-based vs Peri net-based)

Enzymes of the AA model is initialized by Table 3.5 and 3.6 for the strategy Monte Carlo ODE-based and Petri net-based (MPN), respectively. For three mouse strains in control and DDC-treated conditions, both strategies are used to understand the changes in metabolic concentrations based on perturbation in gene expression data.

We sampled 1000 sets of kinetic parameters from a log-normal distribution (μ=0.35 and σ=3.5) for the Monte Carlo ODE-based approach (Wierling et al., 2012). The geometric mean (geomean) ratios (section 2.5) were calculated for the simulated metabolites.

The AA model is simulated 20000 times using MPN. Each simulation runs for 40 time blocks (section 2.8). Average over 20000 simulations were used as time block-course simulation. After 15 time-blocks a kind of steady state was found for each metabolite of the AA model. To judge a kind of steady state for each metabolite of the AA model, first of all, I calculated the mean concentration value and standard deviation of time block-course simulation from 15 to 40 blocks and then calculated coefficient of variations. Each components of the AA model has <1% variation in concentration between time block 15 to

40. Thus, the AA model reached a kind steady state and aforementioned the mean concentration value is used as the steady state concentration. Ratios (DDC-treated vs. control) of concentrations were calculated for the simulated metabolites.

For mouse strains A/J, C57Bl6 and PWD, relative changes between DDC-treated vs control conditions of metabolic concentrations are shown in Figure 3.27. An analysis of the direction of change was done, in which the metabolites were classified as upregulated or downregulated. Metabolites are considered as upregulated if their log2 ratios were greater than 0.25, as downregulated if their log2 ratios were below -0.25, and as unchanged if their log2 ratios were between -0.25 and 0.25. Based on this classification, metabolites were categorized as in agreement or disagreement between experimental and simulated results. If both, experimental and simulated metabolite changes went into the same direction (up or down), they were considered as correctly predicted.

Out of 13 experimentally measured metabolites of the AA pathway, 5, 5, and 7 metabolites were correctly predicted regarding up/down regulation using Monte Carlo ODE-based for the mice A/J, C57Bl6 and PWD respectively. Using the MPN approach 7, 8, and 9 metabolites were predicted correctly regarding up/down regulation for the mice A/J, C57Bl6 and PWD, respectively. The MPN has higher prediction rate of the metabolic changes than Monte Carlo ODE-based for the AA model. However, both methods are not able to predict the quantitative metabolic concentration ranges.

Using ODE-based Monte Carlo method the metabolites 12-HETE and 14,15-EET were always in agreement in A/J, C57Bl6 and PWD mice. Thromboxane B2 (TXB2) and prostaglandin D2 (PGD2) were in agreement in C57Bl6 and PWD mice. Leukotriene B4 (LTB4) was in agreement in the A/J and PWD mice. Arachidonic acid (AA), and 15- and 5-HPETE were in disagreement in all three mice.

Using Petri net-based (MPN) method the metabolites 15-HETE, 15- and 5-HPETE, and 14,15-EET were always in agreement in A/J, C57Bl6 and PWD mice. AA, 12-HETE and PGD2 were in agreement in the C57Bl6 and PWD mice. Prostaglandin E2 (PGE2) was found to be in agreement in A/J and C57Bl6 mice. Leukotriene B4 was in disagreement in all three mice.

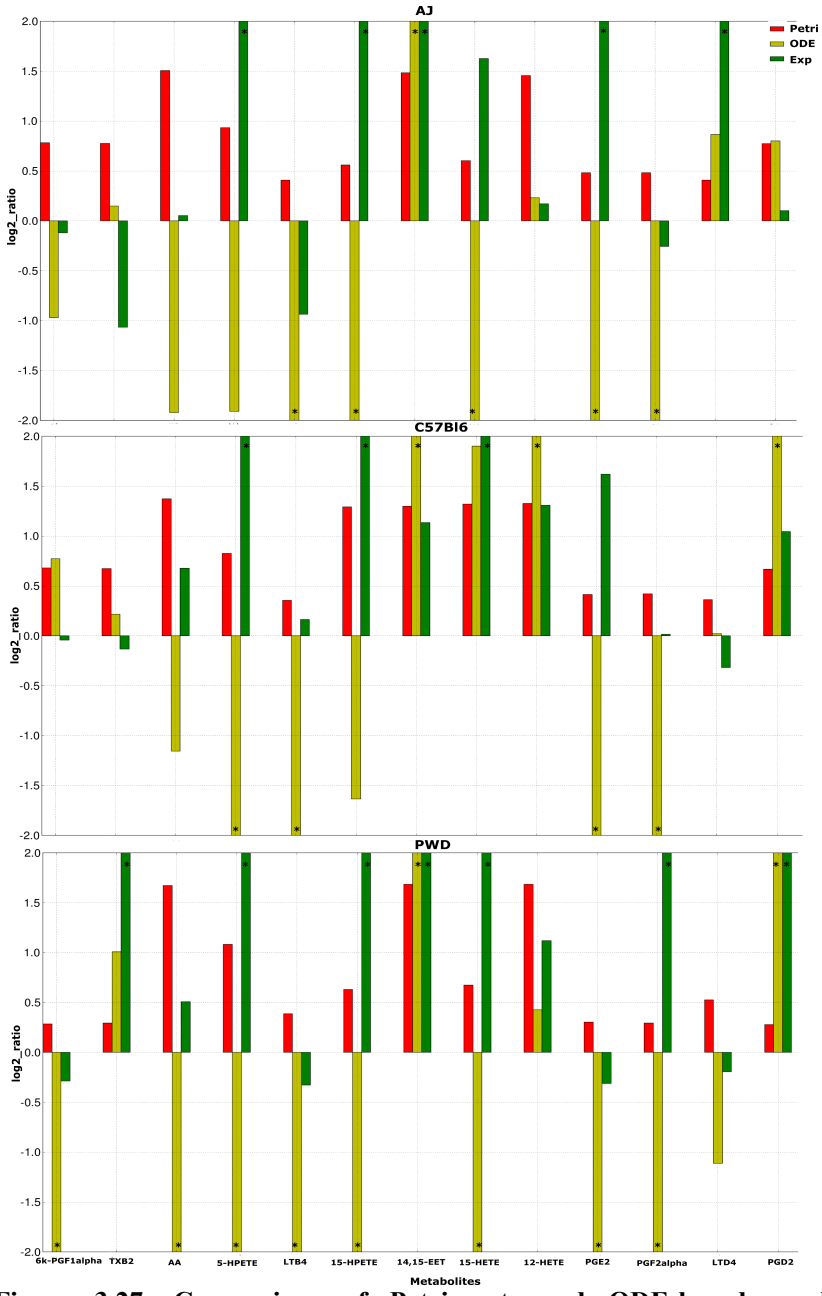

Figure 3.27. Comparison of Petri net- and ODE-based results. Comparison is done for the relative change (DDC-treated vs control) of metabolic concentration using Petri net- and Monte Carlo ODE-based method. Bar associated with '*' indicates that log2 ratio were greater or less than 2 or -2.

One limitation of both simulation approaches is that they can be only used for the qualitative or semi-quantitative prediction. Thus, in the subsequent section I have built a small-scale kinetic model of the AA pathway for a better understanding of metabolic changes as well as the range of the absolute experimental concentration.

3.5.3 Quantitative model for the arachidonic acid metabolism

Based on pathways analysis of gene expression and metabolic profiles arachidonic acid (AA) pathways was identified as an affected candidate due to DDC-treatment (section 3.1). Furthermore, the metabolites PGD2, 5-HPETE, 15-HETE and 15-HPETE of the AA pathway have been observed to be significantly ($p<0.05$) changed in concentration due to DDC-treatment in mouse liver samples. One aim of this study is to construct a small-scale kinetic model that can explain the deregulation of aforementioned metabolites of the AA pathway. In addition, the AA kinetic model can be used for *in silico* knock-out experiment such as the identification of drug targets.

Arachidonic acid (AA) is a fatty acid usually coming from dietary animal sources or being synthesized from dietary linoleic acid. AA is present in cell membranes as a part of phospholipids. AA is released from phospholipids by phospholipase A2 (PLA2) and subsequently it acts as a precursor of prostaglandins and their related compounds, the prostacyclins, thromboxanes and leukotrienes. There are several physiological effects of prostaglandins described in the literature, such as inflammatory response, pain, fever, blood pressure, blood clotting, and regulation of sleep/wake cycle (Hayaishi, 1991; Miller, 2006). Cyclooxygenase-1, also known as prostaglandin H2 synthase 1 (Ptgs1), catalyzes the reaction of arachidonate to prostaglandin H2 (PGH2), which is the precursor of other prostaglandins, prostacyclins, and thromboxanes. The enzyme 5-lipoxygenase (Alox5) produces 5-HPETE, which is a precursor of leukotrienes.

There are only a few kinetic models of the AA metabolism available in the literature to study anti-inflammatory drugs of human polymorphonuclear leukocytes (Yang et al., 2007) and macrophage cells (Gupta et al., 2009). Here, an *in silico* model of the AA metabolism in mouse liver was developed based on information from KEGG (Kanehisa et al., 2012) and a biochemical text book (Michal and Schomburg, 2013) to study the regulation of the perturbed metabolites using the related transcriptomics and proteomics data. The structure of the model is depicted in Figure 3.28.

The downstream synthesis of prostaglandins and leukotrienes is initiated by signal transduction and subsequent release of AA catalyzed by phospholipase A2 (PLA2). 15-lipoxygenase (ALOX15) is an important regulator of inflammation and apoptosis and its expression is regulated by a cytosolic signaling complex with protein kinase C delta (PKCD) and phosphorylated STAT3 (Bhattacharjee et al., 2006). Thus, RPPA data of phosphorylated ERK and STAT3 could be used subsequently for the modeling to describe the activity of PLA2 and PKCD, respectively.

The eicosanoid pathway is regulated by several feedback mechanisms (see Figure 3.28). The release of AA is controlled by an inhibitory link between AA and phospholipase A2 (PLA2, R1 in Figure 3.28, (Geddis et al., 2004)). The production of leukotriene A4 (LTA4) and 5-

HPETE is controlled by inhibiting 5-lipoxygenase (ALOX5; R2, R4, (Aharony et al., 1987; Lepley and Fitzpatrick, 1994)). The synthesis reactions of 15-HPETE, 15-HETE, and PGH2 are controlled by feedback inhibition of 15-lipoxygenase (ALOX15), glutathione peroxide and PTGS1, respectively (R6, R7 and R9, (Mahipal et al., 2007; Yang et al., 2007). The synthesis of leukotrienes from AA via 5-HPETE, such as leukotriene A4, is catalyzed by two different enzymes, ALOX5AP and 5-lipoxygenase (ALOX5; R2, R4).

I hypothesize that changes of metabolite concentrations of the AA pathway can be explained by changes in mRNA expression. To address this hypothesis, changes in Vmax values of enzymatic reactions were approximated by fold changes of RNA-Seq expression of the respective enzymes due to DDC treatment. The fold change of PLA2 was approximated by Pla2g4a which is expressed at a low level (<1 RPKM; Gan et al., 2010) but not differentially. Therefore, I hypothesize that PLA2 is at least present to trigger AA release upon its activity state, but it is not differentially expressed due to DDC treatment. Similarly, ALOX5 and ALOX15 are approximated by Alox5 (p-value=0.13) and Alox15 (p-value=0.62), respectively, which are expressed at low level (<1 RPKM) and we hypothesize that the expression of these genes are not affected due to DDC-treatment. PTGDS is approximated by Ptgds expression, but it is not significantly changed due to DDC treatment (p-value=0.64). If there are several isoenzymes that can catalyze the same reaction the most significantly differentially expressed gene was chosen. For instance, GPX can be approximated by *Gpx1*, *Gpx2*, *Gpx3*, *Gpx6* and *Gpx7*. Since *Gpx3* is significantly affected (p=2.66e-041) and highly expressed it is used to initialize the GPX expression in the model. Fold changes of the enzymes PTGS1, ALOX5AP and PKCD are approximated by expression values of *Ptgs1*, *Alox5ap* and *Prkcd*, respectively.

Within our experiments the protein amount of phosphorylated ERK (p-ERK) and phosphorylated STAT3 (p-STAT3) were measured for both control and DDC-treated mice using the reverse phase protein array (RPPA) technology. p-ERK and p-STAT3 are upregulated due to DDC treatment (fold changes of p-ERK and p-STAT3 in A/J, C57Bl6, and PWD are 1.90, 2.09, 1.98, and 1.41, 1.42, 2.03, respectively). The effect of p-ERK on PLA2 is modeled by the product of the RNA-seq expression value of PLA2 for the DDC-treatment simulation and the p-ERK ratio. Similarly, the effect of the complex formation of PKCD and p-STAT3 on the ALOX15 activity is described by the product of the RNA-seq expression value of ALOX15 of the DDC-treatment and the fold change of p-STAT3 and PKCD, respectively. Phosphatidylcholine (PC) is modeled as a fixed component and as its initial concentration the measured concentration of phosphatidylcholine C33:2 was used (Table 3.7). The experimentally identified fold changes of all enzymes are summarized in Table 3.7 and were used to approximate the fold changes of the maximum reaction rates (Vmax) of the respective reactions.

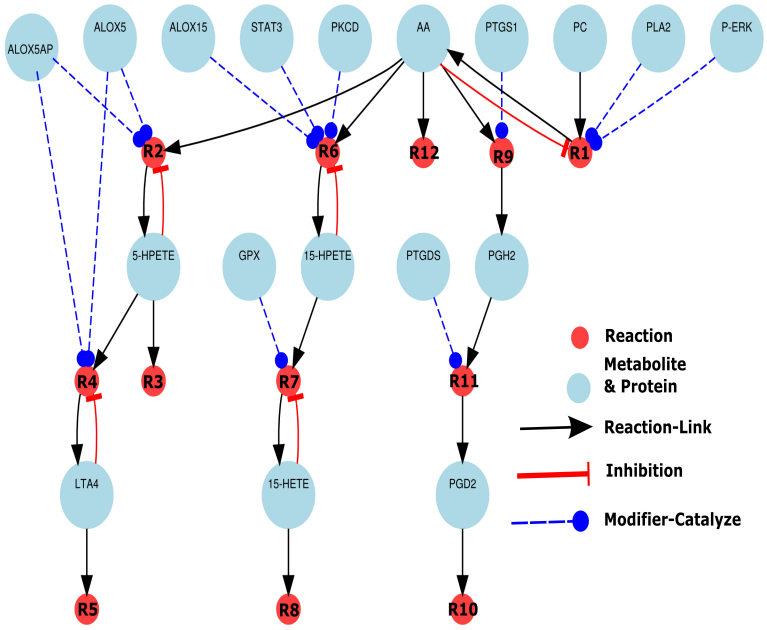

Figure 3.28: **Detailed kinetic model of arachidonic acid / eicosanoid metabolism.** Abbreviations of the AA model: phosphotadicholine (PC), arachidonic acid (AA), phospholipase A2 (PLA2), phosphorylated ERK (P-ERK), Cyclooxygenase-1, also known as prostaglandin G/H synthase 1 (PTGS1), protein kinase C delta (PKCD), signal transducer and activator of transcription 3 (STAT3), arachidonate 15-lipoxygenase (ALOX15), arachidonate 5-lipoxygenase (ALOX5), arachidonate 5-lipoxygenase-activating protein (ALOX5AP), prostaglandin H2 (PGH2), prostaglandin D2 (PGD2), prostaglandin D2 synthase (PTGDS), 15-hydroxyeicosatetraenoic acid (15-HETE), 5- and 15-hydroperoxyeicosatetraenoic acid (5- and 15 -HPETE), leukotriene A4 (LTA4), glutathione peroxidase (GPX).

Table 3.7. Initial values of the arachidonic acid / eicosanoid metabolism model. The enzyme concentrations of the control state were always initialized with 1.0 nM. The DDC initial value always reflects the fold change between the DDC-treated and the control expression value. The absolute concentration of phosphatidylcholine C33:2 (PC*) comes from MS-analysis and is measured in µM for control and DDC-treatment conditions of AJ, B6, and PWD mice

symbol	AJ_ control	AJ_ DDC	B6_ control	B6_ DDC	PWD_ control	PWD_ DDC
PTGS1	1.00	2.13	1.00	1.41	1.00	2.09
ALOX5	1.00	1.00	1.00	1.00	1.00	1.00
PTGDS	1.00	1.00	1.00	1.00	1.00	1.00
STAT3	1.00	1.41	1.00	1.42	1.00	2.03
PKCD	1.00	4.30	1.00	4.48	1.00	3.85
GPX	1.00	6.51	1.00	11.70	1.00	19.42
PLA2	1.00	1.90	1.00	2.09	1.00	1.98
ALOX5AP	1.00	4.81	1.00	6.30	1.00	14.23
ALOX15	1.00	1.00	1.00	1.00	1.00	1.00
PC*	269.00	206.00	249.00	209.00	266.00	255.00

Quantitative modeling requires kinetic parameters for each reaction. Such parameters have been mostly measured for very specific experiments and conditions. The availability of appropriate kinetic parameters to model cellular systems is limited, e.g., kinetic databases like BRENDA (Scheer et al., 2011) and SABIO-RK (Wittig et al., 2012) can only provide a limited number of kinetic parameters. To build a quantitative model for the AA liver metabolism kinetic parameters of the model were optimized. Using data of DDC treatment vs control I built a quantitative model to explain changes of metabolites of the eicosanoid pathway due to treatment by DDC that mimics the NASH phenotype.

Kinetic Equations. Mass action kinetics is used for reactions R3, R5, R8, R11 and R12 as well as R1, R2, R4, R6, R7 and R9 are model with Michaelis-Menten kinetics (Figure 3.27; Table 3.8; Yang et al., 2007). Mass action kinetic is used for R10 which is catalyzed by an enzyme (Figure 3.27; Table 3.8; Gupta et al., 2009).

Model Optimization. Model parameters were optimized using experimentally determined metabolite concentrations of the model components. Metabolite concentrations were measured after 8 weeks of DDC treatment and interpreted as steady state concentrations for the model. The metabolites PGD2, 5-HPETE, 15-HPETE and 15-HETE have been found as being affected due to DDC treatment. These metabolites as well as upstream AA were used for model optimization. Therefore, the model was simulated into its steady state and the euclidean distance between simulated steady state concentrations and experimental

metabolite data of the aforementioned metabolites were used in an objective function to optimize the kinetic parameters using COPASI (Hoops et al., 2006). Parameter optimization was done using the metabolic concentrations of C57Bl6 control and C57Bl6 DDC mice (section 2.9). Kinetic parameters, i.e., the maximum reaction rates (V_{max}), Michaelis-Menten constants (K_m) and inhibition constants (K_i) were optimized based on experimental data.

The following five metabolites of the AA model are used to formulate an objective function: AA, 5-HPETE, and 15-hydroperoxyeicosatetraenoic acid (15-HPETE), 15-Hydroxyeicosatetraenoic (15-HETE), and prostaglandin D2 (PGD2). Let aforementioned metabolites be in a set M. A constructed objective function, *f*, is (section 2.9):

$$f(K) = \mathrm{sqrt}\left(\sum_{x \in M} \left(\frac{\left([x]_{ec} - [x]_{sc,K} \right)}{[x]_{ec}} \right)^2 + \left(\frac{\left([x]_{ed} - [x]_{sd,K} \right)}{[x]_{ed}} \right)^2 \right) \tag{3.6}$$

where, K is set of kinetic parameters of the AA model. “ec” and “sc” represent experimental and simulated steady state concentration in control condition, while “ed” and “sd” represent experimental and simulated steady state concentration in DDC-treated condition. The objective function, *f,* is minimized to obtain optimal kinetic parameters of the AA model using the genetic algorithm of COPASI with a population size of 50 and a generation size 500. The identified kinetic parameters are listed in Table 3.8.

Table 3.8. Kinetic equations and their parameters of the arachidonic acid / eicosanoid metabolism model. *A*, A_1, A_2 and A_3 are the ratios of gene expression or protein levels of the respective enzymes between DDC-treated vs. control mice. *S* is the substrate of the corresponding reaction. Squared brackets refer to concentration values and curly braces indicate fold changes of DDC treatment vs. control. The K_m value of PLA2 is taken from BRENDA (denoted by *), and other parameters of the table were fitted using the AJ mice metabolic concentrations.

Reactions	**Kinetic equations**	**Kinetic parameters**
R1: PC → AA	$\frac{V_{max}\times A_1\times A_2\times S}{\left(S+K_m\times\left(1+\frac{I}{K_i}\right)\right)}$	Vmax=20.56 nM2s-1, Km=2500 nM(*) , Ki=100 nM, A1={PLA2}, A2={PERK}, I=[AA] nM, S=[PC] nM
R2: AA → 5-HPETE	$\frac{V_{max}\times A_1\times A_2\times S}{\left(S+K_m\times\left(1+\frac{I}{K_i}\right)\right)}$	Vmax=0.01 nM2s-1 , Km=0.0107 nM, Ki=8.603 nM, A1={ALOX5}, A2={ALOX5AP}, I=[5-HPETE] nM, S=[AA] nM
R3: 5-HPETE →	$K_{cat}\times S$	S=[5-HPETE], Kcat=0.0012s-1
R4: 5-HPETE → LTA4	$\frac{V_{max}\times A_1\times A_2\times S}{\left(S+K_m\times\left(1+\frac{I}{K_i}\right)\right)}$	Vmax=9.7953 nM2s-1, Km=99.913 nM, Ki=0.709 nM, A1={ALOX5}, A2={ALOX5AP}, I=[LTA4] nM, S=[5-HPETE] nM
R5: LTA4 →	$K_{cat}\times S$	S=[LTA4] nM, Kcat=0.0012 s-1
R6: AA → 15-HPETE	$\frac{V_{max}\times A_1\times A_2\times A_3\times S}{\left(S+K_m\times\left(1+\frac{I}{K_i}\right)\right)}$	Vmax=0.003 nM3s-1, Km=0.067 nM, Ki=1.004 nM, A1={ALOX15}, I=[15-HPETE] nM, A2={PKCD}, A3={PSTAT3}, S=[AA] nM
R7: 15-HPETE → 15-HETE	$\frac{V_{max}\times A\times S}{\left(S+K_m\times\left(1+\frac{I}{K_i}\right)\right)}$	Vmax=1.098 nMs-1, Km=1.58 nM, Ki=0.0106 nM, A1=ALOX15, I=[15-HETE] nM, S=[15-HPETE] nM
R8:15-HETE →	$K_{cat}\times S$	S=[15-HETE], Kcat=0.00127 s-1
R9: AA → PGH2	$\frac{V_{max}\times A\times S}{\left(S+K_m\times\left(1+\frac{I}{K_i}\right)\right)}$	Vmax=0.168 nMs-1, Km=3.876 nM, Ki=0.013 nM, A={PTGS1} I=[PGH2], S=[AA] nM
R10: PGH2 → PGD2	$V_{max}\times S\times A$	Vmax=0.067 nMs-1, A={PTGDS}, S=[PGH2]
R11:PGD2 →	$K_{cat}\times S$	S=[PGD2] nM, Kcat=0.052 s-1
R12: AA →	$K_{cat}\times S$	S=[AA] nM, Kcat=0.00096 s-1

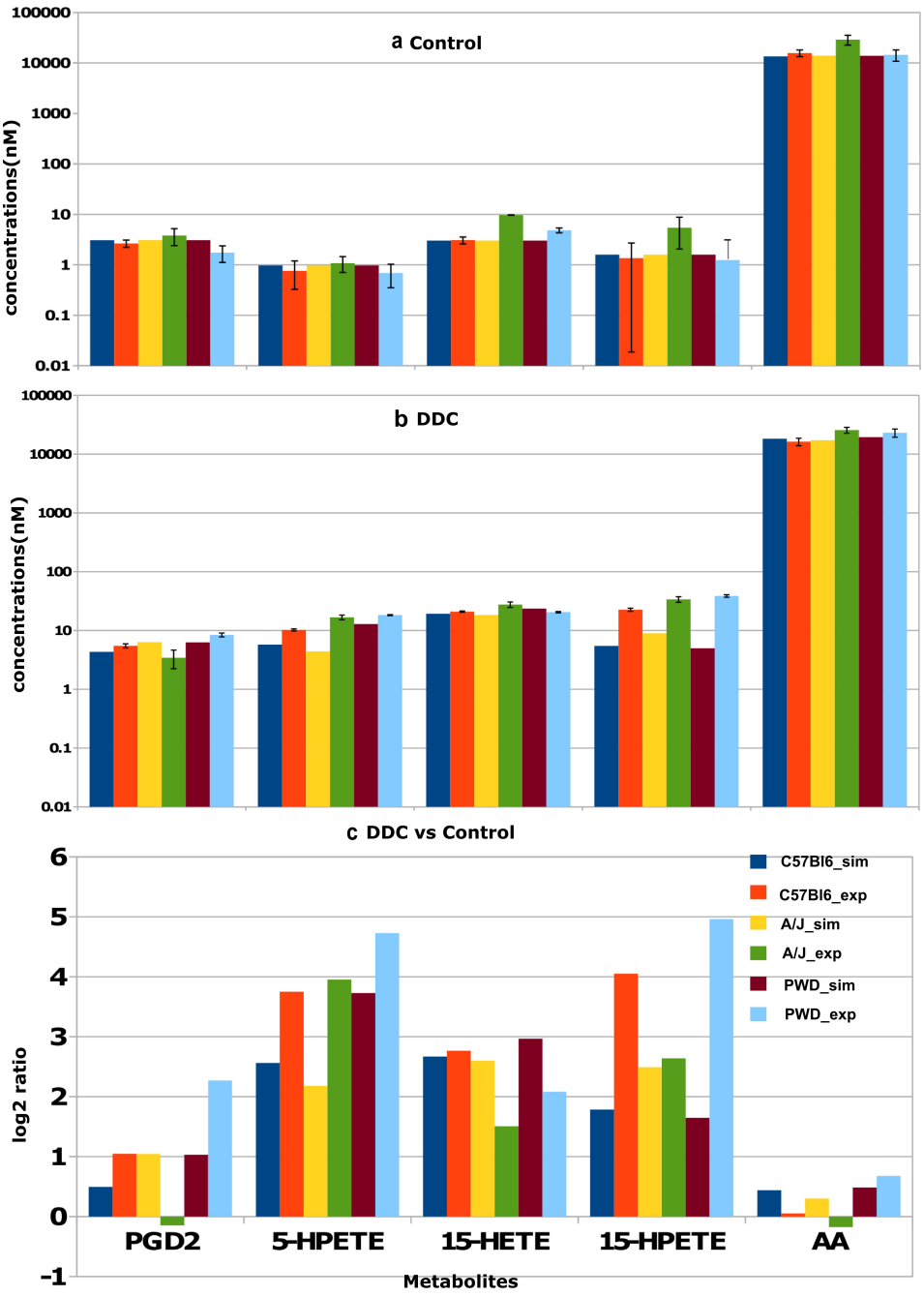

Figure 3.29. Experiment vs simulated metabolites. Comparison of simulated steady state and experimental metabolite concentrations of all three mouse strains A/J, C57BI6, and PWD for **a)** control **b)** DDC treatment, and **c)** their respective ratio.

Results of the fitted model for the metabolites AA, PGD2, 5-HPETE, 15-HPETE, and 15-HETE using the trained C57Bl6 model were compared to the experimental results (Figure 3.29; C57Bl6_sim vs. C57Bl6_exp). Except of AA all of the aforementioned metabolites were upregulated (>1.5 fold) in both, experimental and simulated data, due to DDC in C57Bl6 mice (Figure 3.29c). AA was not found altered in the experimental dataset, but was slightly upregulated in the simulation (1.35-fold).

Using the trained model we predicted DDC-induced metabolite changes for A/J and PWD. We observed an upregulation (>1.5-fold) of 5-HPETE, 15-HPETE and 15-HETE in both strains, which is in line with the experimental data. The concentration of PGD2 is found increased (>1.5-fold) in both experimental and simulated data for PWD (Figure 3.29c). In A/J our modeling approach predicts a 2-fold increase of PGD2 concentration in response to the DDC-treatment, but the experimental data showed no changes (0.9 fold). This disagreement can be strain-specific because one can expect changes in the concentration of PGD2 due to an upregulated gene expression of PTGS1 which is located upstream in the metabolic pathway. AA is upregulated (>=1.5-fold) in DDC-PWD mice, which is concordant with the experimental data, while for A/J, AA is not found as changed (0.9 fold) in the experimental data, whereas a minor upregulation of 1.2-fold was predicted from simulated data (Figure 3.29c).

Key regulatory enzymes of the DDC mouse model. To identify key regulatory enzymes of the AA/eicosanoid pathway in DDC treated mice, a kind of sensitivity analysis of the trained model was performed. Starting with DDC-treated mice of the A/J strain, each enzyme or enzyme combinations were reverted to its/their activity in normal, untreated condition. This analysis was performed for ALOX5AP, GPX, PKCD, PTGS1, pERK, and pSTAT3 whose activity was found to be perturbed due to DDC-treatment. A reference state was defined where the activity of all enzymes is equal to the normal condition. Results of this analysis are depicted in Figure 3.30. Reverting a single enzyme or combinations of two enzymes are not sufficient to bring the DDC-treated state back to normal because it affects only some branches of the model as for example a change in the activity of ALOX5 has an effect on the regulation of downstream metabolites 5-HPETE and LTA4 (cf. Figure 3.28). We found that combinations of the enzymes ALOX5AP, PKCD and PTGS1 with either pERK or pSTAT3 can bring the DDC-treated metabolic state back to normal, with the exception of 15-HETE and AA. Only the combination ALOX5AP, PERK, PKCD, STAT3, and PTGS1 or the combination of all six enzymes is able to bring back the DDC-treated state to normal.

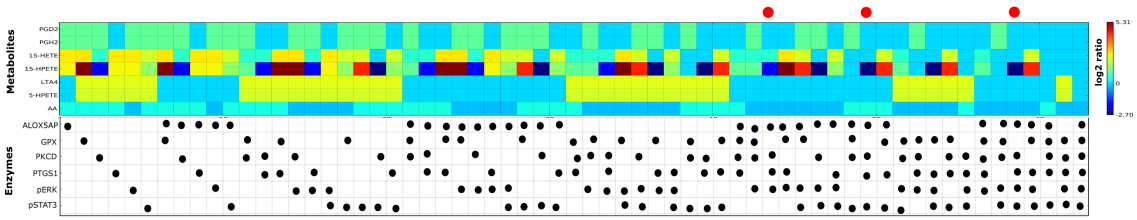

Figure 3.30. Heat map of sensitivity analysis. A sensitivity analysis of the arachidonic acid/ eicosanoid metabolism model was done by simulation of individual models, which brought in normal condition from A/J DDC state by the combination of enzyme as given by their respective expression value of the A/J control data. Color key indicates log 2-ratios. Black dots represent corresponding enzymes got back from DDC-treated state to normal state. Red dots indicate that combination of enzymes are able to bring back DDC-treated to normal conditions.

Simulating drug effects of DDC mouse model. To study potential drug targets for steatosis or inflammation *in silico* drug tests were performed. For instance, cyclooxygenase-2 (COX-2) is a frequent target of anti-inflammatory drugs (Jüni et al., 2002). The activity of each enzyme of the A/J DDC model was inhibited by 3-, 6- and 9-fold, respectively. The effects of those enzyme inhibitions on the regulation of metabolites are shown in Figure 3.31. The inhibition of PTGS1 and ALOX5AP leads to a downregulation of the respective downstream components (either PGH2 and PGD2 or 5-HPETE and LTA4). ALOX15 is activated by PKCD and pSTAT3, and the inhibition of PKCD or pSTAT3 leads to a downregulation of the downstream components 15-HPETE and 15-HETE. It can be expected that the inhibition of glutathione peroxidase (GPX) leads to a downregulation of the downstream component 15-HETE, but the model predicted an upregulation of 15-HPETE and no change in 15-HETE. This is due to the complex regulation of ALOX15, i.e., an activation of the complex PKCD and pSTAT3 and an inhibition of GPX by 15-HETE.

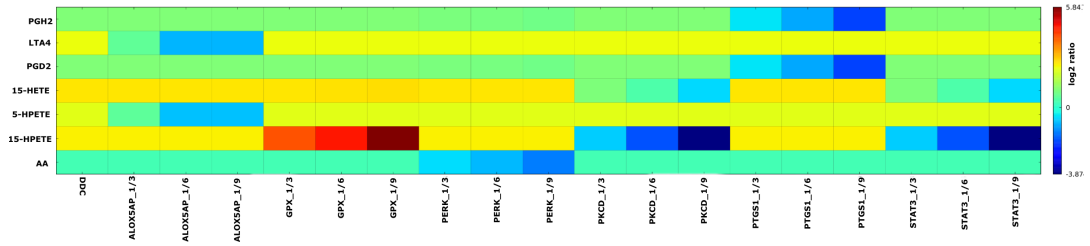

Figure 3.31. *In silico* drug analysis. An *in silico* drug testing of the arachidonic acid/ eicosanoid metabolism model was done by simulation of individual models with a downregulated concentration of each enzyme as given by their respective expression value of the A/J DDC data by a 1/3, 1/6 and 1/9 times.

In conclusion, I propose a novel hypothesis: the prediction of changes in metabolic

concentrations based on gene expression data. The AA model is used to demonstrate the hypothesis. To understand the changes of metabolic concentrations at a qualitative level I used Monte Carlo ODE- and Petri net-based methods. In addition, I modified simulation algorithm of the Petri net-based method SPN in such a way that it can also handle metabolic reactions. The modified simulation method predicts changes of metabolites at qualitative level with higher accuracy than than Monte Carlo ODE-based. Both methods cannot predict the correct range of metabolic concentrations. Thus, I also build a detailed kinetic model of AA to understand quantitative metabolic concentrations and the model is able to predict changes of metabolic concentrations at qualitative and quantitative levels. The model was used to identify key regulatory enzymes and to simulate drug effects of DDC mouse model.

4 Discussion

High-throughput techniques based on RNA-Seq technology provide a flood of data associated with biomolecules at large scale. These data can be analyzed using methods of bioinformatics and systems biology in a variety of ways depending on the purpose of research. Identification of biomarkers for a disease is one of the common tasks where RNA-Seq data can be effectively employed. Several bioinformatics methods for identifying candidate disease genes were published (Chen et al., 2009; van Driel and Brunner, 2006). In systems biology, a living organism is interpreted as an interacting and dynamical network of genes, proteins, and biochemical reactions which regulate the function and behavior of that organism (Chen and Wu, 2012). Systems biology is considered as a useful approach to reveal the essential mechanisms of initiation and progression for complex diseases such as cancer, steatohepatitis, cardiovascular disease, and diabetes (Chen and Wu, 2012). The classical biomedicine based on experimental biology which is mainly based on studying individual genes and proteins as simple systems has made marked progress in the comprehension of diseases progression. However, this biomedicine area still faces challenges in the understanding of complex diseases (Chen and Wu, 2012). Therefore, systems biology techniques may be useful for analyzing the molecular mechanisms of complex diseases and providing new insights to cure these complex diseases.

Nonalcoholic fatty liver disease (NAFLD) comprises a wide spectrum of liver damage, ranging from simple steatosis to steatohepatitis, advanced fibrosis and cirrhosis (Alba and Lindor, 2003). To better understand the NAFLD progression a mouse model system was used. Using molecular data of mouse liver samples I have performed pathway analysis to identify pathways that may be involved in liver disease (section 3.1). Moreover, I performed module analysis to identify most relevant modules related to liver disease based on fully correlated fluxes and elementary flux modes (section 3.2). In section 4.1 most relevant pathways and modules that may be involved in liver disease are discussed.

To identify different degrees of flux regulation in all three mouse strains I have applied flux-based analysis in which gene expression data of mouse liver samples are integrated into a genome-scale metabolic model (section 3.3). I discuss the results of the integration of gene expression data into the metabolic model in section 4.2.

Furthermore, I have identified *in silico* drug targets for the liver disease: steatosis, steatohepatitis and HCC (section 3.4). In section 4.3 most relevant candidates for potential drug targets are discussed.

Finally, I constructed the arachidonic acid (AA) metabolic model that is identified as a deregulated candidate in molecular data of liver. The model was applied for the prediction of changes of metabolic concentrations using gene expression data (section 3.5). In section 4.4, I discuss results related to arachidonic acid metabolic model.

4.1 Pathways and modules related to liver disease

To understand NAFLD progression a mouse model system was used. Three mouse strains were fed a DDC-supplemented diet that produces steatohepatitis-like phenotypes (section 3.1.1). Histidine, beta-alanine and purine metabolisms were identified as relevant pathways that may be involved in liver diseases using gene expression data of mouse liver samples (section 3.1). Furthermore, the hepatic SAMe metabolism was found to be perturbed at different levels among mouse strains (section 3.1). Therefore, the hepatic SAMe metabolism may be linked to the mouse phenotypes. Using expression and metabolic profiles the arachidonic acid (AA) metabolism was found to be deregulated under DDC-treated condition in all three mouse strains. In addition, based on expression profiles fatty acid metabolism and peroxisome proliferator-activated receptors (PPAR) signalling pathways were found to be deregulated. Jin et al., (2012) analyzed microRNA data in a rat model of steatohepatitis and identified beta-alanine, purine and fatty acid metabolism along with PPAR signalling which is inline with my study. Some pathways such as histidine, hepatic SAMe and AA metabolism are identified in my study but not in the study of Jin et al., while calcium signaling pathway was identified in the study of Jin et al. but not in my study. This discrepancy might be due to the different systems were analysed, namely in this study mouse data is used while in Jin et al. human data is used. Furthermore, Haque et al., (2010) identified pathways including xenobiotics by cytochrome P450, arachidonic acid metabolism, PPAR signaling pathway, and pyrimidine pathway in a mouse model of steatohepatitis which is induced by methionine and choline (MCD) diet. These pathways are inline with my analysis in which mice were fed a DDC diet. This indicate that DDC and MCD diets affect similar set of genes.

The genes *Mat1a*, *Srm*, *Sms*, *Dnmt1*, *Achy* and *Bhmt* and metabolites spermidine, spermine, putrescine and methionine of SAMe metabolism along with prostaglandin E2 and leukotriene D4 of AA metabolism were found to be deregulated at different levels among mouse strains (Figures 1e & f). SAMe is a key methyl-group donor for phosphatidylcholine synthesis that is required, e.g., for the export of very-low-density lipoproteins (VLDL) and triglycerides from the liver (Cano et al., 2011). SAMe and methionine metabolism was found to be perturbed in NASH patients (Kalhan et al., 2011) and it may play role in development of NAFLD, such as NASH (Mato et al., 2013). A study suggests that depletion of hepatic anti-oxidants (e.g. reduced glutathione and SAMe) promotes oxidative stress and may induce a histological steatohepatitis (Anstee and Goldin, 2006). Gene *Mata1* is expressed in liver and is a marker of hepatocyte differentiation (Mato et al., 2002). The expression of *Mata1* was found to be downregulated in A/J mice and it may lead to reduction of SAMe production from dietary methionine. When the level of SAMe expression reduces markedly then it causes liver injury and subsequently leads to the development of steatohepatitis, cirrhosis and HCC (Mato et al., 2002). Thus, SAMe metabolism may explain susceptibility of A/J mice to DDC-treatment.

The concentration of putrescine was found to be upregulated for this study in A/J mice

(Figure 1e) and the Dettmer et al., (2011) study of HFD-fed mice. On the contrary, the concentrations of spermidine was found to be upregulated (Figure 1e) which is not inline with the study of Dettmer et al., (2011). This discrepancy may be due to different type of model systems.

Prostaglandin E2 (PGE2) is found to be upregulated in only A/J and C57Bl6 mice (Figure 1e). PGE2 promotes inflammation after binding to prostaglandin E receptor 2 (EP2) (Miller, 2006) which regulates an inflammatory marker tumor necrosis factor alpha (TNFα) in NASH mouse model (Tosello-Trampont et al., 2012). To explain inflammation phenotypes through a path PGE2-EP2-TNFα I observed mRNA expression of *tnf* which corresponds to TNFα, which is markedly upregulated for A/J and C57Bl6 mice. Upregulation of PGE2 and TNFα for A/J and C57Bl6 mice may explain NASH development by inflammation.

Identification of modules. Fully coupled modules may capture the functional associations between genes and consequently may provide refined insights into the modes of transcriptional regulation of metabolism. Identification of fully coupled modules is introduced in the study of Papin et al., (2004). Fully coupled modules explain better co-regulation of metabolic genes than the modules that are identified by network distance (Notebaart et al., 2008). To address the co-regulation of fully coupled modules, I used gene-expression data of mouse liver samples and found moderately high co-expression of metabolic genes which is inline with the study of Notebaart et al., (2008).

To the best of my knowledge the identification of most relevant fully coupled modules that may be involved in liver disease using gene expression data of mouse liver samples was not performed before. There is a tool, ConsensusPathDB (Kamburov et al., 2011a), that can perform a module analysis based on enrichment analysis. In ConsensusPathDB modules are predefined while in my study modules are identified based on the structure of metabolic networks. In this study I identified important fully coupled modules related to liver disease. For example, a module of cholesterol metabolism was identified and linked to the apoptosis (Figure 3.4). Furthermore, some other modules related to pathways such as fatty acid, D-alanine and creatine metabolism were identified and linked to oxidative stress (Figure 3.4). The transport of carnitine shuttle system was found to be affected (Figure 3.4). Moreover, the non-oxidative phase of pentose phosphate pathway and a module of pyrimidine catabolism were found to be perturbed (Figure 3.4). In addition, a module of cysteine that may be linked to the SAMe metabolism was identified. The SAMe metabolism may be involved in the development of steatohepatitis (Mato et al., 2013).

I have identified the most relevant modules that may be related to liver disease using elementary flux modes and liver-specific molecular data (Figure 3.10). For example, a module presented in Figure 3.10a comprises L-arginine, ornithine and putrescine, another module (Figure 3.10b) comprises D-glucose, L-glutamine and L-asparagine. The

aforementioned metabolites along with some genes of the modules were found to be differentially regulated in all three mouse strains (Figure 3.11a). Furthermore, some of genes of the module in Figure 3.10a were found to be deregulated at different levels among mice; also this module comprises several cross-talks of pathways such as arginine metabolism, urea cycle, vitamine A metabolism (Figure 3.10a). These modules might be good candidates to construct more detailed kinetic models for understanding the regulation of metabolic concentrations. Rezola et al., (2013) proposed an approach to obtain tissue specific elementary flux modes that uses tissue specific gene expression data. They applied the approach to data of liver samples and identified elementary flux modes (EFMs) associated with bile acid synthesis which is one of the main functions of hepatocytes as well as urea cycle. I identified EFMs associated with the urea cycle and degradation of various amino acids: glutamine, ornithine and asparagine which is inline with the study of Rezola et al.. However, my approach of EFMs enrichment differs from the approach of Rezola et al. (section 3.3.2). Here, I propose to enrich EFMs using deregulated gene expression data and metabolic profiles, while Rezola et al. uses gene expression data. The identified EFMs (Figure 3.10) which have deregulation at upstream (e.g., gene expression) and downstream (e.g., metabolic profiles) levels covered by molecular data can be more convincing for further experimental investigation.

Fatty acid module. A module of fatty acid metabolism comprises the process of α-oxidation, that converts pristanic acid from phytanic acid after removal of its one carbon atom (Figure 3.4a). Phytanic acid induces proliferator-activated receptor alpha (PPARalpha) *in vivo* and *in vitro* (Ellinghaus et al., 1999). PPARalpha deficient mice have defects in fatty acid oxidation that leads to severe fatty acid overload in liver, which causes severe steatosis or steatohepatitis (Reddy, 2001). PPARalpha activation is sufficiently powerful to prevent the development of dietary steatohepatitis (Ip et al., 2003). Furthermore, PPARalpha may be involved in the progression of NAFLD (Domenici et al., 2013). In short, one may expect different levels of fluxes throughout the module and concentrations of phytanic acid due to changes at different levels of *Aladh3a2* expression among mouse strains, which may affect PPARalpha by phytanic acid and subsequently may develop steatohepatitis-like phenotype. *Ehhadh* and *Acaa1a* are involved in the fatty acid β-oxidation in peroxisome. The elevated expression of *Ehhadh* and *Acaa1a* in mice with high-glucose diet suggests an increased oxidative stress (Du et al., 2010). Oxidative stress can be induced by lipids that causes steatohepatitis in mice fed an atherogenic diet (Matsuzawa et al., 2007). Thus, the elevated expression of *Ehhadh* in A/J and C57Bl6 mice with a DCC supplemented diet may lead to steatohepatitis, which is inline with their steatohepatitis-like phenotype. Furthermore, Weng et al., (2013) studied pathogenic link between nonalcoholic fatty liver and peroxisome biogenesis and in this study an elevated expression of *Acaa1a* was observed. However, in my study a downregulation of *Acaa1a* expression was observed in A/J and C57Bl6 mice which is not inline with the study of Weng et al., (2013).

Carnitine shuttle system. I have identified another module which may affect the transport of fatty acid in the cell which acts through carnitine shuttle system (Figure 3.4d). The expression of *Acadm* was found to be elevated in PWD and C57Bl6 mice, while it is downregulated in mice with reduced hepatic steatosis and lipid peroxidation (Harano et al., 2006). The elevated expression of *Acadm* was found in a study of NAFLD patients (Greco et al., 2008; Mitsuyoshi et al., 2009). Elevated expression of *Acadm* for PWD and C57Bl6 mice in my study was inline with the study of Greco et al. and Mitsuyoshi et al, while results for A/J mice did not show this behavior. Based on expression patterns of *Acadm* in the aforementioned studies along with my study for PWD and C57Bl6 mice one might speculate that the flux through the carnitine shuttle system is disturbed.

Creatine metabolic module. Another module that was identified is related to creatine metabolism (Figure 3.4b). The concentration of creatinine was found to be upregulated (>1.5-fold) only in liver samples of A/J mice. This is inline with the study of Clarke et al., (2013) in NASH patients. Furthermore, the elevated concentration of plasma homocysteine is a predictor of steatohepatitis in patients with NAFLD (Gulsen et al., 2005). However, there is a study where creatinine is not found in abnormal range in NASH patients (Crespo et al., 2001). A transgenic mouse model of Huntington's disease (HD), where HD is characterized by weight loss despite of caloric intake, was administrated with creatine analog beta guanidinopropionic acid (GPA) that reduced ATP and phosphocreatine levels in both the cerebral cortex and striatum (Chaturvedi et al., 2010). Furthermore, this administration impairs activation of liver PPARγ coactivator-1α and promotes development of hepatic steatosis with accumulation of lipids (Chaturvedi et al., 2010). The reduced creatine kinase activity described in Chaturvedi et al. is inline with the expression of *Ckm* for all three mice along with the expression of *Ckmt2* for A/J and C57Bl6 mice in my study (Figure 3.5). However, the elevated expression of *Ckmt2* for PWD mice that is characterized by high steatosis, is not inline with Chaturvedi et al. Moreover, a study reports that an elevated alanine aminotransferase activity is a predictor of NASH; however, levels of creatine kinase and creatine did not changed markedly (Puljiz et al., 2010). In contrast with the study of Puljiz et al. the expression of *Ckm* and *Ckmt2* was found to be perturbed in my study. In short, molecular data of creatine module are perturbed in my study as well as in the study of Chaturvedi et al., suggesting that the flux throughout the creatine metabolic system system may be perturbed.

Glycogen module. A module related to glycogen metabolism was identified in analysis. (Figure 3.4c). The elevated expression of a gene of the module glucan (1,4-alpha-), branching enzyme 1 (*Gbe1*) was found in A/J and C57Bl6 mice, and its expression was downregulated for PWD mice (Figure 3.8). The elevated expression of *Gbe1* was found in mice fed a high-fat diet that induced hepatosteatosis (Oyadomari et al., 2008), which is inline with the expression of A/J and C57Bl6 mice. Furthermore, a low-choline diet supports

development of hepatosteatosis, liver and muscle damage, and lymphocyte apoptosis in humans (Niculescu et al., 2007). In the study of Niculescu et al. *Gbe1* was found to be overexpressed which is inline with my study in A/J and C57Bl6 mice. Genes *Pygb* and *Pygl* of the module catalyze the synthesis of glycogen. The elevated expression of *Pygb* was found for all three mice (Figure 3.8). The expression of glycogen *Pygl* was found to be downregulated in C57Bl6 and PWD mice. Sun et al., (2012) showed that mice with liver-specific deletion of histone deacetylase 3 display severe hepatosteatosis and elevation of the *Pygl* expression which is not inline with the expression of *Pygl* observed in my study, but its inline with the expression of Pygb in my study. Expression data of glycogen module is perturbed in my study as well as in the aforementioned study. These findings raise the possibility for further experimental investigation of the flux through the glycogen module.

D-alanine metabolism. The expression of *Slc16a7* was found to be upregulated for all three mose strains under DDC-treatment in the module of D-alanine metabolism (Figure 3.4e & 3.5). The regulation of *Slc16a7* expression in the study of Jansen et al., (2009) is inline with my study for all three strains. Although supplied with sufficient monocarboxylate-derived energy, glucose-deprived cleavage stage mouse embryos undergo oxidative stress and exhibit elevated reactive oxygen species which cause elevation of PPARalpha and *Slc16a7* that are markers of peroxisome proliferation. Thus, this module might be linked to oxidative stress.

Cholesterol metabolic module. This module is a part of cholesterol biosynthesis (Figure 3.4f). The expression of *Dhcr24* in the module is decreased after bariatric surgery in obese subjects with type 2 diabetes (Berisha et al., 2011). Obesity and type 2 diabetes are know to be an important factors for NAFLD (Youssef and McCullough, 2002). Based on aforementioned studies one can speculate the overexpression of *Dhcr24* may cause NAFLD and interestingly *Dhcr24* is found to be overexpressed for all three mice in my study. The overexpression of *Dhcr24* is a consequence of oxidative stress (Wu et al., 2004). Thus, this module might be linked to fatty liver disease as well as to oxidative stress.

Pyrimidine catabolism. This module is related to pyrimidine metabolism which may play an important role in liver. The expression of gene *Dpyd* was found to be elevated in the study of Yoon et al., (2011) in subject of HCC patients, while its expression in our study was found to be downregulated for all three mouse strains. Furthermore, hepatic steatosis and plasma dyslipidemia is induced by a high sucrose diet in rats and is reversed by the acute leptin administration because leptin plays a major role in weight loss and in regulation of fat deposition (Sharma et al., 2010). The expression of *Dpyd* gene was found to be downregulated in HepG2 cells treated with baicalein which is hepatoprotective against liver injury in mice (Qin et al., 2012). The expression of *Dpyd* is downregulated in both studies of Sharma et al. and Qin et al. which is in disagreement with the expression for all the mice of my analysis. It seems to be pyrimidine catabolism is deregulated differently in my study

compared to other studies such as Sharma et al and Qin et al. This might occur because in my study molecular data were derived from mouse strains while in Qin et al. it was derived from human HepG2 cells.

Cysteine metabolic module. This module comprises a part of the L-cysteine and L-methionine synthesis pathway (Figure 3.4h). *Cbs* deficiency causes severe hyperhomocysteinemia that is characterized by an abnormally high level of homocystein in the blood and is a consequence of impaired methionine/cysteine metabolism. Hyperhomocysteinemia in liver of mice is induced by *Cbs* deficiency, that promotes oxidative stress, fibrosis and steatosis (Namekata et al., 2004; Robert et al., 2005). In my analysis, *Cbs* was found to be downregulated for A/J and C57Bl6 mice, which may lead to steatosis and oxidative stress that fulfill partially the two-hit theory of NAFLD. Mice fed a high fat-diet display hyperglycemia, hyperinsulinemia, reduced glucose tolerance and hepatic triacylglycerol accumulation, while the *Cth* expression did not change (Dahlhoff et al., 2013) which is inline with the C57Bl6 mice in my analysis. Furthermore, Ames dwarf mice showed mitochondrial oxidant generation and oxidative damage (Brown-Borg et al., 2001) and the expression of *Cth* is upregulated in the Ames dwarf mouse relative to wild-type mice, which is inline with the PWD mice in my study.

Pentose metabolic module. The nonoxidative phase of pentose phosphate pathway (PPP) is identified as an important module that may be involved in liver disease progression (Figure 3.4i). In order to clarify metabolic pathways involved in hepatic lipoapoptosis Noguchi et al. (2009) performed a study using metabolic flux analysis in H4IIEC3 hepatoma cells treated with either palmitate alone (PA) or both palmitate and oleate (PA/OA) in combination (Noguchi et al., 2009). It showed that PA induces the decoupling of glycolysis and TCA cycle fluxes which leads to the onset of apoptosis, whereas PA/OA treatment restores fluxes to normal levels which results in lipid accumulation and prevents apoptosis (Noguchi et al., 2009). Furthermore, metabolomics were measured in H4IIEC3 hepatoma cells that developed lipoapoptosis and steatosis (Noguchi et al., 2011). The correlation networks of metabolites demonstrate that PA treated cells induce a dissociation among PPP metabolites, whereas PA/OA treated cells restore this dissociation (Noguchi et al., 2011). Furthermore, abnormal fluxes of PPP along with increased adenosine level, may suggest the decoupling of glycolysis, resulting in lipoapoptosis (Noguchi et al., 2011). The elevated expression of *Tkt* of this module may alter the fluxes of PPP module to an abnormal level along with the elevated combined concentration PPP components ribose 5-phosphate and ribulose phosphate in A/J and C57Bl6 mice, may indicate apoptosis which is inline with the study of Noguchi et al.. However, aforementioned combined concentration was found to be elevated in PWD mice which is not inline with the apoptosis phenotype (section 3.1.1).

4.2 Inclusion of gene expression into flux balance analysis

In order to understand the regulation of fluxes I integrated liver-specific gene expression data into a metabolic model (section 2.1) using the methods mGX-FBA and E-Flux (section 2.3 & 2.4). These methods can be used to explore metabolic states (which is relatively difficult to measure) from gene expression state (which is relatively simple to measure in many conditions). I applied these methods to gene expression datasets containing two different states (control and DDC-treatment) for three mouse strains (A/J, C57Bl6 and PWD). After integration of aforementioned datasets I identified different degrees of flux regulation among mouse strains through glucose metabolism such as glycolysis, gluconeogenesis, glycogen synthesis and degradation, TCA cycle, and pentose phosphate pathway along with lipid metabolism such as carnitine shuttle system, cholesterol biosynthesis, bile synthesis and fatty acid activation. Perturbation in experimental metabolic concentration may be expected due to the flux through the glycolysis and gluconeogenesis was found to be perturbed my *in silico* analysis. The flux through the pentose phosphate pathway and experimental concentrations of ribose 5-phosphate and ribulose 5-phosphate together were found to be perturbed. Furthermore, desmosterol is a downstream component of cholesterol biosynthesis and the flux through the pathway was found to be perturbed (section 3.3).

Challenges in the integration of expression data. Both methods, integrating expression profiles with the constraint-based metabolic models, depend on the assumption that expression profiles are a strong indicator for the amount of protein activity. For instance, E-Flux method assumes that the expression profile indicates the amount of available enzyme which can be proportional to maximum flux (e.g. Vmax). Hence, expression profiles approximate as an *upper bound* on the potential flux through the corresponding reactions. mGX-FBA method uses relative levels (between two conditions) of expression profiles within the rigid framework of metabolic constraints for optimization problem. However, some studies acknowledge that the validity of this assumption, which correlates mRNA expression to protein levels, is highly questionable. Using the Pearson product moment correlation coefficient, a measure of linear dependence between two variables, the correlation coefficient of mRNA and protein expression was estimated to be 0.356 for yeast data (Gygi et al., 1999). Furthermore, yeast cells were grown on galactose and ethanol and mRNA and protein level were measured (Griffin and Vidal-Puig, 2008). Using Spearman rank correlation method that is a non-parametric approach, the correlation of ratios (galactose *vs.* ethanol) between mRNA and protein level was found to be 0.21 for yeast data (Griffin and Vidal-Puig, 2008). Another study indicated that the correlation coefficient was 0.45 between mRNA and protein expression for yeast data using Spearman rank correlation method (Washburn et al., 2003). Furthermore, in mice, the average correlation coefficient of transcripts and proteins was calculated to be 0.27 using a R package, and the authors argue that the coefficient is varying depending on the cellular location and biological function of the gene (Ghazalpour et al., 2011). Taken together, these studies suggest that only a modest

correlation between transcripts and proteins can be found in yeast and mammals. A possible reason of weak correlation is the presence of many regulatory levels in cells after gene transcription, such as, post-transcriptional, translational, post-translational regulation and metabolite feedback regulation (Zhang et al., 2010). These studies acknowledge a strong discrepancy between studies where expression profile was used to determine protein level or upper bound on fluxes. However, integrating expression data into flux balance models improves simple FBA, which can benefit the understanding of the cellular metabolism on large scale.

Cyclic adenosine monophosphate (cAMP) signalling and transcriptional regulator of gluconeogenesis

The flux through glycolysis and gluconeogenesis as well as experimental glucose concentration were found to be perturbed in all three mouse strains (section 3.3, Figure 3.12). This flux can be affected by transcriptional regulator of gluconeogenesis that is interacted with cAMP signalling. The interaction network of cAMP signalling and transcriptional regulator of glycolysis and gluconeogenesis is shown in Figure 4.1a. Ratio (DDC-treated vs control) of cAMP concentrations and expression of genes that participate in the interaction network is shown in Figure 4.1b. The concentration of cAMP was found to upregulated for all three mouse strains due to DDC-treatment (Figure 4.1b). cAMP was identified as a second messenger that acts to transmit signals from a receptor to a target. It is involved in many biological processes, e.g., the cellular response to many hormones and neurotransmitters (Fimia and Sassone-Corsi, 2001).

The cAMP level is regulated by the activity of the enzyme adenylyl cyclase (AC) that is encoded by a large number of genes. The expression patterns of these genes and their mechanisms of regulation can differ (Fimia and Sassone-Corsi, 2001). Protein kinase A (PKA) was identified as a target of cAMP. In turn, many transcription factors (e.g. CREB and ATF-1) are regulated by PKA which is mainly achieved by direct phosphorylation of transcription factors. The phosphorylated PKA activates CREB and ATF-1 that were originally identified as the activators. A promoter analysis revealed that the CREB directly activates the transcription of gluconeogenic genes including PEPCK, G6Pase, Fbpase, and PC (Fimia and Sassone-Corsi, 2001; Hanson and Reshef, 1997).

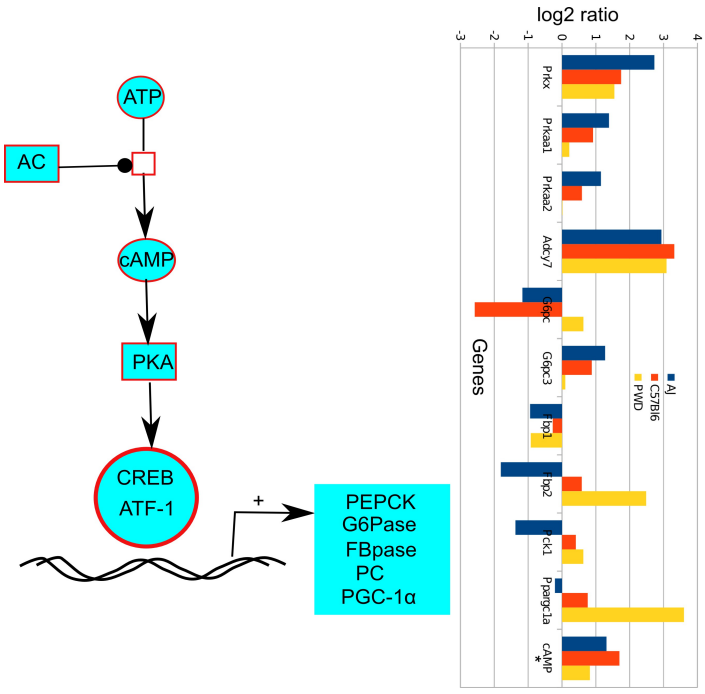

Figure 4.1. cAMP signalling network and its genes. a) AC, adenylyl cyclase; PKA, protein kinase A; CREB, cAMP response element binding protein; ATF-1, A-kinase anchoring protein; PEPCK, phosphoenolpyruvate carboxykinase; G6Pase, glucose-6-phosphatase; 1,6-bisphosphatase, FBPase; PC, pyruvate carboxylase; PGC-1α, peroxisome proliferator-activated receptor gamma co-activator 1 alpha. cAMP is produced by ATP by AC. Subsequently, cAMP activates PKA that further activates transcription factors CREB and ATF-1 which can lead to activation in gluconeogenesis genes, e.g., PEPCK, FBPase and G6Pase. **b)** cAMP is a metabolite (*). Enzymes are encoded by genes, such that AC, Adcy7; PKA, Prkx, Prkaa1, Prkaa2; G6Pase, G6pc, G6pc3; FBPase,Fbp1,Fbp2; PEPCK, Pck1; PGC-1α, Ppargc1a.

PKA is encoded by genes *Prkx, Prkaa1*, and *Prkaa2* that are upregulated for A/J and C57Bl6 mice (Figure 4.1b). The expression of *Prkx* is up regulated for PWD mice (Figure 4.1b). The quantity of upregulation (ratios) of these genes differs between all three mouse strains. In the study of Ryan et al., (2011) which concerns hepatic steatosis in human, the expression of cAMP-activated kinases (Prkaa1) is not differentially expressed which is inline with the expression of PWD mice in my study. An upregulation of Prkaa1 was found in A/J and C57Bl6 mice which is inline with the study of Kursawe et al., (2010) that is associated with insulin resistance and hepatic steatosis of obese adolescents, while it is not inline with the pattern observed in PWD mice. An upregulation of Prkaa1 leads to downregulation of sterol regulatory element binding protein-1 (SREBF1), because Prkaa1 regulates SREBF1 negatively (Rayner et al., 2011). Interestingly, the expression of SREBF1 is found to be downregulated in A/J (0.77-fold), C57Bl6 (0.48-fold), and PWD (0.56-fold) mice. SRBEF1 is a transcriptional regulator of enzymes acetyl-CoA carboxylase alpha (ACACA) and fatty acid Synthase (FASN) (Croft et al., 2011) that is involved in *de novo* synthesis of fatty acids.

Downstream of cAMP signalling (e.g., CREB) activates peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC-1α) that is a protein encoded by the *Ppargc1a* gene. A role of PGC-1α is the formation of lipid droplets in mice and human (Koves et al., 2013). The formation of lipid droplets is enhanced by the mRNA and protein level of fat storage-inducing transmembrane protein 1(FITM1/FIT1) which raises PGC-1α (Mormeneo et al., 2012). Furthermore, PGC-1α controls the expression of cell death-inducing DFFA-like Effector A (CIDEA) that is an important regulatory factor in adipose cell function and obesity (Hallberg et al., 2008). Ectopic expression of CIDEA induces the formation of lipid droplets in nonadipogenic cell lines. Lipid droplets are a marker of hepatic steatosis (Anderson and Borlak, 2008). The expression of *Ppargc1a* is upregulated for C57Bl6 and PWD mice, while slightly downregulated in A/J mice (Fig 4.1 b). Upregulation of *Ppargc1a* in PWD mice may be an explanation of the lipid droplets formation which is inline with its steatosis phenotypes (section 3.1.1).

A marked reduction was found in A/J mice in the expressions of *G6pc*, *Pck1* and *Fbp1* that encode G6pase, PEPCK, and fructose-1,6-phosphatase, respectively. Wang et al., (2012) proposed inhibition of gluconeogenesis induces hepatocellular carcinoma (HCC) in mouse model. Thus, A/J mice have a potential to develop HCC due to high steatohepatitis-like phenotypes observed in this strain (section 3.1). Interestingly, the expression of gluconeogenesis enzymes (*G6pc*, *Pck1* and *Fbp1*) was found to be downregulated which is inline with the study of Wang et al., (2012). However, Kuo et al., (2012) proposed that curcumin protects hepatocytes from mitochondrial dysfunction and free fatty acid-mediated hepatic lipoapoptosis by inhibiting key enzymes of gluconeogenesis, e.g., phosphoenol pyruvate carboxykinase (PEPCK) and glucose-6-phosphatase (G6pase). According to the study of Kuo et al. downregulation of gluconeogenesis can protect against apoptosis. As A/J mice show a higher degree of apoptosis (section 3.1) one can expect that the expression of *G6pc3* in A/J mice is in agreement with the study of Kuo et al., (2012), while the expression of gluconeogenesis enzymes (*G6pc*, *Pck1* and *Fbp1*) is not in agreement (Figure 4.1b).

Cholesterol biosynthesis. The flux through the reaction of cholesterol biosynthesis, LSTO1r (chstol/Cholesta-7,24-dien-3beta-ol -> ddsmsterol/ desmosterol), was found to be upregulated in C57Bl6 and PWD mice. Experimentally, the concentration of desmosterol was found to be upregulated in C57Bl6 and PWD mice. The serum desmosterol, a precursor of cholesterol in the cholesterol biosynthesis pathway, was found to be increased in the subject of NASH compared to the subject of simple steatosis and normal human liver (Simonen et al., 2013). An increased concentration of desmosterol in C57Bl6 mice was found in my study which is in agreement with the study of Simonen et al. due to C57Bl6 mice show low steatohepatitis-like phenotypes, while it is not inline with A/J and PWD mouse strains because desmosterol concentration did not change in A/J mouse strain although it shows high steatohepatitis-like phenotypes and desmosterol concentration is

elevated in PWD mouse strains although it shows unspecific steatohepatitis-like phenotypes. Cholesterol synthesis was found to be increased for a mouse model of hepatic steatosis which is induced by high-fat diets (Oosterveer et al., 2009). Due to increased cholesterol synthesis one may expect an increase in its precursor desmosterol which is inline with the expectation in PWD mice (section 3.1.1).

Carnitine shuttle system. The flux through a part of fatty acid metabolism (carnitine shuttle system) was found to be perturbed in all three mouse strains (Figure 3.13a). Due to an increase in flux through the carnitine shuttle system the production of carnitine is expected to be higher in PWD and C57Bl6 mice. Malaguarnera et al., (2010) proposed that a carnitine supplemented diet improves liver function, glucose plasma level, lipid profile, and histological manifestations of NASH. Thus, one may expect less steatohepatitis in PWD and C57Bl6 mice than in A/J mice which is inline with steatohepatitis-like phenotypes of all three mouse strains (section 3.1.1).

Fatty acid activation. The flux through the reactions FACOAL1812 and FACOAL1813 that are catalyzed by acyl-CoA synthetase long-chain family member 1 (ACSL1) was found to be upregulated in C57BL6 and PWD mice, while it was downregulated in A/J mice (Fig 3.16 c). ACSL1 converts free long-chain fatty acid to acyl-CoA, thereby plays an important role in lipid biosynthesis and fatty acid degradation. Gene expression of ACSL1 was found to be downregulated (< 2-fold) in all three mouse strains. Although the expression of ACSL1 was downregulated, the flux through the ACSL1 catalyzed reactions (FACOAL1812 and FACOAL1813 in Figure 3.16c) was found to be upregulated for C57Bl6 and PWD mice. This result suggests that the regulation of flux may not be judged based on single perturbation of enzyme. Thus, the complex interaction of metabolism plays an important role in the regulation of flux. Li et al., (2009) studied the function of ACSL1 in liver where liver-specific loss of ACSL1 decreases triacylglycerol synthesis, acyl-CoA, and beta-oxidation and alters phospholipid fatty acid composition. Beta-oxidation is the major process by which fatty acids are broken down to generate energy, especially when the glucose level is low during periods of starvation. Beta-oxidation of fatty acid occurs in both mitochondria and peroxisomes. Diet derived short-, medium-, and long-chain fatty acids are broken down in mitochondria which generates acetyl-CoA, and this pathway constitutes the major process by which fatty acids are oxidized to generate energy. Peroxisomes are involved in the beta-oxidation where very long chain fatty acids shorten and in the process H_2O_2 is produced. The first step of peroxisomal beta-oxidation process catalyzed by acyl-CoA oxidase (AOX) which converts fatty acyl-CoA to 2-trans-enoyl-CoA. Fan et al., (1998) proposed that mice deficient in AOX exhibit steatohepatitis. Diet derived fatty acids, e.g., elaidic acid and vaccenic acid that are involved in FACOAL1813 and FACOAL1812, respectively, can be oxidized to generate energy using beta-oxidation pathway. The flux through the FACOAL1813 and FACOAL1812 was found to be decreased in A/J mice indicating a

decrease in beta-oxidation which may develop steatohepatitis. The latter finding is inline with the steatohepatitis-like phenotype of A/J mice.

Bile acid synthesis. Bile acids are the end products of cholesterol utilization. Bile acids are used as signaling molecules and inflammatory agents that activate nuclear receptors and cell signaling pathways that regulate several metabolic pathways such as lipid, glucose, and energy metabolism. Bile acids synthesis is the major pathway of cholesterol catabolism in mammals. The excretion of excess cholesterol is occurred in the form of bile acid. Although several of the enzymes that are involved in synthesis of bile acid are active in many tissues, the liver is the only organ where *de novo* synthesis of bile acid takes place (Chiang, 2009). The genes encoding the enzymes of bile acid synthesis pathway should be tightly controlled in order to ensure the necessary level of bile acid production that coordinates changes of the metabolic conditions. Many bile acid metabolites are cytotoxic and due to this their synthesis needs to be tightly controlled. However, bile acid was found to be protective against the fructose-induced hepatic steatosis in mice (Volynets et al., 2010). Metabolite taurocholic acid/tchola is a yellowish crystalline bile acid, it is the end product derived from cholesterol/chsterol (Figure 3.16b). The flux through the bile acid synthesis pathway was found to be downregulated in A/J mice, and to be upregulated in PWD mice, and unchanged in C57Bl6 mice (Figure 3.16b). Levels of bile acid were found to be elevated in the liver of patients with steatohepatitis (Aranha et al., 2008) and in a murine steatohepatitis model induced by the methionine-choline-deficient (MCD) diet (Gyamfi et al., 2008). Experimental concentration of taurocholic acid was found to be upregulated in AJ mice which can be in agreement with the study of Aranha et al., (2008) and Gyamfi et al., (2008). Farnesoid X receptor (FXR), a ligand-activated nuclear receptor belonging to the nuclear receptor superfamily, is a central regulator and sensor for bile acids. FXR deficiency leds to the development of nonalcoholic steatohepatitis in low-density lipoprotein receptor-knockout mice fed a high-fat diet by altering the bile acid homeostasis (Kong et al., 2009). Perturbation of the flux through bile acid synthesis (Figure 3.16b) among mice may alter bile acid homeostasis which is inline with the study of Kong et al., (2009). Cholesterol 7α-hydroxylase (Cyp7a1) catalyzes the rate-limiting step that converts cholesterol to bile acids in the liver. After feeding a DDC supplemented diet the expression of Cyp7a1 is found to be upregulated 2.85-fold and 1.45-fold in A/J and C57Bl6 mice, respectively, and to be downregulated 0.63-fold in PWD mice. Chronic cholesterol feeding in mice results in suppression of hepatic Cyp7a1 expression that is associated with an increase in hepatic cholesterol content (Henkel et al., 2011), which is in agreement with the expression and steatosis phenotypes of PWD mice. On the contrary, Henkel et al., (2011) proposed that the acute cholesterol feeding results in upregulation of hepatic Cyp7a1 expression. Furthermore, Ha and Kim, (2013) studied the effects of fucoxanthin on gene expression related to lipid metabolism in rats with a high-fat diet. Consumption of fucoxanthin improves lipid and cholesterol metabolism in rats with a high fat diet and in addition to this it increases the

mRNA expression *Cyp7a1* that is inline with an upregulation of *Cyp7a1* expression in A/J and C57Bl6 mice with a DDC supplemented diet. FXR plays an important role in the regulation of cholesterol and glucose metabolism (Lambert et al., 2003; Ma et al., 2006). Hepatic FXR responds to bile acids and represses the *Cyp7a1* expression (Chiang et al., 2000). Histological features of liver of FXR deficient (FXR KO) mice display higher degree of steatosis, perisinusoidal/sinusoidal foam cells, ballooning degeneration and lobular inflammation compared to wild mice, which may indicate liver damage and steatohepatitis-like phenotypes in FXR KO mice (Bjursell et al., 2013).

Furthermore, Tanaka et al., (2013) proposed that restricted intake of dietary phosphate plays an important role for the development of fatty liver disease which is induced by a high-cholesterol diet and the hepatic mRNA levels of *Cyp7a1* are observed to be increased. The expression of *Cyp7a1* in A/J and C57Bl6 mice is inline with the study of Bjursell et al., (2013) and Tanaka et al., (2013).

4.3 Drug target candidates in liver disease

There are few drugs to cure the liver disease: steatosis, steatohepatitis and HCC. Thus, I have performed an in *silico* analysis in which the metabolic candidate genes are identified as potential drug targets in aforementioned liver disease (section 3.4). I have formulated biomass functions for steatosis and steatohepatitis and these biomass functions are used to identify potential drug targets of diseases. 26 drug targets were found to be common in steatosis, steatohepatitis, and HCC (Table 3.1). Out of 26 drug targets, 13 were identified in cholesterol metabolism, while 6 and 3 were triacylglycerol synthesis and inositol phosphate metabolism, respectively. Folger et al., (2011) performed an *in silico* drug target analysis in cancer using human metabolic network and found that a marked number of drug targets is from cholesterol metabolism which is inline with my study. Interestingly, there are some common targets such as *Fdps*, *Hmgcr*, *Lss.* Folger et al. also recognized *NSDHL* and *SOAT1* as potential drug targets, which were not identified in my work. On the other hand in this study I identified potential drug targets such as *Agpat1* and *Agpat2* form triacylglycerol synthesis along with potential drug targets such as *Isyna1* and *Impa1* from inositol phosphate metabolism which were not found in the study of Folger et al.. Furthermore, *Acaca* is a key enzyme gene of fatty acid metabolism and this is identified as a potential drug target in my study (steatosis and steatohepatitis but not in HCC) as well as in the study of Folger et al.. Moreover, drug targets such as *Sptlc1*, *Sptlc2* and *Kdsr* were found to be from sphingolipid metabolism in the study of Folger et al. and in my study.

Fdps is involved in cholesterol metabolism and is identified as a potential candidate for drugs in liver disease. There are some drugs such as pamidronate, zoledronate, alendronate, ibandronate, and risedronate which target *Fdps* (Ahmed et al., 2011). Zein et al., (2005)

proposed that alendronate improves bone mineral density in primary biliary cirrhosis (PBC). PBC is an autoimmune disease characterized by inflammation of small bile ducts, and in some patients the development of fibrosis and cirrhosis (Hirschfield and Gershwin, 2013). Furthermore, another candidate of cholesterol metabolism, *Hmgcr*, was identified as a potential candidate for drugs such as NADH, pravastatin, lovastatin (Ahmed et al., 2011). Statins such as pravastatin can be used safely in patients with NAFLD (Calderon et al., 2010). Gene *Sc4mol* is another candidate of cholesterol metabolism which can be targeted by drugs NADH and indomethacin (Ahmed et al., 2011). Murali et al.,(2012) proposed that fish oil and indomethacin in combination potently reduce dyslipidemia and hepatic steatosis in mice. Taken together, there are some drugs for genes of cholesterol metabolism such as *Hmgcr*, *Fdps*, *Sc4mol* which might be used for liver disease such as steatosis, steatohepatitis, and HCC but of course this needs an experimental validation. Some other candidates of cholesterol metabolism such as *Mvd* and *Lss* were identified as a target in this study. These candidates are novel and may be used as a potential drug targets in liver disease but this need further experimental investigation.

L-serine is a drug which can inhibit *Sptlc1*, *Sptlc2* (Ahmed et al., 2011). L-serine is synthesized from glycine or threonine. L-Serine is non-essential and it can also be derived from the diet and the degradation of protein and/or phospholipids (Mardinoglu et al., 2014). L-serine is involved in the formation of lipids such as sphingosine and phosphatidylserine as well as other building blocks and cofactors, such as protein (glycine and cysteine), creatine, porphyrins, glutathione and nucleotides (Mardinoglu et al., 2014). The expression of *Sptlc1* in the study of Mardinoglu et al., (2014) were found to be upregulated in NASH patient which is inline with expression pattern, observed in A/J and C57Bl6 mouse strains in my study (section 3.1.1). Furthermore, Mardinoglu et al. identified phosphoserine phosphatase (*PSPH*) as potential therapeutic targets for the treatment of NASH which is inline with my study where *Psph* (Table 3.1) is identified for the treatment of HCC. The amplification of *PSPH* in the serine synthesis pathway may also decrease the flux through pyruvate and lactate formation in cytosol (Mardinoglu et al., 2014). The increased pyruvate and lactate levels were previously reported in NASH patients (Kalhan et al., 2011b). The experimentally measured concentrations of pyruvate together with oxaloacetate were found to be upregulated (> 1.5-fold) in mouse strain A/J and C57Bl6 which is inline with the study of Kalhan et al., (2011b). Taken together, both results of my work along with the work of Mardinoglu et al., (2014) indicate that one needs to validate experimentally *PSPH* as potential therapeutic targets for the treatment of NASH. Moreover, Mardinoglu et al. identified *SHMT1* and *BCAT1* as potential therapeutic targets for the treatment of NASH, while these targets are not identified in my *in silico* drug target analysis it might be due to use of different metabolic models in both studies.

4.4 Arachidonic acid metabolism

In pathway analysis AA metabolism was identified as being the most affected due to DDC treatment in all of three mouse strains. I proposed a dynamic model of AA metabolism to study the DDC treatment effects. The fitted model provides good correspondence between the experimental and simulated metabolic data using the respective gene expression data.

To determine regulatory roles of important enzymes in the upregulation of metabolic levels of the AA/eicosanoid pathway due to DDC-treatment, a sensitivity analysis was performed by reverting to the activity of the enzymes at normal, untreated conditions. The combination of the enzymes ALOX5AP, PERK, PKCD, STAT3, and PTGS1 is able to revert the DDC-treated metabolic state to normal. This multi-target inhibition might likely become relevant in the context of individualized medicine. A study suggests that multi-target cancer therapy has higher probability of success than single-target, because it can better reflect the diversity in genetics among individuals (Broekman et al., 2011; Lu et al., 2012).

Furthermore, I performed *in silico* drug testing of several model enzymes. Inhibition of PTGS1 and ALOX5AP caused downregulation of PGH2 and PGD2 or 5-HPETE and LTA4, suggesting that PTGS1 and ALOX5AP might be potential drug targets of steatohepatitis. PTGS1 (COX-1), for instance, can be inhibited by mofezolac, SC-560, and others (Perrone et al., 2010). Inhibition of ALOX5AP (also known as 5-lipoxygenase activating protein or FLAP) may be useful in the prevention of hepatotoxin-induced necro-inflammatory injury (Titos et al., 2005). Drug molecules can interact with multiple targets to alter the state and function of the associated biological network. Licofelone is a novel 5-LOX/Cox-inhibitor which inhibits two enzymes to avoid side effects (Ulbrich et al., 2005). Overall, the dynamics of the AA model can be used for *in silico* drug studies of multiple drugs and potential drug targets. AA model can confirm the upregulation of PGD2 due to DDC treatment found experimentally. It is connected with the key transcription factors/ligand-activated nuclear receptors such as PPARδ which was implicated as a key regulator of energy homeostasis and may represent future research avenues to study the interaction of metabolic and signaling pathways (Wang and DuBois, 2010).

The spectrum of NAFLD and NASH can be characterized by specific alterations in hepatic lipid composition. A comprehensive analysis of plasma lipids and eicosanoids in human revealed a stepwise increase in lipoxygenase metabolites 5-HETE, 8-HETE and 15-HETE in NAFLD and NASH (Puri et al., 2009). This correlates with our observations in the experimental and simulated data, where concentrations of 5-HETE and 15-HETE were increased in DDC mice. Another study reported overexpression of cyclooxygenase-2 (COX-2) in hepatocellular carcinoma (HCC) patients (Giannitrapani et al., 2009). Using immunohistochemistry they studied COX-2 overexpression in different chronic liver diseases including NASH, chronic hepatitis, and liver cirrhosis. Here, I found overexpression of PTGS1 (COX-2) in DDC mice, which is a key regulator of prostaglandin formation.

Martínez-Clemente et al., (2010) demonstrated that hyperlipidemia-prone apolipoprotein E-deficient (ApoE(-/-)) mice exhibit hepatic steatosis and increased susceptibility to hepatic inflammation and advanced fibrosis. They experimentally found that the proinflammatory 5-lipoxygenase (5-LO) pathway is upregulated and thus causes liver inflammation and fibrogenesis. They also found that the inhibition of the 5-LO pathway results in a significant reduction in liver inflammation. Our data supports an upregulation of ALOX5AP through 5-LO pathway due to DDC treatment in A/J, C57Bl6, and PWD, leading to an upregulation of the downstream component 5-HPETE in the model that is supported by our experimental data.

In conclusion, considering mRNA expression data for mathematical modeling of metabolic systems provides a very useful way to understand cellular metabolism, although the correlation between transcripts and proteins can deviate depending on cellular location, biological function, and organism (Ghazalpour et al., 2011). I identified AA/eicosanoid metabolism as highly perturbed in DDC-induced mice using a combination of *in silico* and experimental approaches. The analysis of the AA/eicosanoid metabolic pathway presented in this work suggests that 5-HETE, 15-HETE and PGD2 are perturbed in DDC mice. Therefore this work demonstrates that a dynamic model can be used for qualitative prediction of metabolic changes based on transcriptomics data in a disease-related context.

Bibliography

Adams, L.A., Lymp, J.F., St Sauver, J., Sanderson, S.O., Lindor, K.D., Feldstein, A., and Angulo, P. (2005). The natural history of nonalcoholic fatty liver disease: a population-based cohort study. Gastroenterology 129, 113–121.

Aharony, D., Redkar-Brown, D.G., Hubbs, S.J., and Stein, R.L. (1987). Kinetic studies on the inactivation of 5-lipoxygenase by 5(S)-hydroperoxyeicosatetraenoic acid. Prostaglandins *33*, 85–100.

Ahmed, J., Meinel, T., Dunkel, M., Murgueitio, M.S., Adams, R., Blasse, C., Eckert, A., Preissner, S., and Preissner, R. (2011). CancerResource: a comprehensive database of cancer-relevant proteins and compound interactions supported by experimental knowledge. Nucleic Acids Res. *39*, D960–967.

Ahn, Y.-Y., Bagrow, J.P., and Lehmann, S. (2010). Link communities reveal multiscale complexity in networks. Nature *466*, 761–764.

Alba, L.M., and Lindor, K. (2003). Review article: Non-alcoholic fatty liver disease. Aliment. Pharmacol. Ther. *17*, 977–986.

Albert, R. (2007). Network Inference, Analysis, and Modeling in Systems Biology. Plant Cell Online *19*, 3327–3338.

Aldana, M., and Cluzel, P. (2003). A natural class of robust networks. Proc. Natl. Acad. Sci. *100*, 8710–8714.

Amarapurkar, D., Kamani, P., Patel, N., Gupte, P., Kumar, P., Agal, S., Baijal, R., Lala, S., Chaudhary, D., and Deshpande, A. (2007). Prevalence of non-alcoholic fatty liver disease: population based study. Ann. Hepatol. *6*, 161–163.

Amberger, J., Bocchini, C.A., Scott, A.F., and Hamosh, A. (2009). McKusick's Online Mendelian Inheritance in Man (OMIM). Nucleic Acids Res. *37*, D793–796.

Anderson, N., and Borlak, J. (2008). Molecular Mechanisms and Therapeutic Targets in Steatosis and Steatohepatitis. Pharmacol. Rev. *60*, 311–357.

Anfuso, C.D., Lupo, G., Romeo, L., Giurdanella, G., Motta, C., Pascale, A., Tirolo, C., Marchetti, B., and Alberghina, M. (2007). Endothelial cell-pericyte cocultures induce PLA2 protein expression through activation of PKCalpha and the MAPK/ERK cascade. J. Lipid Res. *48*, 782–793.

Angulo, P. (2002). Nonalcoholic fatty liver disease. N. Engl. J. Med. *346*, 1221–1231.

Anstee, Q.M., and Goldin, R.D. (2006). Mouse models in non-alcoholic fatty liver disease and steatohepatitis research. Int. J. Exp. Pathol. *87*, 1–16.

Aranha, M.M., Cortez-Pinto, H., Costa, A., da Silva, I.B.M., Camilo, M.E., de Moura, M.C., and Rodrigues, C.M.P. (2008). Bile acid levels are increased in the liver of patients with steatohepatitis. Eur. J. Gastroenterol. Hepatol. *20*, 519–525.

Avigan, J., Steinberg, D., Gutman, A., Mize, C.E., and Milne, G.W.A. (1966). Alpha-decarboxylation, an important pathway for degradation of phytanic acid in animals. Biochem. Biophys. Res. Commun. *24*, 838–844.

Bartz, R., Li, W.-H., Venables, B., Zehmer, J.K., Roth, M.R., Welti, R., Anderson, R.G.W., Liu, P., and Chapman, K.D. (2007). Lipidomics reveals that adiposomes store ether lipids and mediate phospholipid traffic. J. Lipid Res. *48*, 837–847.

Becker, S.A., and Palsson, B.O. (2008). Context-Specific Metabolic Networks Are Consistent with Experiments. PLoS Comput Biol *4*, e1000082.

Becker, S.A., Feist, A.M., Mo, M.L., Hannum, G., Palsson, B.Ø., and Herrgard, M.J. (2007). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. Nat. Protoc. *2*, 727–738.

Bedogni, G., Miglioli, L., Masutti, F., Tiribelli, C., Marchesini, G., and Bellentani, S. (2005). Prevalence of and risk factors for nonalcoholic fatty liver disease: The Dionysos nutrition and liver study. Hepatology *42*, 44–52.

Begriche, K., Massart, J., Abbey-Toby, A., Igoudjil, A., Lettéron, P., and Fromenty, B. (2008). Beta-aminoisobutyric acid prevents diet-induced obesity in mice with partial leptin deficiency. Obes. Silver Spring Md *16*, 2053–2067.

Berisha, S.Z., Serre, D., Schauer, P., Kashyap, S.R., and Smith, J.D. (2011). Changes in Whole Blood Gene Expression in Obese Subjects with Type 2 Diabetes Following Bariatric Surgery: a Pilot Study. PLoS ONE *6*, e16729.

Bhattacharjee, A., Xu, B., Frank, D.A., Feldman, G.M., and Cathcart, M.K. (2006). Monocyte 15-Lipoxygenase Expression Is Regulated by a Novel Cytosolic Signaling Complex with Protein Kinase C δ and Tyrosine-Phosphorylated Stat3. J. Immunol. *177*, 3771–3781.

Bjursell, M., Wedin, M., Admyre, T., Hermansson, M., Bottcher, G., Goransson, M., Linden, D., Bamberg, K., Oscarsson, J., and Bohlooly-Y, M. (2013). Ageing Fxr Deficient Mice Develop Increased Energy Expenditure, Improved Glucose Control and Liver Damage Resembling NASH. PLoS ONE *8*.

Bray, G.A., Nielsen, S.J., and Popkin, B.M. (2004). Consumption of high-fructose corn syrup in beverages may play a role in the epidemic of obesity. Am. J. Clin. Nutr. *79*, 537–543.

Brink, D.M. van den, and Wanders, R.J.A. (2006). Phytanic acid: production from phytol, its breakdown and role in human disease. Cell. Mol. Life Sci. *63*, 1752–1765.

Broekman, F., Giovannetti, E., and Peters, G.J. (2011). Tyrosine kinase inhibitors: Multi-targeted or single-targeted? World J. Clin. Oncol. *2*, 80–93.

Brown-Borg, D.H., Johnson, W.T., Rakoczy, S., and Romanick, M. (2001). Mitochondrial oxidant generation and oxidative damage in Ames dwarf and GH transgenic mice. J. Am. Aging Assoc. *24*, 85–96.

Browning, J.D., Szczepaniak, L.S., Dobbins, R., Nuremberg, P., Horton, J.D., Cohen, J.C., Grundy, S.M., and Hobbs, H.H. (2004). Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. Hepatol. Baltim. Md *40*, 1387–1395.

Bugianesi, E., Leone, N., Vanni, E., Marchesini, G., Brunello, F., Carucci, P., Musso, A., De Paolis, P., Capussotti, L., Salizzoni, M., et al. (2002). Expanding the natural history of nonalcoholic steatohepatitis: from cryptogenic cirrhosis to hepatocellular carcinoma. Gastroenterology *123*, 134–140.

Burgard, A.P., Nikolaev, E.V., Schilling, C.H., and Maranas, C.D. (2004). Flux coupling analysis of genome-scale metabolic network reconstructions. Genome Res. *14*, 301–312.

Butcher, E.C., Berg, E.L., and Kunkel, E.J. (2004). Systems biology in drug discovery. Nat. Biotechnol. *22*, 1253–1259.

Calderon, R.M., Cubeddu, L.X., Goldberg, R.B., and Schiff, E.R. (2010). Statins in the Treatment of Dyslipidemia in the Presence of Elevated Liver Aminotransferase Levels: A Therapeutic Dilemma. Mayo Clin. Proc. *85*, 349–356.

Canbay, A., Bechmann, L., and Gerken, G. (2007). Lipid metabolism in the liver. Z. Für Gastroenterol. *45*, 35–41.

Cano, A., Buqué, X., Martínez-Uña, M., Aurrekoetxea, I., Menor, A., García-Rodríguez, J.L., Lu, S.C., Martínez-Chantar, M.L., Mato, J.M., Ochoa, B., et al. (2011). Methionine adenosyltransferase 1A gene deletion disrupts hepatic very low-density lipoprotein assembly in mice. Hepatol. Baltim. Md *54*, 1975–1986.

Cha, J.-Y., and Repa, J.J. (2007). The liver X receptor (LXR) and hepatic lipogenesis. The carbohydrate-response element-binding protein is a target gene of LXR. J. Biol. Chem. *282*, 743–751.

Chaturvedi, R.K., Calingasan, N.Y., Yang, L., Hennessey, T., Johri, A., and Beal, M.F. (2010). Impairment of PGC-1alpha expression, neuropathology and hepatic steatosis in a transgenic mouse model of Huntington's disease following chronic energy deprivation. Hum. Mol. Genet. *19*, 3190–3205.

Cheang, M.C.U., van de Rijn, M., and Nielsen, T.O. (2008). Gene Expression Profiling of Breast Cancer. Annu. Rev. Pathol. Mech. Dis. *3*, 67–97.

Chen, L., and Wu, J. (2012). Systems biology for complex diseases. J. Mol. Cell Biol. *4*, 125–126.

Chen, J., Aronow, B.J., and Jegga, A.G. (2009). Disease candidate gene identification and prioritization using protein interaction networks. BMC Bioinformatics *10*, 73.

Chiang, J.Y.L. (2009). Bile acids: regulation of synthesis. J. Lipid Res. *50*, 1955–1966.

Chiang, J.Y.L., Kimmel, R., Weinberger, C., and Stroup, D. (2000). Farnesoid X Receptor Responds to Bile Acids and Represses Cholesterol 7α-Hydroxylase Gene (CYP7A1) Transcription. J. Biol. Chem. *275*, 10918–10924.

Chuang, H.-Y., Hofree, M., and Ideker, T. (2010). A Decade of Systems Biology. Annu. Rev. Cell Dev. Biol. *26*, 721–744.

Clària, J., and Planagumà, A. (2005). Liver: The formation and actions of aspirin-triggered lipoxins. Prostaglandins Leukot. Essent. Fatty Acids *73*, 277–282.

Clarke, J.D., Novak, P., Lake, A.D., Shipkova, P., Aranibar, N., Robertson, D., Severson, P.L., Reily, M.D., Futscher, B.W., Lehman-McKeeman, L.D., et al. (2013). Characterization of Hepatocellular Carcinoma Related Genes and Metabolites in Human Nonalcoholic Fatty Liver Disease. Dig. Dis. Sci.

Cohen, J.C., Horton, J.D., and Hobbs, H.H. (2011). Human Fatty Liver Disease: Old Questions and

New Insights. Science *332*, 1519–1523.

Colijn, C., Brandes, A., Zucker, J., Lun, D.S., Weiner, B., Farhat, M.R., Cheng, T.-Y., Moody, D.B., Murray, M., and Galagan, J.E. (2009). Interpreting Expression Data with Metabolic Flux Models: Predicting Mycobacterium tuberculosis Mycolic Acid Production. PLoS Comput Biol *5*, e1000489.

Crespo, J., Cayón, A., Fernández-Gil, P., Hernández-Guerra, M., Mayorga, M., Domínguez-Díez, A., Fernández-Escalante, J.C., and Pons-Romero, F. (2001). Gene expression of tumor necrosis factor alpha and TNF-receptors, p55 and p75, in nonalcoholic steatohepatitis patients. Hepatol. Baltim. Md *34*, 1158–1163.

Croft, D., O'Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., Caudy, M., Garapati, P., Gopinath, G., Jassal, B., et al. (2011). Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res. *39*, D691–697.

Cunningham, C.C., and Van Horn, C.G. (2003). Energy availability and alcohol-related liver pathology. Alcohol Res. Health J. Natl. Inst. Alcohol Abuse Alcohol. *27*, 291–299.

Dahlhoff, C., Desmarchelier, C., Sailer, M., Fürst, R.W., Haag, A., Ulbrich, S.E., Hummel, B., Obeid, R., Geisel, J., Bader, B.L., et al. (2013). Hepatic Methionine Homeostasis Is Conserved in C57BL/6N Mice on High-Fat Diet Despite Major Changes in Hepatic One-Carbon Metabolism. PLoS ONE *8*, e57387.

David, L., Marashi, S.-A., Larhlimi, A., Mieth, B., and Bockmayr, A. (2011). FFCA: a feasibility-based method for flux coupling analysis of metabolic networks. BMC Bioinformatics *12*, 236.

Day, C.P. (2005). Natural history of NAFLD: remarkably benign in the absence of cirrhosis. Gastroenterology *129*, 375–378.

Day, C.P., and James, O.F. (1998). Steatohepatitis: a tale of two "hits"? Gastroenterology *114*, 842–845.

De Figueiredo, L.F., Podhorski, A., Rubio, A., Kaleta, C., Beasley, J.E., Schuster, S., and Planes, F.J. (2009). Computing the shortest elementary flux modes in genome-scale metabolic networks. Bioinforma. Oxf. Engl. *25*, 3158–3165.

Delerive, P., Fruchart, J.C., and Staels, B. (2001). Peroxisome proliferator-activated receptors in inflammation control. J. Endocrinol. *169*, 453–459.

De Medina, P., Silvente-Poirot, S., and Poirot, M. (2009). Tamoxifen and AEBS ligands induced apoptosis and autophagy in breast cancer cells through the stimulation of sterol accumulation. Autophagy *5*, 1066–1067.

Denk, H., Stumptner, C., and Zatloukal, K. (2000). Mallory bodies revisited. J. Hepatol. *32*, 689–702.

Dettmer, K., Stevens, A., Gronwald, W., Hellerbrand, C., and Oefner, P. (2011). Mass spectrometry based metabolomics to study the development of Non-alcoholic steatohepatitis. Z. Für Gastroenterol. *49*.

Deutscher, D., Meilijson, I., Kupiec, M., and Ruppin, E. (2006). Multiple knockout analysis of genetic robustness in the yeast metabolic network. Nat. Genet. *38*, 993–998.

Dierkes, T., Wade, M., Nowak, U., and Röblitz, S. (2011). BioPARKIN - Biology-related Parameter

Identification in Large Kinetic Networks. 11–15.

Doi, A., Nagasaki, M., Matsuno, H., and Miyano, S. (2011). Simulation-Based Validation of the p53 Transcriptional Activity with Hybrid Functional Petri Net. Stud. Health Technol. Inform. *162*, 130–142.

Domenici, F.A., Brochado, M.J.F., Martinelli, A. de L.C., Zucoloto, S., da Cunha, S.F. de C., and Vannucchi, H. (2013). Peroxisome proliferator-activated receptors alpha and gamma2 polymorphisms in nonalcoholic fatty liver disease: A study in Brazilian patients. Gene *529*, 326–331.

Dowman, J.K., Tomlinson, J.W., and Newsome, P.N. (2011). Systematic review: the diagnosis and staging of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis. *33*, 525–540.

Van Driel, M.A., and Brunner, H.G. (2006). Bioinformatics methods for identifying candidate disease genes. Hum. Genomics *2*, 429–432.

Du, D., Shi, Y.-H., and Le, G.-W. (2010). Oxidative stress induced by high-glucose diet in liver of C57BL/6J mice and its underlying mechanism. Mol. Biol. Rep. *37*, 3833–3839.

Duarte, N.C., Becker, S.A., Jamshidi, N., Thiele, I., Mo, M.L., Vo, T.D., Srivas, R., and Palsson, B.Ø. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proc. Natl. Acad. Sci. *104*, 1777–1782.

Eberlé, D., Hegarty, B., Bossard, P., Ferré, P., and Foufelle, F. (2004). SREBP transcription factors: master regulators of lipid homeostasis. Biochimie *86*, 839–848.

Ekstedt, M., Franzén, L.E., Mathiesen, U.L., Holmqvist, M., Bodemar, G., and Kechagias, S. (2007). Statins in non-alcoholic fatty liver disease and chronically elevated liver enzymes: a histopathological follow-up study. J. Hepatol. *47*, 135–141.

Ellinghaus, P., Wolfrum, C., Assmann, G., Spener, F., and Seedorf, U. (1999). Phytanic acid activates the peroxisome proliferator-activated receptor alpha (PPARalpha) in sterol carrier protein 2-/ sterol carrier protein x-deficient mice. J. Biol. Chem. *274*, 2766–2772.

Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. Nature *403*, 335–338.

Fan, C.Y., Pan, J., Usuda, N., Yeldandi, A.V., Rao, M.S., and Reddy, J.K. (1998). Steatohepatitis, spontaneous peroxisome proliferation and liver tumors in mice lacking peroxisomal fatty acyl-CoA oxidase. Implications for peroxisome proliferator-activated receptor alpha natural ligand metabolism. J. Biol. Chem. *273*, 15639–15645.

Farrell, G.C., and Larter, C.Z. (2006). Nonalcoholic fatty liver disease: from steatosis to cirrhosis. Hepatol. Baltim. Md *43*, S99–S112.

Feldstein, A.E., Werneburg, N.W., Canbay, A., Guicciardi, M.E., Bronk, S.F., Rydzewski, R., Burgart, L.J., and Gores, G.J. (2004). Free fatty acids promote hepatic lipotoxicity by stimulating TNF-alpha expression via a lysosomal pathway. Hepatol. Baltim. Md *40*, 185–194.

Fell, D.A., and Thomas, S. (1995). Physiological control of metabolic flux: the requirement for multisite modulation. Biochem. J. *311*, 35–39.

Fimia, G.M., and Sassone-Corsi, P. (2001). Cyclic AMP signalling. J. Cell Sci. *114*, 1971–1972.

Fischer, E., and Sauer, U. (2003). A novel metabolic cycle catalyzes glucose oxidation and anaplerosis in hungry Escherichia coli. J. Biol. Chem. *278*, 46446–46451.

Folger, O., Jerby, L., Frezza, C., Gottlieb, E., Ruppin, E., and Shlomi, T. (2011). Predicting selective drug targets in cancer through metabolic networks. Mol. Syst. Biol. *7*, 501.

Gagneur, J., and Klamt, S. (2004). Computation of elementary modes: a unifying framework and the new binary approach. BMC Bioinformatics *5*, 175.

Gan, Q., Schones, D.E., Eun, S.H., Wei, G., Cui, K., Zhao, K., and Chen, X. (2010). Monovalent and unpoised status of most genes in undifferentiated cell-enriched Drosophila testis. Genome Biol. *11*, R42.

Garcia-Ruiz, C., and Fernandez-Checa, J.C. (2006). Mitochondrial glutathione: hepatocellular survival-death switch. J. Gastroenterol. Hepatol. *21 Suppl 3*, S3–6.

Geddis, M.S., Tornieri, K., Giesecke, A., and Rehder, V. (2004). PLA2 and secondary metabolites of arachidonic acid control filopodial behavior in neuronal growth cones. Cell Motil. Cytoskeleton *57*, 53–67.

Ghazalpour, A., Bennett, B., Petyuk, V.A., Orozco, L., Hagopian, R., Mungrue, I.N., Farber, C.R., Sinsheimer, J., Kang, H.M., Furlotte, N., et al. (2011). Comparative Analysis of Proteome and Transcriptome Variation in Mouse. PLoS Genet *7*, e1001393.

Giannitrapani, L., Ingrao, S., Soresi, M., Florena, A.M., Spada, E.L., Sandonato, L., D'Alessandro, N., Cervello, M., and Montalto, G. (2009). Cyclooxygenase-2 Expression in Chronic Liver Diseases and Hepatocellular Carcinoma. *1155*, 293–299.

Girvan, M., and Newman, M.E.J. (2002). Community structure in social and biological networks. Proc. Natl. Acad. Sci. *99*, 7821–7826.

Grattagliano, I., Portincasa, P., Palmieri, V.O., and Palasciano, G. (2007). Managing nonalcoholic fatty liver disease: recommendations for family physicians. Can. Fam. Physician Médecin Fam. Can. *53*, 857–863.

Greco, D., Kotronen, A., Westerbacka, J., Puig, O., Arkkila, P., Kiviluoto, T., Laitinen, S., Kolak, M., Fisher, R.M., Hamsten, A., et al. (2008). Gene expression in human NAFLD. Am. J. Physiol. Gastrointest. Liver Physiol. *294*, G1281–1287.

Griffin, J.L., and Vidal-Puig, A. (2008). Current challenges in metabolomics for diabetes research: a vital functional genomic tool or just a ploy for gaining funding? Physiol. Genomics *34*, 1–5.

Gulsen, M., Yesilova, Z., Bagci, S., Uygun, A., Ozcan, A., Ercin, C.N., Erdil, A., Sanisoglu, S.Y., Cakir, E., Ates, Y., et al. (2005). Elevated plasma homocysteine concentrations as a predictor of steatohepatitis in patients with non-alcoholic fatty liver disease. J. Gastroenterol. Hepatol. *20*, 1448–1455.

Guo, Y., Cordes, K.R., Farese, R.V., and Walther, T.C. (2009). Lipid droplets at a glance. J. Cell Sci. *122*, 749–752.

Gupta, S., Maurya, M.R., Stephens, D.L., Dennis, E.A., and Subramaniam, S. (2009). An integrated model of eicosanoid metabolism and signaling based on lipidomics flux analysis. Biophys. J. *96*, 4542–4551.

Gyamfi, M.A., Damjanov, I., French, S., and Wan, Y.-J.Y. (2008). The pathogenesis of ethanol versus methionine and choline deficient diet-induced liver injury. Biochem. Pharmacol. *75*, 981–995.

Gygi, S.P., Rochon, Y., Franza, B.R., and Aebersold, R. (1999). Correlation between Protein and mRNA Abundance in Yeast. Mol. Cell. Biol. *19*, 1720–1730.

Ha, A.W., and Kim, W.K. (2013). The effect of fucoxanthin rich power on the lipid metabolism in rats with a high fat diet. Nutr. Res. Pract. *7*, 287–293.

Halama, A., Riesen, N., Möller, G., Hrabě de Angelis, M., and Adamski, J. (2013). Identification of biomarkers for apoptosis in cancer cell lines using metabolomics: tools for individualized medicine. J. Intern. Med. *274*, 425–439.

Hallberg, M., Morganstein, D.L., Kiskinis, E., Shah, K., Kralli, A., Dilworth, S.M., White, R., Parker, M.G., and Christian, M. (2008). A Functional Interaction between RIP140 and PGC-1α Regulates the Expression of the Lipid Droplet Protein CIDEA. Mol. Cell. Biol. *28*, 6785–6795.

Haluzik, M., Colombo, C., Gavrilova, O., Chua, S., Wolf, N., Chen, M., Stannard, B., Dietz, K.R., Le Roith, D., and Reitman, M.L. (2004). Genetic background (C57BL/6J versus FVB/N) strongly influences the severity of diabetes and insulin resistance in ob/ob mice. Endocrinology *145*, 3258–3264.

Hanson, R.W., and Reshef, L. (1997). Regulation of Phosphoenolpyruvate Carboxykinase (gtp) Gene Expression. Annu. Rev. Biochem. *66*, 581–611.

Haque, J.A., McMahan, R.S., Campbell, J.S., Shimizu-Albergine, M., Wilson, A.M., Botta, D., Bammler, T.K., Beyer, R.P., Montine, T.J., Yeh, M.M., et al. (2010). Attenuated progression of diet-induced steatohepatitis in glutathione-deficient mice. Lab. Invest. *90*, 1704–1717.

Harano, Y., Yasui, K., Toyama, T., Nakajima, T., Mitsuyoshi, H., Mimani, M., Hirasawa, T., Itoh, Y., and Okanoue, T. (2006). Fenofibrate, a peroxisome proliferator-activated receptor α agonist, reduces hepatic steatosis and lipid peroxidation in fatty liver Shionogi mice with hereditary fatty liver. Liver Int. *26*, 613–620.

Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W. (1999). From molecular to modular cell biology. Nature *402*, C47–52.

Hayaishi, O. (1991). Molecular mechanisms of sleep-wake regulation: roles of prostaglandins D2 and E2. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. *5*, 2575–2581.

Henkel, A.S., Anderson, K.A., Dewey, A.M., Kavesh, M.H., and Green, R.M. (2011). A chronic high-cholesterol diet paradoxically suppresses hepatic CYP7A1 expression in FVB/NJ mice. J. Lipid Res. *52*, 289–298.

Hidvegi, T., Ewing, M., Hale, P., Dippold, C., Beckett, C., Kemp, C., Maurice, N., Mukherjee, A., Goldbach, C., Watkins, S., et al. (2010). An Autophagy-Enhancing Drug Promotes Degradation of Mutant α1-Antitrypsin Z and Reduces Hepatic Fibrosis. Science *329*, 229–232.

Higuchi, H., and Gores, G.J. (2003). Mechanisms of liver injury: an overview. Curr. Mol. Med. *3*, 483–490.

Hirschfield, G.M., and Gershwin, M.E. (2013). The Immunobiology and Pathophysiology of Primary Biliary Cirrhosis. Annu. Rev. Pathol. Mech. Dis. *8*, 303–330.

Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P., and Kummer, U. (2006). COPASI--a COmplex PAthway SImulator. Bioinforma. Oxf. Engl. *22*, 3067–3074.

Hoppe, A. (2012). What mRNA Abundances Can Tell us about Metabolism. Metabolites *2*, 614–631.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Hyduke, D., Schellenberger, J., Que, R., Fleming, R., Thiele, I., Orth, J., Feist, A., Zielinski, D., Bordbar, A., Lewis, N., et al. (2011). COBRA Toolbox 2.0.

Ide, T., Shimano, H., Yoshikawa, T., Yahagi, N., Amemiya-Kudo, M., Matsuzaka, T., Nakakuki, M., Yatoh, S., Iizuka, Y., Tomita, S., et al. (2003). Cross-talk between peroxisome proliferator-activated receptor (PPAR) alpha and liver X receptor (LXR) in nutritional regulation of fatty acid metabolism. II. LXRs suppress lipid degradation gene promoters through inhibition of PPAR signaling. Mol. Endocrinol. Baltim. Md *17*, 1255–1267.

Ideker, T., Galitski, T., and Hood, L. (2001). A NEW APPROACH TO DECODING LIFE: Systems Biology. Annu. Rev. Genomics Hum. Genet. *2*, 343–372.

Ip, E., Farrell, G.C., Robertson, G., Hall, P., Kirsch, R., and Leclercq, I. (2003). Central role of PPARalpha-dependent hepatic lipid turnover in dietary steatohepatitis in mice. Hepatol. Baltim. Md *38*, 123–132.

Ishihara, K., Miyazaki, A., Nabe, T., Fushimi, H., Iriyama, N., Kanai, S., Sato, T., Uozumi, N., Shimizu, T., and Akiba, S. (2012). Group IVA phospholipase A2 participates in the progression of hepatic fibrosis. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. *26*, 4111–4121.

Jansen, S., Cashman, K., Thompson, J.G., Pantaleon, M., and Kaye, P.L. (2009). Glucose deprivation, oxidative stress and peroxisome proliferator-activated receptor-α (PPARA) cause peroxisome proliferation in preimplantation mouse embryos. Reproduction *138*, 493–505.

Ji, G., Zhao, X., Leng, L., Liu, P., and Jiang, Z. (2011). Comparison of dietary control and atorvastatin on high fat diet induced hepatic steatosis and hyperlipidemia in rats. Lipids Health Dis. *10*, 23.

Jin, X., Chen, Y., Kong, M., Zheng, L., Yang, Y., and Li, Y. (2012). Transition from hepatic steatosis to steatohepatitis: Unique microRNA patterns and potential downstream functions and pathways. J. Gastroenterol. Hepatol. *27*, 331–340.

Jüni, P., Rutjes, A.W.S., and Dieppe, P.A. (2002). Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? BMJ *324*, 1287–1288.

Kacser, H., and Acerenza, L. (1993). A universal method for achieving increases in metabolite production. Eur. J. Biochem. FEBS *216*, 361–367.

Kalhan, S.C., Edmison, J., Marczewski, S., Dasarathy, S., Gruca, L.L., Bennett, C., Duenas, C., and Lopez, R. (2011a). Methionine and protein metabolism in non-alcoholic steatohepatitis: evidence for lower rate of transmethylation of methionine. Clin. Sci. Lond. Engl. 1979 *121*, 179–189.

Kalhan, S.C., Guo, L., Edmison, J., Dasarathy, S., McCullough, A.J., Hanson, R.W., and Milburn, M. (2011b). Plasma metabolomic profile in nonalcoholic fatty liver disease. Metabolism *60*, 404–413.

Kamburov, A., Pentchev, K., Galicka, H., Wierling, C., Lehrach, H., and Herwig, R. (2011a). ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Res. *39*, D712–717.

Kamburov, A., Cavill, R., Ebbels, T.M.D., Herwig, R., and Keun, H.C. (2011b). Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. Bioinforma. Oxf. Engl. *27*, 2917–2918.

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. *28*, 27–30.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. *40*, D109–114.

Kasprzyk, A. (2011). BioMart: driving a paradigm change in biological data management. Database J. Biol. Databases Curation *2011*, bar049.

Kauffman, S., Peterson, C., Samuelsson, B., and Troein, C. (2004). Genetic networks with canalyzing Boolean rules are always stable. Proc. Natl. Acad. Sci. U. S. A. *101*, 17102–17107.

Kern, P.A., Saghizadeh, M., Ong, J.M., Bosch, R.J., Deem, R., and Simsolo, R.B. (1995). The expression of tumor necrosis factor in human adipose tissue. Regulation by obesity, weight loss, and relationship to lipoprotein lipase. J. Clin. Invest. *95*, 2111–2119.

Kessler, T., Hache, H., and Wierling, C. (2013). Integrative analysis of cancer-related signaling pathways. Front. Physiol. *4*, 124.

Khatri, P., Sirota, M., and Butte, A.J. (2012). Ten Years of Pathway Analysis: Current Approaches and Outstanding Challenges. PLoS Comput Biol *8*, e1002375.

Kim, Y.W., Kim, Y.M., Yang, Y.M., Kim, T.H., Hwang, S.J., Lee, J.R., Kim, S.C., and Kim, S.G. (2010). Inhibition of SREBP-1c-mediated hepatic steatosis and oxidative stress by sauchinone, an AMPK-activating lignan in Saururus chinensis. Free Radic. Biol. Med. *48*, 567–578.

Kirsch, R., Clarkson, V., Shephard, E.G., Marais, D.A., Jaffer, M.A., Woodburne, V.E., Kirsch, R.E., and Hall, P. de la M. (2003). Rodent nutritional model of non-alcoholic steatohepatitis: species, strain and sex difference studies. J. Gastroenterol. Hepatol. *18*, 1272–1282.

Kitano, H. (2002a). Systems Biology: A Brief Overview. Science *295*, 1662–1664.

Kitano, H. (2002b). Computational systems biology. Nature *420*, 206–210.

Klamt, S., and Stelling, J. (2002). Combinatorial complexity of pathway analysis in metabolic networks. Mol. Biol. Rep. *29*, 233–236.

Klamt, S., Stelling, J., Ginkel, M., and Gilles, E.D. (2003). FluxAnalyzer: exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps. Bioinforma. Oxf. Engl. *19*, 261–269.

Klemm, K., and Bornholdt, S. (2005). Topology of biological networks and reliability of information processing. Proc. Natl. Acad. Sci. U. S. A. *102*, 18414–18419.

Klipp, E., Liebermeister, W., Wierling, C., Axel, K., Lehrach, H., and Herwig, R. (2009). Systems

Biology- A Textbook (Wiley-VCH Verlag GmbH & C0.kGaA).

Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., et al. (2011). DrugBank 3.0: a comprehensive resource for “omics” research on drugs. Nucleic Acids Res. *39*, D1035–1041.

Koek, G.H., Liedorp, P.R., and Bast, A. (2011). The role of oxidative stress in non-alcoholic steatohepatitis. Clin. Chim. Acta *412*, 1297–1305.

Kong, B., Luyendyk, J.P., Tawfik, O., and Guo, G.L. (2009). Farnesoid X Receptor Deficiency Induces Nonalcoholic Steatohepatitis in Low-Density Lipoprotein Receptor-Knockout Mice Fed a High-Fat Diet. J. Pharmacol. Exp. Ther. *328*, 116–122.

Koves, T.R., Sparks, L.M., Kovalik, J.P., Mosedale, M., Arumugam, R., DeBalsi, K.L., Everingham, K., Thorne, L., Phielix, E., Meex, R.C., et al. (2013). PPARγ coactivator-1α contributes to exercise-induced regulation of intramuscular lipid droplet programming in mice and humans. J. Lipid Res. *54*, 522–534.

Kuhajda, F.P., Aja, S., Tu, Y., Han, W.F., Medghalchi, S.M., El Meskini, R., Landree, L.E., Peterson, J.M., Daniels, K., Wong, K., et al. (2011). Pharmacological glycerol-3-phosphate acyltransferase inhibition decreases food intake and adiposity and increases insulin sensitivity in diet-induced obesity. Am. J. Physiol. Regul. Integr. Comp. Physiol. *301*, R116–130.

Kuo, J.-J., Chang, H.-H., Tsai, T.-H., and Lee, T.-Y. (2012). Curcumin ameliorates mitochondrial dysfunction associated with inhibition of gluconeogenesis in free fatty acid-mediated hepatic lipoapoptosis. Int. J. Mol. Med. *30*, 643–649.

Kursawe, R., Eszlinger, M., Narayan, D., Liu, T., Bazuine, M., Cali, A.M.G., D’Adamo, E., Shaw, M., Pierpont, B., Shulman, G.I., et al. (2010). Cellularity and Adipogenic Profile of the Abdominal Subcutaneous Adipose Tissue From Obese Adolescents: Association With Insulin Resistance and Hepatic Steatosis. Diabetes *59*, 2288–2296.

Ladias, J.A., Hadzopoulou-Cladaras, M., Kardassis, D., Cardot, P., Cheng, J., Zannis, V., and Cladaras, C. (1992). Transcriptional regulation of human apolipoprotein genes ApoB, ApoCIII, and ApoAII by members of the steroid hormone receptor superfamily HNF-4, ARP-1, EAR-2, and EAR-3. J. Biol. Chem. *267*, 15849–15860.

Lam, B., and Younossi, Z.M. (2010). Treatment options for nonalcoholic fatty liver disease. Ther. Adv. Gastroenterol. *3*, 121–137.

Lambert, G., Amar, M.J.A., Guo, G., Brewer, H.B., Jr, Gonzalez, F.J., and Sinal, C.J. (2003). The farnesoid X-receptor is an essential regulator of cholesterol homeostasis. J. Biol. Chem. *278*, 2563–2570.

Larhlimi, A., David, L., Selbig, J., and Bockmayr, A. (2012). F2C2: a fast tool for the computation of flux coupling in genome-scale metabolic networks. BMC Bioinformatics *13*, 57.

Lawler, J.M., Barnes, W.S., Wu, G., Song, W., and Demaree, S. (2002). Direct antioxidant properties of creatine. Biochem. Biophys. Res. Commun. *290*, 47–52.

Le, T.T., Ziemba, A., Urasaki, Y., Hayes, E., Brotman, S., and Pizzorno, G. (2013). Disruption of uridine homeostasis links liver pyrimidine metabolism to lipid accumulation. J. Lipid Res. *54*, 1044–1057.

Leamy, A.K., Egnatchik, R.A., and Young, J.D. (2013). Molecular mechanisms and the role of saturated fatty acids in the progression of non-alcoholic fatty liver disease. Prog. Lipid Res. *52*, 165–174.

Lee, D., Smallbone, K., Dunn, W.B., Murabito, E., Winder, C.L., Kell, D.B., Mendes, P., and Swainston, N. (2012). Improving metabolic flux predictions using absolute gene expression data. BMC Syst. Biol. *6*, 73.

Le Novère, N., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A., Demir, E., Wegner, K., Aladjem, M.I., Wimalaratne, S.M., et al. (2009). The Systems Biology Graphical Notation. Nat. Biotechnol. *27*, 735–741.

Lepley, R.A., and Fitzpatrick, F.A. (1994). Irreversible inactivation of 5-lipoxygenase by leukotriene A4. Characterization of product inactivation with purified enzyme and intact leukocytes. J. Biol. Chem. *269*, 2627–2631.

Li, G.-C. (2013). Tumor markers for hepatocellular carcinoma (Review). Mol. Clin. Oncol.

Li, Z., and Chan, C. (2009). Systems biology for identifying liver toxicity pathways. BMC Proc. *3*, S2.

Li, L.O., Ellis, J.M., Paich, H.A., Wang, S., Gong, N., Altshuller, G., Thresher, R.J., Koves, T.R., Watkins, S.M., Muoio, D.M., et al. (2009). Liver-specific Loss of Long Chain Acyl-CoA Synthetase-1 Decreases Triacylglycerol Synthesis and β-Oxidation and Alters Phospholipid Fatty Acid Composition. J. Biol. Chem. *284*, 27816–27826.

Lin, C.-W., Zhang, H., Li, M., Xiong, X., Chen, X., Chen, X., Dong, X.C., and Yin, X.-M. (2013). Pharmacological promotion of autophagy alleviates steatosis and injury in alcoholic and non-alcoholic fatty liver conditions in mice. J. Hepatol. *58*, 993–999.

Lindén, D., William-Olsson, L., Ahnmark, A., Ekroos, K., Hallberg, C., Sjögren, H.P., Becker, B., Svensson, L., Clapham, J.C., Oscarsson, J., et al. (2006). Liver-directed overexpression of mitochondrial glycerol-3-phosphate acyltransferase results in hepatic steatosis, increased triacylglycerol secretion and reduced fatty acid oxidation. FASEB J. *20*, 434–443.

Loscalzo, J., and Barabasi, A.-L. (2011). Systems biology and the future of medicine. Wiley Interdiscip. Rev. Syst. Biol. Med. *3*, 619–627.

Lu, J.-J., Pan, W., Hu, Y.-J., and Wang, Y.-T. (2012). Multi-Target Drugs: The Trend of Drug Research and Development. PLoS ONE *7*, e40262.

Ma, H.-W., Zhao, X.-M., Yuan, Y.-J., and Zeng, A.-P. (2004). Decomposition of metabolic network into functional modules based on the global connectivity structure of reaction graph. Bioinformatics *20*, 1870–1876.

Ma, K., Saha, P.K., Chan, L., and Moore, D.D. (2006). Farnesoid X receptor is essential for normal glucose homeostasis. J. Clin. Invest. *116*, 1102–1109.

MacAulay, J., Thompson, K., Kiberd, B.A., Barnes, D.C., and Peltekian, K.M. (2006). Serum creatinine in patients with advanced liver disease is of limited value for identification of moderate renal dysfunction: Are the equations for estimating renal function better? Can. J. Gastroenterol. *20*, 521–526.

Machado, D., Soons, Z., Patil, K.R., Ferreira, E.C., and Rocha, I. (2012). Random sampling of elementary flux modes in large-scale metabolic networks. Bioinformatics *28*, i515–i521.

Machado, M., Marques-Vidal, P., and Cortez-Pinto, H. (2006). Hepatic histology in obese patients undergoing bariatric surgery. J. Hepatol. *45*, 600–606.

Mahadevan, R., and Schilling, C.H. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab. Eng. *5*, 264–276.

Mahipal, S.V.K., Subhashini, J., Reddy, M.C., Reddy, M.M., Anilkumar, K., Roy, K.R., Reddy, G.V., and Reddanna, P. (2007). Effect of 15-lipoxygenase metabolites, 15-(S)-HPETE and 15-(S)-HETE on chronic myelogenous leukemia cell line K-562: Reactive oxygen species (ROS) mediate caspase-dependent apoptosis. Biochem. Pharmacol. *74*, 202–214.

Malaguarnera, M., Gargante, M.P., Russo, C., Antic, T., Vacante, M., Malaguarnera, M., Avitabile, T., Li Volti, G., and Galvano, F. (2010). L-carnitine supplementation to diet: a new tool in treatment of nonalcoholic steatohepatitis--a randomized and controlled clinical trial. Am. J. Gastroenterol. *105*, 1338–1345.

Mann, R.E., Smart, R.G., and Govoni, R. (2003). The epidemiology of alcoholic liver disease. Alcohol Res. Health J. Natl. Inst. Alcohol Abuse Alcohol. *27*, 209–219.

Marchesini, G., Bugianesi, E., Forlani, G., Cerrelli, F., Lenzi, M., Manini, R., Natale, S., Vanni, E., Villanova, N., Melchionda, N., et al. (2003). Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. Hepatol. Baltim. Md *37*, 917–923.

Mardinoglu, A., Agren, R., Kampf, C., Asplund, A., Uhlen, M., and Nielsen, J. (2014). Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. Nat. Commun. *5*.

Marra, F., Gastaldelli, A., Baroni, G.S., Tell, G., and Tiribelli, C. (2008). Molecular basis and mechanisms of progression of non-alcoholic steatohepatitis. Trends Mol. Med. *14*, 72–81.

Martínez-Chantar, M.L., Latasa, M.U., Varela-Rey, M., Lu, S.C., García-Trevijano, E.R., Mato, J.M., and Avila, M.A. (2003). L-Methionine Availability Regulates Expression of the Methionine Adenosyltransferase 2A Gene in Human Hepatocarcinoma Cells ROLE OF S-ADENOSYLMETHIONINE. J. Biol. Chem. *278*, 19885–19890.

Martínez-Clemente, M., Ferré, N., González-Périz, A., López-Parra, M., Horrillo, R., Titos, E., Morán-Salvador, E., Miquel, R., Arroyo, V., Funk, C.D., et al. (2010). 5-lipoxygenase deficiency reduces hepatic inflammation and tumor necrosis factor alpha-induced hepatocyte damage in hyperlipidemia-prone ApoE-null mice. *51*, 817–827.

Mason, E.F., and Rathmell, J.C. (2011). Cell metabolism: an essential link between cell growth and apoptosis. Biochim. Biophys. Acta *1813*, 645–654.

Mato, J.M., Corrales, F.J., Lu, S.C., and Avila, M.A. (2002). S-Adenosylmethionine: a control switch that regulates liver function. FASEB J. *16*, 15–26.

Mato, J.M., Martínez-Chantar, M.L., and Lu, S.C. (2013). S-adenosylmethionine metabolism and liver disease. *12*, 183–189.

Matsuno, H., Tanaka, Y., Aoshima, H., Doi, A., Matsui, M., and Miyano, S. (2003). Biopathways

representation and simulation on hybrid functional Petri net. In Silico Biol. *3*, 389–404.

Matsuzawa, N., Takamura, T., Kurita, S., Misu, H., Ota, T., Ando, H., Yokoyama, M., Honda, M., Zen, Y., Nakanuma, Y., et al. (2007). Lipid-induced oxidative stress causes steatohepatitis in mice fed an atherogenic diet. Hepatol. Baltim. Md *46*, 1392–1403.

Mehendale, H.M. (2000). PPAR-alpha: a key to the mechanism of hepatoprotection by clofibrate. Toxicol. Sci. Off. J. Soc. Toxicol. *57*, 187–190.

Meléndez-Hevia, E., Waddell, T.G., and Montero, F. (1994). Optimization of Metabolism: The Evolution of Metabolic Pathways Toward Simplicity Through the Game of the Pentose Phosphate Cycle. J. Theor. Biol. *166*, 201–220.

Michal,1998. Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology (Wiley-Spektrum).

Michal, G., and Schomburg, D. (2013). Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology (John Wiley & Sons).

Miller, S.B. (2006). Prostaglandins in Health and Disease: An Overview. Semin. Arthritis Rheum. *36*, 37–49.

Mishra, P., and Younossi, Z.M. (2007). Current treatment strategies for non-alcoholic fatty liver disease (NAFLD). Curr. Drug Discov. Technol. *4*, 133–140.

Mitsuyoshi, H., Yasui, K., Harano, Y., Endo, M., Tsuji, K., Minami, M., Itoh, Y., Okanoue, T., and Yoshikawa, T. (2009). Analysis of hepatic genes involved in the metabolism of fatty acids and iron in nonalcoholic fatty liver disease. Hepatol. Res. *39*, 366–373.

Miyazaki, M., Flowers, M.T., Sampath, H., Chu, K., Otzelberger, C., Liu, X., and Ntambi, J.M. (2007). Hepatic Stearoyl-CoA Desaturase-1 Deficiency Protects Mice from Carbohydrate-Induced Adiposity and Hepatic Steatosis. Cell Metab. *6*, 484–496.

Moon, Y.-A., Liang, G., Xie, X., Frank-Kamenetsky, M., Fitzgerald, K., Koteliansky, V., Brown, M.S., Goldstein, J.L., and Horton, J.D. (2012). The Scap/SREBP pathway is essential for developing diabetic fatty liver and carbohydrate-induced hypertriglyceridemia in animals. Cell Metab. *15*, 240–246.

Mormeneo, E., Jimenez-Mallebrera, C., Palomer, X., De Nigris, V., Vázquez-Carrera, M., Orozco, A., Nascimento, A., Colomer, J., Lerín, C., and Gómez-Foix, A.M. (2012). PGC-1α Induces Mitochondrial and Myokine Transcriptional Programs and Lipid Droplet and Glycogen Accumulation in Cultured Human Skeletal Muscle Cells. PLoS ONE *7*, e29985.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods *5*, 621–628.

Motojima, K., and Hirai, T. (2006). Peroxisome proliferator-activated receptor alpha plays a vital role in inducing a detoxification system against plant compounds with crosstalk with other xenobiotic nuclear receptors. FEBS J. *273*, 292–300.

Mu, J., and Roach, P.J. (1998). Characterization of Human Glycogenin-2, a Self-glucosylating Initiator of Liver Glycogen Metabolism. J. Biol. Chem. *273*, 34850–34856.

Murali, G., Milne, G.L., Webb, C.D., Stewart, A.B., McMillan, R.P., Lyle, B.C., Hulver, M.W., and Saraswathi, V. (2012). Fish oil and indomethacin in combination potently reduce dyslipidemia and hepatic steatosis in LDLR(-/-) mice. J. Lipid Res. *53*, 2186–2197.

Muriel, P., and Rivera-Espinoza, Y. (2008). Beneficial drugs for liver diseases. J. Appl. Toxicol. *28*, 93–103.

Müller, A.C., and Bockmayr, A. (2013). Fast thermodynamically constrained flux variability analysis. Bioinformatics *29*, 903–909.

Müller, A.C., and Bockmayr, A. Flux modules in metabolic networks. J. Math. Biol. 1–29.

Namekata, K., Enokido, Y., Ishii, I., Nagai, Y., Harada, T., and Kimura, H. (2004). Abnormal Lipid Metabolism in Cystathionine β-Synthase-deficient Mice, an Animal Model for Hyperhomocysteinemia. J. Biol. Chem. *279*, 52961–52969.

Navid, A., and Almaas, E. (2012). Genome-level transcription data of Yersinia pestis analyzed with a New metabolic constraint-based approach. BMC Syst. Biol. *6*, 150.

Neuschwander-Tetri, B.A. (2005). Nonalcoholic steatohepatitis and the metabolic syndrome. Am. J. Med. Sci. *330*, 326–335.

Niculescu, M.D., Costa, K.-A. da, Fischer, L.M., and Zeisel, S.H. (2007). Lymphocyte gene expression in subjects fed a low-choline diet differs between those who develop organ dysfunction and those who do not. Am. J. Clin. Nutr. *86*, 230–239.

Niederberger, P., Prasad, R., Miozzari, G., and Kacser, H. (1992). A strategy for increasing an in vivo flux by genetic manipulations. The tryptophan system of yeast. Biochem. J. *287 (Pt 2)*, 473–479.

Nielsen, J. (2003). It Is All about Metabolic Fluxes. J. Bacteriol. *185*, 7031–7035.

Noguchi, Y., Young, J.D., Aleman, J.O., Hansen, M.E., Kelleher, J.K., and Stephanopoulos, G. (2009). Effect of Anaplerotic Fluxes and Amino Acid Availability on Hepatic Lipoapoptosis. J. Biol. Chem. *284*, 33425–33436.

Noguchi, Y., Young, J.D., Aleman, J.O., Hansen, M.E., Kelleher, J.K., and Stephanopoulos, G. (2011). Tracking cellular metabolomics in lipoapoptosis- and steatosis-developing liver cells. Mol. Biosyst. *7*, 1409.

Notebaart, R.A., Teusink, B., Siezen, R.J., and Papp, B. (2008). Co-Regulation of Metabolic Genes Is Better Explained by Flux Coupling Than by Network Distance. PLoS Comput Biol *4*, e26.

Oh, Y.-K., Palsson, B.O., Park, S.M., Schilling, C.H., and Mahadevan, R. (2007). Genome-scale Reconstruction of Metabolic Network in Bacillus subtilis Based on High-throughput Phenotyping and Gene Essentiality Data. J. Biol. Chem. *282*, 28791–28799.

Olivier, B.G., Rohwer, J.M., and Hofmeyr, J.-H.S. (2005). Modelling cellular systems with PySCeS. Bioinforma. Oxf. Engl. *21*, 560–561.

Ong, J.P., Elariny, H., Collantes, R., Younoszai, A., Chandhoke, V., Reines, H.D., Goodman, Z., and Younossi, Z.M. (2005). Predictors of nonalcoholic steatohepatitis and advanced fibrosis in morbidly obese patients. Obes. Surg. *15*, 310–315.

Onofrei, M.D., Butler, K.L., Fuke, D.C., and Miller, H.B. (2008). Safety of statin therapy in patients with preexisting liver disease. Pharmacotherapy *28*, 522–529.

Oosterveer, M.H., van Dijk, T.H., Tietge, U.J.F., Boer, T., Havinga, R., Stellaard, F., Groen, A.K., Kuipers, F., and Reijngoud, D.-J. (2009). High Fat Feeding Induces Hepatic Fatty Acid Elongation in Mice. PLoS ONE *4*, e6066.

Orth, J.D., Thiele, I., and Palsson, B.Ø. (2010). What is flux balance analysis? Nat. Biotechnol. *28*, 245–248.

Oyadomari, S., Harding, H.P., Zhang, Y., Oyadomari, M., and Ron, D. (2008). Dephosphorylation of Translation Initiation Factor 2α Enhances Glucose Tolerance and Attenuates Hepatosteatosis in Mice. Cell Metab. *7*, 520–532.

Papin, J.A., Reed, J.L., and Palsson, B.O. (2004). Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. Trends Biochem. Sci. *29*, 641–647.

Paweletz, C.P., Charboneau, L., Bichsel, V.E., Simone, N.L., Chen, T., Gillespie, J.W., Emmert-Buck, M.R., Roth, M.J., Petricoin III, E.F., and Liotta, L.A. (2001). Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. Oncogene *20*, 1981–1989.

Perrone, M.G., Scilimati, A., Simone, L., and Vitale, P. (2010). Selective COX-1 inhibition: A therapeutic target to be reconsidered. Curr. Med. Chem. *17*, 3769–3805.

Pessayre, D., Mansouri, A., and Fromenty, B. (2002). V. Mitochondrial dysfunction in steatohepatitis. Am. J. Physiol. - Gastrointest. Liver Physiol. *282*, G193–G199.

Pfeiffer, T., and Bonhoeffer, S. (2004). Evolution of cross-feeding in microbial populations. Am. Nat. *163*, E126–135.

Pfeiffer, T., Sánchez-Valdenebro, I., Nuño, J.C., Montero, F., and Schuster, S. (1999). METATOOL: for studying metabolic networks. Bioinforma. Oxf. Engl. *15*, 251–257.

Price, N.D., Reed, J.L., and Palsson, B.Ø. (2004). Genome-scale models of microbial cells: evaluating the consequences of constraints. Nat. Rev. Microbiol. *2*, 886–897.

Puljiz, Z., Stimac, D., Kovac, D., Puljiz, M., Bratanić, A., Kovacić, V., Kardum, D., Bonacin, D., and Hozo, I. (2010). Predictors of nonalcoholic steatohepatitis in patients with elevated alanine aminotransferase activity. Coll. Antropol. *34 Suppl 1*, 33–37.

Puri, P., Mirshahi, F., Cheung, O., Natarajan, R., Maher, J.W., Kellum, J.M., and Sanyal, A.J. (2008). Activation and dysregulation of the unfolded protein response in nonalcoholic fatty liver disease. Gastroenterology *134*, 568–576.

Puri, P., Wiest, M.M., Cheung, O., Mirshahi, F., Sargeant, C., Min, H.-K., Contos, M.J., Sterling, R.K., Fuchs, M., Zhou, H., et al. (2009). The plasma lipidomic signature of nonalcoholic steatohepatitis. Hepatol. Baltim. Md *50*, 1827–1838.

Qin, S., Chen, J., Tanigawa, S., and Hou, D.-X. (2012). Gene expression profiling and pathway network analysis of hepatic metabolic enzymes targeted by baicalein. J. Ethnopharmacol. *140*, 131–140.

Quackenbush, J. (2006). Microarray Analysis and Tumor Classification. N. Engl. J. Med. *354*, 2463–2472.

Rader, D.J., and Daugherty, A. (2008). Translating molecular discoveries into new therapies for atherosclerosis. Nature *451*, 904–913.

Rakhshandehroo, M., Knoch, B., Müller, M., and Kersten, S. (2010). Peroxisome proliferator-activated receptor alpha target genes. PPAR Res. *2010*.

Rallidis, L.S., Drakoulis, C.K., and Parasi, A.S. (2004). Pravastatin in patients with nonalcoholic steatohepatitis: results of a pilot study. Atherosclerosis *174*, 193–196.

Rao, M.S., and Reddy, J.K. (2001). Peroxisomal beta-oxidation and steatohepatitis. Semin. Liver Dis. *21*, 43–55.

Rasmussen, B.B., and Wolfe, R.R. (1999). Regulation of fatty acid oxidation in skeletal muscle. Annu. Rev. Nutr. *19*, 463–484.

Raue, A., Schilling, M., Bachmann, J., Matteson, A., Schelke, M., Kaschek, D., Hug, S., Kreutz, C., Harms, B.D., Theis, F.J., et al. (2013). Lessons Learned from Quantitative Dynamical Modeling in Systems Biology. PLoS ONE *8*, e74335.

Rayner, K.J., Esau, C.C., Hussain, F.N., McDaniel, A.L., Marshall, S.M., van Gils, J.M., Ray, T.D., Sheedy, F.J., Goedeke, L., Liu, X., et al. (2011). Inhibition of miR-33a/b in non-human primates raises plasma HDL and lowers VLDL triglycerides. Nature *478*, 404–407.

Reddy, J.K. (2001). Nonalcoholic steatosis and steatohepatitis. III. Peroxisomal beta-oxidation, PPAR alpha, and steatohepatitis. Am. J. Physiol. Gastrointest. Liver Physiol. *281*, G1333–1339.

Reddy, J.K., and Rao, M.S. (2006). Lipid Metabolism and Liver Inflammation. II. Fatty liver disease and fatty acid oxidation. Am. J. Physiol. - Gastrointest. Liver Physiol. *290*, G852–G858.

Resendis-Antonio, O., Hernández, M., Mora, Y., and Encarnación, S. (2012). Functional Modules, Structural Topology, and Optimal Activity in Metabolic Networks. PLoS Comput Biol *8*, e1002720.

Rezola, A., Pey, J., Figueiredo, L.F. de, Podhorski, A., Schuster, S., Rubio, A., and Planes, F.J. (2013). Selection of human tissue-specific elementary flux modes using gene expression data. Bioinformatics *29*, 2009–2016.

Rives, A.W., and Galitski, T. (2003). Modular organization of cellular networks. Proc. Natl. Acad. Sci. U. S. A. *100*, 1128–1133.

Robert, K., Nehmé, J., Bourdon, E., Pivert, G., Friguet, B., Delcayre, C., Delabar, J.-M., and Janel, N. (2005). Cystathionine beta synthase deficiency promotes oxidative stress, fibrosis, and steatosis in mice liver. Gastroenterology *128*, 1405–1415.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Rohr, C., Marwan, W., and Heiner, M. (2010). Snoopy—a unifying Petri net framework to investigate biomolecular networks. Bioinformatics *26*, 974–975.

Ruths, D., Muller, M., Tseng, J.-T., Nakhleh, L., and Ram, P.T. (2008). The Signaling Petri Net-Based

Simulator: A Non-Parametric Strategy for Characterizing the Dynamics of Cell-Specific Signaling Networks. PLoS Comput Biol *4*, e1000005.

Ryan, M.C., Desmond, P.V., Slavin, J.L., and Congiu, M. (2011). Expression of genes involved in lipogenesis is not increased in patients with HCV genotype 3 in human liver. J. Viral Hepat. *18*, 53–60.

Saltiel, A.R., and Kahn, C.R. (2001). Insulin signalling and the regulation of glucose and lipid metabolism. Nature *414*, 799–806.

Scheer, M., Grote, A., Chang, A., Schomburg, I., Munaretto, C., Rother, M., Söhngen, C., Stelzer, M., Thiele, J., and Schomburg, D. (2011). BRENDA, the enzyme information system in 2011. Nucleic Acids Res. *39*, D670–676.

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

Schulz, M. (2012). Identification of potential drug targets in kinetic networks described by ordinary differential equations.

Schuppan, D., Gorrell, M.D., Klein, T., Mark, M., and Afdhal, N.H. (2010). The challenge of developing novel pharmacological therapies for non-alcoholic steatohepatitis. Liver Int. Off. J. Int. Assoc. Study Liver *30*, 795–808.

Schuster, S., and Hilgetag, C. (1994). ON ELEMENTARY FLUX MODES IN BIOCHEMICAL REACTION SYSTEMS AT STEADY STATE. J. Biol. Syst. *02*, 165–182.

Schuster, S., Dandekar, T., and Fell, D.A. (1999). Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. Trends Biotechnol. *17*, 53–60.

Schwab, J.M., and Serhan, C.N. (2006). Lipoxins and new lipid mediators in the resolution of inflammation. Curr. Opin. Pharmacol. *6*, 414–420.

Sharma, A., Bartell, S.M., Baile, C.A., Chen, B., Podolsky, R.H., McIndoe, R.A., and She, J.-X. (2010). Hepatic Gene Expression Profiling Reveals Key Pathways Involved in Leptin-Mediated Weight Loss in ob/ob Mice. PLoS ONE *5*, e12147.

Sigurdsson, M.I., Jamshidi, N., Steingrimsson, E., Thiele, I., and Palsson, B.Ø. (2010). A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1. BMC Syst. Biol. *4*, 140.

Simonen, M., Männistö, V., Leppänen, J., Kaminska, D., Kärjä, V., Venesmaa, S., Käkelä, P., Kuusisto, J., Gylling, H., Laakso, M., et al. (2013). Desmosterol in human nonalcoholic steatohepatitis. Hepatology *58*, 976–982.

Singh, D.K., Rastogi, A., Sakhuja, P., Gondal, R., and Sarin, S.K. (2010). Comparison of clinical, biochemical and histological features of alcoholic steatohepatitis and non-alcoholic steatohepatitis in Asian Indian patients. Indian J. Pathol. Microbiol. *53*, 408–413.

Singh, R., Wang, Y., Xiang, Y., Tanaka, K.E., Gaarde, W.A., and Czaja, M.J. (2009). Differential effects of JNK1 and JNK2 inhibition on murine steatohepatitis and insulin resistance. Hepatol. Baltim. Md *49*, 87–96.

Spurrier, B., Ramalingam, S., and Nishizuka, S. (2008). Reverse-phase protein lysate microarrays for cell signaling analysis. Nat. Protoc. *3*, 1796–1808.

Srivastava, S., and Chan, C. (2007). Hydrogen peroxide and hydroxyl radicals mediate palmitate-induced cytotoxicity to hepatoma cells: relation to mitochondrial permeability transition. Free Radic. Res. *41*, 38–49.

Starmann, J., Fälth, M., Spindelböck, W., Lanz, K.-L., Lackner, C., Zatloukal, K., Trauner, M., and Sültmann, H. (2012). Gene Expression Profiling Unravels Cancer-Related Hepatic Molecular Signatures in Steatohepatitis but Not in Steatosis. PLoS ONE *7*, e46584.

Stavrum, A.-K., Heiland, I., Schuster, S., Puntervoll, P., and Ziegler, M. (2013). Model of Tryptophan Metabolism, Readily Scalable using Tissue-Specific Gene Expression Data. J. Biol. Chem. jbc.M113.474908.

Strohman, R. (2002). Maneuvering in the complex path from genotype to phenotype. Science *296*, 701–703.

Sun, Z., Miller, R.A., Patel, R.T., Chen, J., Dhir, R., Wang, H., Zhang, D., Graham, M.J., Unterman, T.G., Shulman, G.I., et al. (2012). Hepatic Hdac3 promotes gluconeogenesis by repressing lipid synthesis and sequestration. Nat. Med. *18*, 934–942.

Tanaka, N., Matsubara, T., Krausz, K.W., Patterson, A.D., and Gonzalez, F.J. (2012). Disruption of phospholipid and bile acid homeostasis in mice with nonalcoholic steatohepatitis. Hepatol. Baltim. Md *56*, 118–129.

Tanaka, S., Yamamoto, H., Nakahashi, O., Kagawa, T., Ishiguro, M., Masuda, M., Kozai, M., Ikeda, S., Taketani, Y., and Takeda, E. (2013). Dietary phosphate restriction induces hepatic lipid accumulation through dysregulation of cholesterol metabolism in mice. Nutr. Res. N. Y. N *33*, 586–593.

Titos, E., Clària, J., Planagumà, A., López-Parra, M., González-Périz, A., Gaya, J., Miquel, R., Arroyo, V., and Rodés, J. (2005). Inhibition of 5-lipoxygenase-activating protein abrogates experimental liver injury: role of Kupffer cells. J. Leukoc. Biol. *78*, 871–878.

Tosello-Trampont, A.-C., Landes, S.G., Nguyen, V., Novobrantseva, T.I., and Hahn, Y.S. (2012). Kuppfer cells trigger nonalcoholic steatohepatitis development in diet-induced mouse model through tumor necrosis factor-α production. J. Biol. Chem. *287*, 40161–40172.

Tyson, J.J., Chen, K.C., and Novak, B. (2003). Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr. Opin. Cell Biol. *15*, 221–231.

Ulbrich, H., Soehnlein, O., Xie, X., Eriksson, E.E., Lindbom, L., Albrecht, W., Laufer, S., and Dannhardt, G. (2005). Licofelone, a novel 5-LOX/COX-inhibitor, attenuates leukocyte rolling and adhesion on endothelium under flow. Biochem. Pharmacol. *70*, 30–36.

Unger, R.H., and Orci, L. (2002). Lipoapoptosis: its mechanism and its diseases. Biochim. Biophys. Acta *1585*, 202–212.

Valenti, L., Fracanzani, A.L., Dongiovanni, P., Santorelli, G., Branchi, A., Taioli, E., Fiorelli, G., and Fargion, S. (2002). Tumor necrosis factor alpha promoter polymorphisms and insulin resistance in nonalcoholic fatty liver disease. Gastroenterology *122*, 274–280.

Varma, A., and Palsson, B.O. (1994). Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use. Nat. Biotechnol. *12*, 994–998.

Vera, J.C., Wheat, C.W., Fescemyer, H.W., Frilander, M.J., Crawford, D.L., Hanski, I., and Marden, J.H. (2008). Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. Mol. Ecol. *17*, 1636–1647.

Volynets, V., Spruss, A., Kanuri, G., Wagnerberger, S., Bischoff, S.C., and Bergheim, I. (2010). Protective effect of bile acids on the onset of fructose-induced hepatic steatosis in mice. J. Lipid Res. *51*, 3414–3424.

Voss, K., Heiner, M., and Koch, I. (2011). Petri nets for steady state analysis of metabolic systems. Stud. Health Technol. Inform. *162*, 56–76.

Wakil, S.J., and Abu-Elheiga, L.A. (2009). Fatty acid metabolism: target for metabolic syndrome. J. Lipid Res. *50 Suppl*, S138–143.

Wang, D., and DuBois, R.N. (2010). Eicosanoids and cancer. Nat. Rev. Cancer *10*, 181–193.

Wang, B., Hsu, S.-H., Frankel, W., Ghoshal, K., and Jacob, S.T. (2012). Stat3-mediated activation of microRNA-23a suppresses gluconeogenesis in hepatocellular carcinoma by down-regulating glucose-6-phosphatase and peroxisome proliferator-activated receptor gamma, coactivator 1 alpha. Hepatol. Baltim. Md *56*, 186–197.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. *10*, 57–63.

Want, E.J., Cravatt, B.F., and Siuzdak, G. (2005). The Expanding Role of Mass Spectrometry in Metabolite Profiling and Characterization. ChemBioChem *6*, 1941–1951.

Washburn, M.P., Koller, A., Oshiro, G., Ulaszek, R.R., Plouffe, D., Deciu, C., Winzeler, E., and Yates, J.R. (2003). Protein pathway and complex clustering of correlated mRNA and protein expression analyses in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. *100*, 3107–3112.

Weng, H., Ji, X., Naito, Y., Endo, K., Ma, X., Takahashi, R., Shen, C., Hirokawa, G., Fukushima, Y., and Iwai, N. (2013). Pex11α deficiency impairs peroxisome elongation and division and contributes to nonalcoholic fatty liver in mice. Am. J. Physiol. - Endocrinol. Metab. *304*, E187–E196.

Wierling, C., Kühn, A., Hache, H., Daskalaki, A., Maschke-Dutz, E., Peycheva, S., Li, J., Herwig, R., and Lehrach, H. (2012). Prediction in the face of uncertainty: a Monte Carlo-based approach for systems biology of cancer treatment. Mutat. Res. *746*, 163–170.

Wijeratne, S.S.K., Cuppett, S.L., and Schlegel, V. (2005). Hydrogen Peroxide Induced Oxidative Stress Damage and Antioxidant Enzyme Response in Caco-2 Human Colon Cells. J. Agric. Food Chem. *53*, 8768–8774.

Wishart, D.S., Tzur, D., Knox, C., Eisner, R., Guo, A.C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., et al. (2007). HMDB: the Human Metabolome Database. Nucleic Acids Res. *35*, D521–526.

Wittig, U., Kania, R., Golebiewski, M., Rey, M., Shi, L., Jong, L., Algaa, E., Weidemann, A., Sauer-Danzwith, H., Mir, S., et al. (2012). SABIO-RK--database for biochemical reaction kinetics. Nucleic Acids Res. *40*, D790–D796.

Wolf-Yadlin, A., Kumar, N., Zhang, Y., Hautaniemi, S., Zaman, M., Kim, H.-D., Grantcharova, V., Lauffenburger, D.A., and White, F.M. (2006). Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. Mol. Syst. Biol. *2*, 54.

Wolkenhauer, O. (2014). Why model? Front. Physiol. *5*.

Wu, C., Miloslavskaya, I., Demontis, S., Maestro, R., and Galaktionov, K. (2004). Regulation of cellular response to oncogenic and oxidative stress by Seladin-1. Nature *432*, 640–645.

Xanthakos, S., Miles, L., Bucuvalas, J., Daniels, S., Garcia, V., and Inge, T. (2006). Histologic spectrum of nonalcoholic fatty liver disease in morbidly obese adolescents. Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc. *4*, 226–232.

Yang, K., Ma, W., Liang, H., Ouyang, Q., Tang, C., and Lai, L. (2007). Dynamic simulations on the arachidonic acid metabolic network. PLoS Comput. Biol. *3*, e55.

Yetukuri, L., Katajamaa, M., Medina-Gomez, G., Seppänen-Laakso, T., Vidal-Puig, A., and Oresic, M. (2007). Bioinformatics strategies for lipidomics analysis: characterization of obesity related hepatic steatosis. BMC Syst. Biol. *1*, 12.

Yoon, S., Jung, J., Kim, T., Park, S., Chwae, Y.-J., Shin, H.-J., and Kim, K. (2011). Adiponectin, a downstream target gene of peroxisome proliferator-activated receptor γ, controls hepatitis B virus replication. Virology *409*, 290–298.

Youssef, W.I., and McCullough, A.J. (2002). Steatohepatitis in obese individuals. Best Pract. Res. Clin. Gastroenterol. *16*, 733–747.

Yu, J., Ip, E., Dela Peña, A., Hou, J.Y., Sesha, J., Pera, N., Hall, P., Kirsch, R., Leclercq, I., and Farrell, G.C. (2006). COX-2 induction in mice with experimental nutritional steatohepatitis: Role as pro-inflammatory mediator. Hepatol. Baltim. Md *43*, 826–836.

Zein, C.O., Jorgensen, R.A., Clarke, B., Wenger, D.E., Keach, J.C., Angulo, P., and Lindor, K.D. (2005). Alendronate improves bone mineral density in primary biliary cirrhosis: A randomized placebo-controlled trial. Hepatology *42*, 762–771.

Zhang, W., Li, F., and Nie, L. (2010). Integrating multiple “omics” analysis for microbial biology: application and methodologies. Microbiol. Read. Engl. *156*, 287–301.

Zhao, J., Ding, G.-H., Tao, L., Yu, H., Yu, Z.-H., Luo, J.-H., Cao, Z.-W., and Li, Y.-X. (2007). Modular co-evolution of metabolic networks. BMC Bioinformatics *8*, 311.

Zhou, G., Myers, R., Li, Y., Chen, Y., Shen, X., Fenyk-Melody, J., Wu, M., Ventre, J., Doebber, T., Fujii, N., et al. (2001). Role of AMP-activated protein kinase in mechanism of metformin action. J. Clin. Invest. *108*, 1167–1174.

Zou, J., Zheng, M.-W., Li, G., and Su, Z.-G. (2013). Advanced Systems Biology Methods in Drug Discovery and Translational Biomedicine. BioMed Res. Int. *2013*.

Abbreviations

AA	arachidonic acid
ALD	alcoholic liver disease
ATL	aspirin-triggered LXA_4
CV	coefficients of variation
DDC	3,5-diethoxycarbonyl-1,4-dihydrocollidine
EFM	elementary flux modes
EGFR	epidermal growth factor receptor
FBA	flux balance analysis
FCA	Flux coupling analysis
FDA	Food and Drug Administration
Fox	forkhead protein family
GMM	genome scale mouse metabolic model
HCC	hepatocellular carcinoma
HDL	high density lipoprotein
HFD	high fat diet
HNF	hepatic nuclear factor
IL-6	interleukin 6
JNK	jun N-terminal kinase
LDL	low-density lipoprotein
LXA_4	lipoxin
MAPK	mitogen-activated protein kinase
MCD	methionine-choline deficient

MPN	Petri net-based method for metabolic pathways
MS	mass spectrometry
NAFLD	nonalcoholic fatty liver disease
NASH	nonalcoholic steatohepatitis
NEFA	non-esterified fatty acids
NF-κB	nuclear factor-κB
ODE	ordinary differential equations
ORA	over-representation analysis
PCA	principle component analysis
PGC1	peroxisome proliferator-activated receptor gamma coactivator 1
PKC	protein kinase C
PPAR	peroxisome proliferator-activated receptor
PPP	pentose phosphate pathway
RPPA	reverse phase protein array
SAMe	S-adenosylmethionine
SREBP	sterol regulatory element binding protein
TCA	citric acid cycle
TNF-α	tumor necrosis factor α
VLDL	very-low-density lipoprotein

Declaration

I hereby declare that I am the sole author of this thesis and that has not been submitted, either in the same or a different form, to this or any other university for a degree. I also declare that the content of this thesis is the product of my own work and effort. Where other sources of information have been used, they have been acknowledged.

Signature:

Place: Berlin Date: 08/07/2014