THE APPLICATION OF METABOLIC NETWORK ANALYSIS IN UNDERSTANDING AND TARGETING METABOLISM FOR DRUG DISCOVERY

by

Jiangxia Liu

B.S., Wuhan University, 2005

Submitted to the Graduate Faculty of

Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in Cellular and Molecular Pathology

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Jiangxia Liu

It was defended on

December 03, 2010

and approved by

Dr. Bennett Van Houten, PhD, Departments of Pharmacology and Chemical Biology, Biochemistry and Molecular Genetics

Dr. Marie DeFrances, MD, PhD, Department of Pathology

Dr. Alexander Doemling, PhD, Department of Pharmacology and Chemical Biology

Committee Chair: Dr. Robert Bowser, PhD, Department of Pathology

Dissertation Advisor: Dr. Zoltan N. Oltvai, MD, Department of Pathology

THE APPLICATION OF METABOLIC NETWORK ANALYSIS IN UNDERSTANDING AND TARGETING METABOLISM FOR DRUG DISCOVERY

Jiangxia Liu, Ph.D.

University of Pittsburgh, 2010

Metabolic networks provide a vital framework for understanding the cellular metabolism in both physiological and pathophysiological states, which will ultimately facilitate network analysis-based drug discovery. In this thesis, we aim to employ a metabolic network analysis approach to study cancer metabolism (a pathophysiological state) and the metabolism of the bacterial pathogen, *S. aureus* (a physiological state), in order to understand, predict, and ultimately target cell metabolism for drug discovery.

Cancer cells have distinct metabolism that highly depend on glycolysis instead of mitochondrial oxidative phosphorylation alone, even in the presence of oxygen, also called aerobic glycolysis or the Warburg effect, which may offer novel therapeutic opportunities. However, the origin of the Warburg effect is only partially understood. To understand the origin of cancer metabolism, our theoretical collaborator, Prof. Alexei Vazquez, developed a reduced flux balance model of human cell metabolism incorporating the macromolecular crowding (MC) constraint and the maximum glucose uptake constraint. The simulations successfully captured the main characteristics of cancer metabolism (aerobic glycolysis), indicating that MC constraint may be a potential origin of the Warburg effect. Notably, when we experimentally tested the model with mammalian cells from low to high growth rates as a proxy of MC alteration, we find that, consistent with the model, faster growing cells indeed have increased aerobic glycolysis.

Moreover, the metabolic network analysis approach has also been shown to be capable of predicting the drug targets against pathogen metabolism when completely reconstructed metabolic networks are available. We deduced common antibiotic targets in *Escherichia coli* and *Staphylococcus aureus* by identifying shared tissue-specific or uniformly essential metabolic reactions in their metabolic networks. We then predicted through virtual screening dozens of potential

inhibitors for several enzymes of these reactions and demonstrated experimentally that a subset of these inhibited both enzyme activities in vitro and bacterial cell viability.

Our results indicate that the metabolic network analysis approach is able to facilitate the understanding of cellular metabolism by identifying potential constraints and predicting as well as ultimately targeting the metabolism of the organisms whose complete metabolic networks are available through the seamless integration of virtual screening with experimental validation.

TABLE OF CONTENTS

PREFACE

My Ph.D. is a journey that I wouldn't have been able to complete without the support of many people.

First, I would like to express my appreciation and gratitude to my Ph.D. mentor, Dr. Zoltan N. Oltvai for his encouragement, guidance, and financial support throughout my thesis work. I would also like to thank Professors Marie DeFrances, Robert Bowser, Alexander Doemling, Xiaoming Yin and Bennett Van Houten for their service on my dissertation committee. Their suggestions, comments and continuous supports have been a critical part in achieving my goals.

I would like to thank all the previous and present members in the Oltvai Lab, Yi Zhou, Basak Isin, Krin Kay, and Qasim Beg, without whom my work could not have been accomplished. In recognition of their scientific and intellectual contributions to this research, special thanks go out to: Yao Shen, Singe Henriksen, Guillerma Estiu and Dr. Olaf Wiest in University of Notre Dame; Dr. Alexei Vazquez in The Cancer Institute of New Jersey; Vinayak Kapatral in Integrated Genomics Inc.; and Dr. Deok-Sun Lee, Dr. Yong-Yeoh Anh and Dr. Albert-László Barabási in Northeastern University. I particularly appreciate Yi and her family for all the enjoyable scientific and non-scientific discussions as well as great time together.

Finally, and most importantly, I extend my thanks to all members in my family, especially my father Guomeng Liu, my mother Yunzhen Liang, my brother Xingming Liu, my lovely daughter Iris Cao and my husband Liangliang Cao. I could have not succeeded without their unconditional love and support. Thank you all!

1.0 INTRODUCTION

Cell metabolism, including both catabolism and anabolism, is the foundation of life, allowing cells to reproduce and divide, and to maintain cell structures and functions. Cell metabolism in physiological or pathophysiological states has been studied extensively using reductionist methodologies, but the understanding of cell metabolism as a complete interconnected system by such approaches remains incomplete. To overcome these limitations more recently developed network-based approaches - integrating genomic, transcriptomic, proteomic and metabolomic data - offer a new framework for understanding physiological or pathophysiological metabolisms, which in turn may ultimately contribute to the development of novel therapeutic approaches. This thesis focuses on employing a systems biology approach to better understand cancer metabolism (a pathophysiological state) and the metabolism of the bacterium, *S. aureus* (a physiological state), and to demonstrate that the derived insights can contribute to drug discovery.

The following introduction describes the basics of the systems-level analysis of metabolic networks, the biology of cancer metabolism, and the antibiotic resistance mechanisms of *S. aureus.* This background information represents the main elements required to understand the significance of this thesis.

1.1 METABOLIC NETWORK ANALYSIS

1.1.1 Metabolic networks

Cell metabolism, the integrated interconversion of thousands of metabolites through enzyme-catalyzed biochemical reactions, can be reduced to a complex network of interconnected molecules and reactions that function to generate mass, energy, information transfer and cell-fate specification. While reductionist approaches remain critically important to develop an understanding of the individual functions of genes and proteins, inevitably most cellular components exert their function through a network of metabolic and genetic interactions (Barabasi and Oltvai 2004; Kroemer and Pouyssegur 2008; Brezina 2010; Tyson and Novak 2010). Recently, genome-driven *in silico* modeling and analysis of metabolic networks have been recognized as a promising approach for characterizing complex cellular systems and for deriving novel strategies for biomedical and biotechnological applications in the context of systems biology and biotechnology (Stelling, Klamt et al. 2002; Bugrim, Nikolskaya et al. 2004; Feist and Palsson 2008; Mo and Palsson 2009; Boran and Iyengar 2010).

Metabolic network is a reduced representation of cellular metabolism at a highly abstract level, where each node represents a metabolite that is connected to others by links, each of which represents a metabolic reaction catalyzed by a gene-encoded metabolic enzyme (Barabasi and Oltvai 2004). With the availability of full-genome sequences since the mid-1990s (Fleischmann, Adams et al. 1995), organism-specific metabolic networks can be reconstructed to delineate all the genetic products and enzymatic reactions that determine the physiological and biochemical processes of a cell (Feist, Herrgard et al. 2009; Thiele and Palsson 2010). In brief, as depicted in Figure 1.1, the metabolic network reconstruction process is initiated by acquiring the genetic

content from the annotated whole genome sequence of a specific organism and/or strain. The available literature and metabolic pathway databases, such as KEGG (Kanehisa and Goto 2000), BRENDA (Schomburg, Chang et al. 2004), MetaCyc (Caspi, Foerster et al. 2006), are then utilized to facilitate the identification of enzymes and the metabolic reactions they catalyze together with the biochemical stoichiometry, cofactors and substrates information. Once the gene-protein-reaction association is established for each gene, the metabolic network can be reconstructed by integrating all the metabolic reactions (links) and metabolites (nodes). There have been multiple automated reconstruction tools available to aid the reconstruction process (Feist, Herrgard et al. 2009). However, extensive manual curation and different computational tools are needed to transform the draft reconstruction to a high quality metabolic network (Pitkanen, Rousu et al. 2010; Thiele and Palsson 2010).

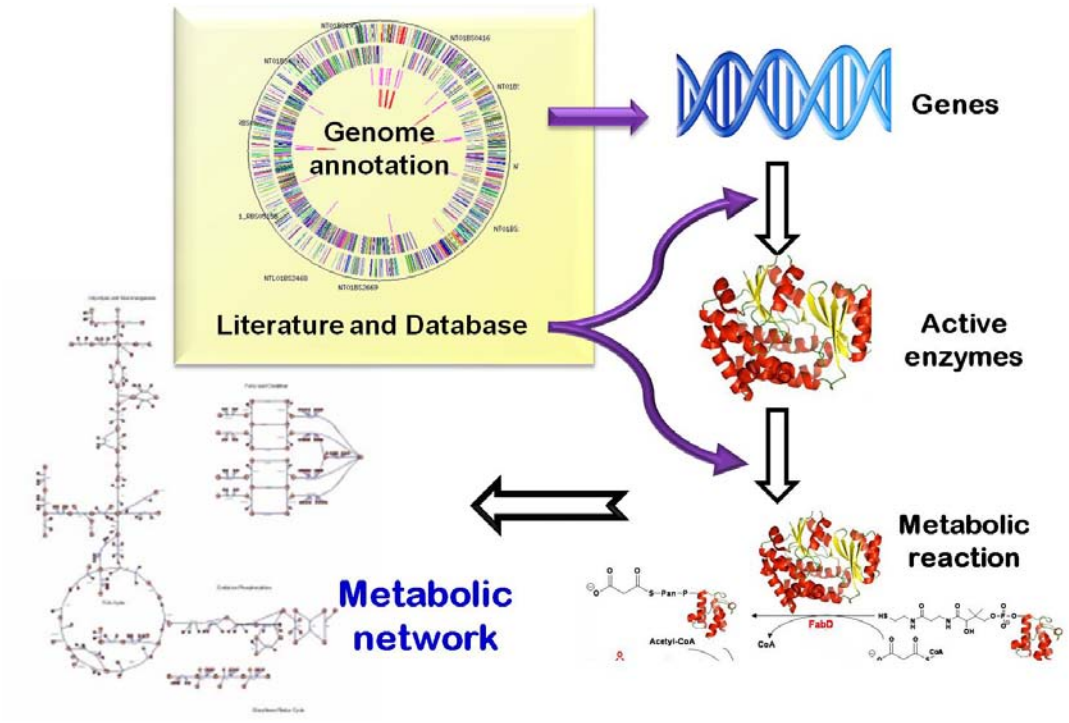


Figure 1.1 Metabolic network reconstruction

The metabolic network reconstruction of a given organism starts from its annotated genome where a parts list (genes) can be obtained. With the use of published literature and available data bases, the association

of the genome-encoded metabolic enzyme and its corresponding catalyzed metabolic reaction can be established. Finally, a metabolic network can be reconstructed to contain all the gene-enzyme-metabolic reaction associations.

Currently, such genome-scale models have been reconstructed for more than 30 microbes (Durot, Bourguignon et al. 2009), and for the generic mouse (Sheikh, Forster et al. 2005; Selvarasu, Wong et al. 2009; Selvarasu, Karimi et al. 2010), cattle (Seo and Lewin 2009) and human genomes (Duarte, Becker et al. 2007; Ma, Sorokin et al. 2007). These reconstructed metabolic networks provide valuable insights into the cellular metabolism and behaviors under various conditions.

1.1.2 Flux balance analysis

The predictive powers of metabolic network analyses are realized through mathematical computational modeling tools (reviewed in (Maertens and Vanrolleghem 2010)). One of the most extensively used network analysis tools is *flux balance analysis* (FBA), which is based on appropriate constraints to derive the steady-state metabolic capabilities of the system without the need for accurate kinetic data (Raman and Chandra 2009).

The main underlying principle of FBA is that under any given environmental condition, the organism will reach a steady-state that satisfies all the physiochemical constraints of the cell, such as the stoichiometric balance of mass, and energy (Kauffman, Prakash et al. 2003). Figure 1.2 briefly illustrates the steps of FBA (Schilling and Palsson 1998). Once the metabolic network has been reconstructed, the first step of FBA is to define the system boundary that includes all the internal metabolic reactions but not exchange processes across the system (Fig 1.2a, b). Once all the internal reactions and exchange mechanisms are identified, a mass balance equation, that is, an equilibrium of in-flux and out-flux is derived for each metabolite in the system (Fig 1.2c).

The set of coupled ordinary differential equations are then represented in a matrix format, where "**S**" is the stoichiometric matrix and "**V**" is the flux matrix (Fig 1.2d). Without any constraint, the solution for the metabolic fluxes is underdetermined (Fig 1.2e), that is, an infinitely large solution "space" is present. When additional constraints, e.g., the maximum uptake rate of a given substrate is applied, the totality of solutions that satisfy the constraints form an allowable solution space (Fig 1.2f) and a flux cone corresponding to the metabolic capacity under the specified constraints can be calculated (Schilling, Edwards et al. 2000). In turn, given a metabolic optimization objective, such as the maximal biomass production rate, FBA can identify the optimal set of the internal flux distribution through linear programming, which is a mathematical method for optimization of a linear objective function (Fig 1.2g).

Due to its simple fundamentals, FBA does not require kinetic data and can be computed very quickly. Since its initial development about 25 years ago (Fell 1986), FBA has proved to be a robust and diverse tool to predict the physiologically meaningful steady state to achieve a given metabolic objective, e.g., ATP production rate or growth rate (Raman and Chandra 2009). It also proved successful in predicting metabolic and cellular phenotypes with different perturbations, such as different growth conditions or various genetic manipulations in *E. coli* and other organisms (Ishii, Nakahigashi et al. 2007; Selvarasu, Wong et al. 2009). For example, the most well studied *E. coli* metabolic network has been extensively exploited for characterizing its cellular metabolism (Edwards, Ibarra et al. 2001; Covert and Palsson 2002; Fong, Marciniak et al. 2003) and its behavior under different perturbations and growth conditions (Fong, Marciniak et al. 2003; Ishii, Nakahigashi et al. 2007; Holm, Blank et al. 2010; Jozefczuk, Klie et al. 2010). In addition, the *in silico* models can also be used to infer the transcriptional and translational regulation and functional modules (Covert and Palsson 2002; Ravasz, Somera et al. 2002;

Stelling, Klamt et al. 2002; Herrgard, Lee et al. 2006). Last, but not least, FBA provides the ability to predict the pathways or core reactions of an organisms' metabolism with obvious implications for metabolic engineering (Almaas, Oltvai et al. 2005; Deutscher, Meilijson et al. 2006; Resendis-Antonio, Reed et al. 2007; Motter, Gulbahce et al. 2008; Ohlsen, Dandekar et al. 2008). Taken together, metabolic network analysis by FBA is a powerful modeling approach to investigate cellular metabolism, including the identification of potential enzyme targets from a systems biology perspective.

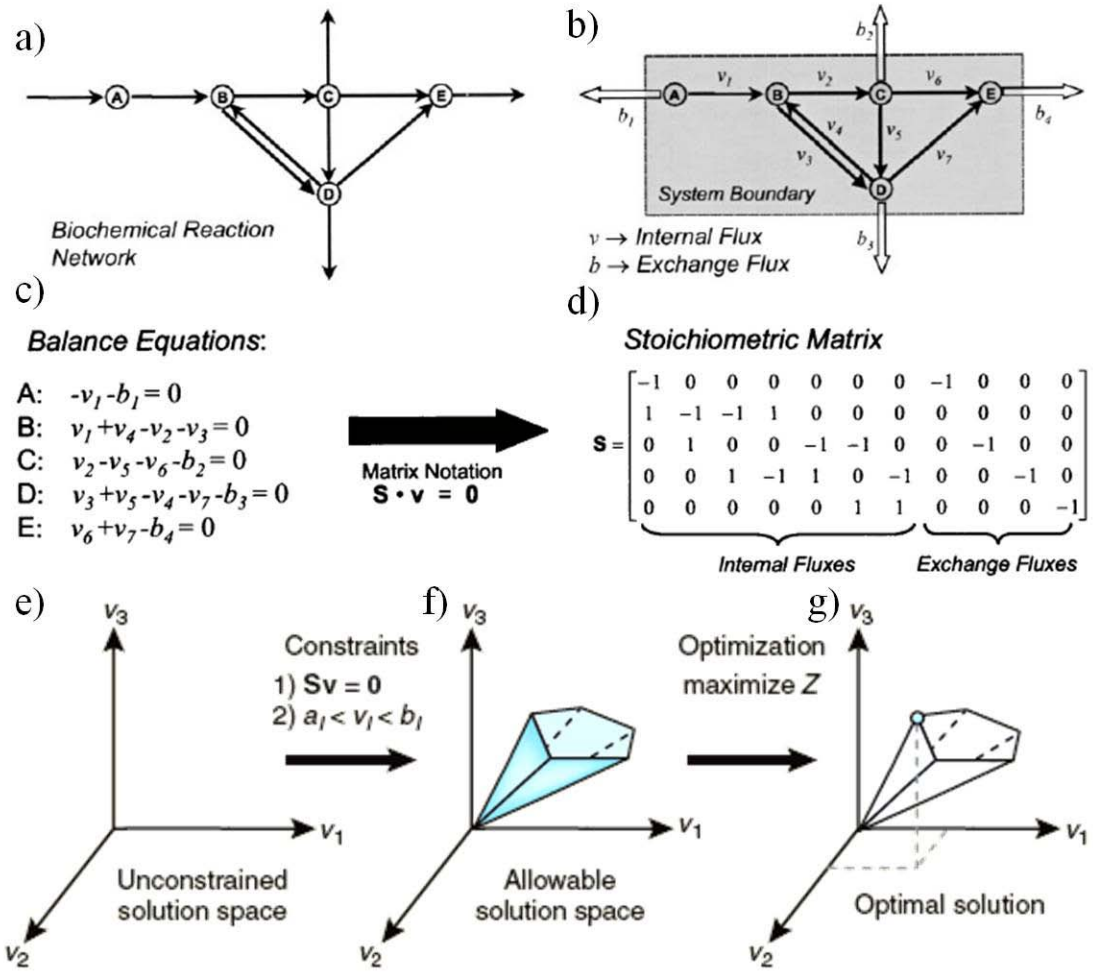


Figure 1.2 Flux balance analysis

a) A given metabolic network; b) the system boundary that includes all the internal metabolic reactions but not exchange processes across the system is defined; c) An equilibrium of in-flux and out-flux is

derived for each metabolite in the system; d) The set of coupled ordinary differential equations are then represented in a matrix format, where "**S**" is the stoichiometric matrix and "**V**" is the flux matrix; e) Without any constraint, the solution for the metabolic fluxes is underdetermined; f) With additional constraints, e.g., the maximum uptake rate of a specified substrate, all the solutions that satisfy the constraints form an allowable solution space; g) Given an optimization objective, such as the maximal biomass production rate, FBA can identify the optimal solution of flux distribution within the allowable solution space. Reproduced with permission from (Orth, 2010; Schilling, 1998).

However, traditional FBA without additional constraints or biological knowledge may not be enough to characterize the effect of mutations or shifts in the steady states (Segre, Vitkup et al. 2002). Indeed, during the last several years, additional constraints or alternative objective functions have been gradually incorporated into FBA to enrich its predictive capabilities and accuracy. For example, transcriptional and translational regulatory constraints have been considered through Boolean logic in the FBA model that captures the dynamic shifts of gene regulation over time (Covert and Palsson 2002; Mahadevan, Edwards et al. 2002). Besides regulatory constraints, the incorporation of additional thermodynamic constraints (i.e. energy balance) (Beard, Liang et al. 2002) and physicochemical constraints (Beg, Vazquez et al. 2007) also holds the potential to improve the modeling of metabolic behaviors that cannot be interpreted by traditional FBA. Of note, each method has its own advantages and limitations. Therefore, optimal FBA methods and objective functions should be carefully selected for investigation to achieve respective research aims.

1.1.3 Molecular crowding

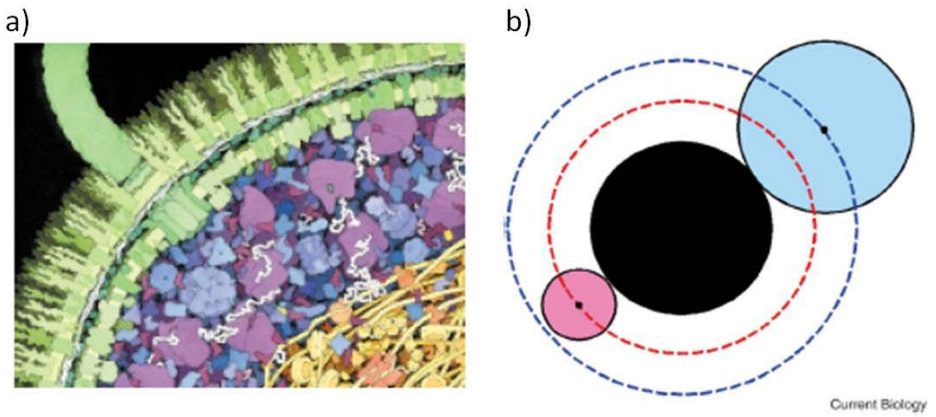


Figure 1.3 Molecular crowding

a) Representation of the approximate numbers, shapes and density of packing of macromolecules inside an *Escherichia coli* cell. Small molecules are not shown. Reproduced from with permission (Hoppert and Mayer 1999); **b**) Examples of excluded volume. In this figure, the spheres represent three different species of roughly spherical globular protein. The volume excluded by the black sphere to the red sphere is a spherical volume bounded by the dashed red line. The volume excluded by the black sphere to the blue sphere is a spherical volume bounded by the dashed blue line. Reproduced with permission from (Minton 2006).

Molecular crowding (MC) is an important characteristic of all cells' interior (Fulton 1982; Ellis 2001) (Fig 1.3a). The term MC means that the high concentration of intracellular macromolecules, such as DNA, RNA, polysaccharides and protein complexes, leads to a significant portion of cell volume physically occupied and thus, unavailable to other molecules (Ellis 2001; Acerenza and Graña 2006). Thus, the effect of MC is also known as the excluded volume effect (Fig 1.3b).

How crowded is the cell cytoplasm? In general, the volume fraction physically occupied by macromolecules ranges from 5% to 40% of the total volume (up to 400g/L) (Ellis and Minton 2003). For example, Zimmerman and coworkers estimated that the total concentration of protein and RNA inside an *E. coli* cell is in the range of 300-400g/L (Zimmerman and Trach 1991). The cytoplasm of a eukaryotic cell is also crowded. For example, red blood cells contain around 350g/L hemoglobin alone (Ellis 2001).

Due to the fundamental influence of thermodynamic activities on biochemical reactions, such crowded intracellular environments have profound effects on many aspects of cellular information processing steps. For example, molecular crowding can stabilize folded DNA, RNA and protein structures and conformations (Hancock 2004; Miyoshi and Sugimoto 2008; Kilburn, Roh et al. 2010); It can also result in the alteration of equilibrium and rate constants of biochemical reactions by up to several orders of magnitude (Minton 2001; Hu, Jiang et al. 2007; Zhou, Rivas et al. 2008); in addition, MC is very important to determine the diffusion/ mobility of metabolites and macromolecules. Therefore, cell signaling and communication as well as regulation could be also impacted by MC (Minton 1993; Bray 1998; Garcia-Perez, Lopez-Beltran et al. 1999; Goins 2007). Recently, there have been some attempts to study the effect of MC on the glycogenolysis enzymes (Chebotareva 2007).

Previous studies from our lab suggested that MC is a physiologically relevant constraint for determining the substrate uptake order of *E. coli* under various growth medium compositions (Beg, Vazquez et al. 2007), and essential for constraining the cellular metabolism especially in fast growing *E. coli* and *S. cerevisiae* cells (Vazquez, Beg et al. 2008; Vazquez, de Menezes et al. 2008). It is therefore likely that it also has a role in constraining the function of mammalian cellular metabolisms with potential contribution to cancer cell metabolism.

1.2 CANCER METABOLISM

Cancer is the second leading cause of death in the United States, resulting in one of every four deaths (Jemal, Siegel et al. 2007). More than 1 million people are diagnosed with cancer each year (Jemal, Siegel et al. 2007). However, not all types of cancer respond to current cancer

therapies, such as chemotherapy and radiation therapy. Moreover, many cancer patients develop drug resistance or suffer severe toxicity, limiting the effect of available chemotherapy and radiation therapy. Thus, the identification of additional novel effective cancer therapeutic approaches is urgently needed.

Cancer is a highly heterogeneous disease, both morphologically and genetically, with multiple genetic and regulatory abnormalities in the cell (Hanahan and Weinberg 2000). A major challenge in the pursuit of effective cancer therapy is the characterization of those common features that distinguish cancer cells from their normal counterparts, and how these differences can be exploited therapeutically to improve the specificity and effectiveness of cancer treatment. The hallmarks of cancer include growth signal independence, limitless replication potential, sustained angiogenesis, resistance to apoptosis, and tissue invasion and metastasis (Hanahan and Weinberg 2000), pathways of which some are already targeted by current cancer therapeutics. However, the molecular pathways underlying these cellular pathways are multifaceted and often redundant. Therefore, the alternative approach—targeting cancer metabolism, which is distinct from that of normal cells, represents a potentially major new therapeutic option.

1.2.1 The Warburg effect

Cellular energy metabolism, especially glucose metabolism, is one of the main processes that differentiate cancer cells from normal cells. Glycolysis is a catabolic process that converts 1 glucose to 2 pyruvates and also yields 2 ATPs. In normal cells, pyruvate is oxidized to CO_2 and H_2O generating 36 ATPs in the mitochondrial oxidative phosphorylation pathway when oxygen is abundant. Alternatively, in the absence of oxygen, pyruvate is catalyzed to lactic acid (or lactate) in the anaerobic glycolysis pathway (Fig 1.4a). As early as the mid 1920s, Otto Warburg

made the seminal discovery that rapidly proliferating ascites tumor cells utilize glycolysis at a much higher rate than normal resting cells, resulting in enhanced lactate secretion, even in the presence of oxygen. This phenomenon in mammalian cells is referred to aerobic glycolysis or the Warburg effect (Fig 1.4b) (Warburg 1956). Indeed, such distinct glucose metabolism has been found in most types of cancers (Kim, Gardner et al. 2005; Hsu and Sabatini 2008; Ortega AD 2009; Zhivotovsky and Orrenius 2009). Since glycolysis is far less efficient for ATP production than mitochondrial oxidative phosphorylation, cancer cells maintain high glycolytic fluxes, and such phenomenon has been clinically exploited to image cancer using the glucose analog 2-(^{18}F)-fluoro-2-deoxy-D-glucose (FDG) by positron emission tomography (PET) (Kroemer and Pouyssegur 2008).

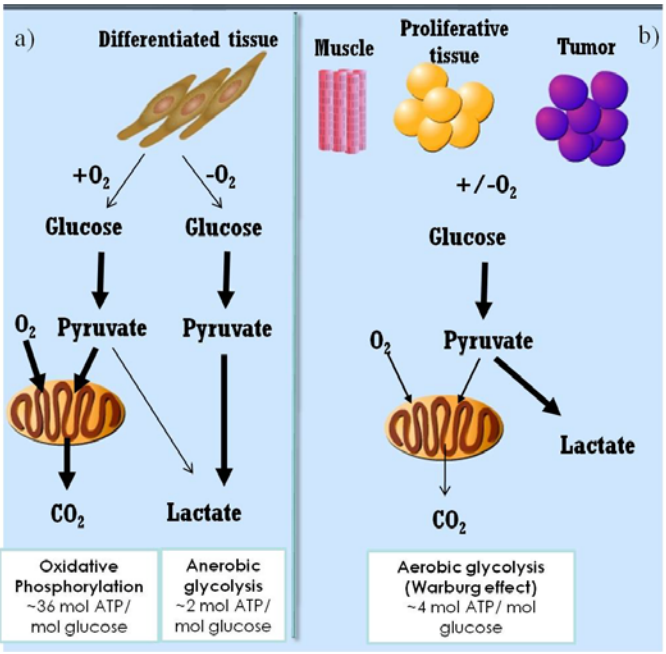


Figure 1.4 Glucose metabolism in normal and cancer cells.

a) In normal cells of differentiated tissue, glucose is converted to pyruvate, which is oxidized to CO_2 and H_2O generating 36 ATPs in the mitochondrial oxidative phosphorylation pathway when oxygen is abundant. In the absence of oxygen, pyruvate is catalyzed to lactic acid (or lactate) in the anaerobic glycolysis pathway, generating 2 ATP per glucose. b) In cancer cells and other proliferative tissues, glucose is converted to pyruvate, most of which is catalyzed to lactic acid even in the presence of oxygen,

a phenomenon called Warburg effect. Modified with permission from (Vander Heiden, Cantley et al. 2009).

Irrespective of the initiating mechanism, the molecular mechanisms underlying the Warburg effect are increasingly understood. Aerobic glycolysis is always associated with the up-regulation of glycolytic enzymes, such as glucose transporters (mainly GLUT-1), hexokinase-2 (HK-2), pyruvate kinase M2 (PK-M2) and lactate dehydrogenase A (LDH-A) (Fig 1.5) (Reviewed in (Gatenby and Gillies 2007; Hsu and Sabatini 2008; Locasale and Cantley 2010)). The increased glucose utilization through upregulated glycolytic enzymes generates metabolic intermediates such as glucose 6-phosphate, which is then used, e.g., for nucleic acid synthesis through pentose phosphate pathway to support the cells' fast proliferation. The abundant production of pyruvate is shunted from the TCA cycle in mitochondria to lactate production pathway through the up-regulation of pyruvate dehydrogenase kinase (PDK) and lactate dehydrogenase A (LDH-A). Part of pyruvate may enter a truncated tricarboxylic acid (TCA) cycle, the result of which stimulates lipid synthesis to sustain the cell membrane formation during division by exporting acetyl-CoA from the mitochondrial matrix to the cytoplasm (Fig 1.5) (Kroemer and Pouyssegur 2008; Frezza and Gottlieb 2009; Mayevsky 2009; Vander Heiden, Cantley et al. 2009; Kaelin and Thompson 2010).

Although inefficient for ATP production per glucose molecule, such distinct glucose metabolism has been found to convey proliferative advantages for tumor growth. First, cells with aerobic glycolysis can be more resistant to hypoxic or anoxic environments, which would otherwise be lethal for normal resting cells (Mazure, Dayan et al. 2006; Pouyssegur, Dayan et al. 2006). Second, aerobic glycolysis leads to accelerated lactate secretion, which can acidify the surrounding extracellular matrix and facilitate tumor metastasis and immune escape (Fischer,

Hoffmann et al. 2007; Swietach, Vaughan-Jones et al. 2007). In addition, aerobic glycolysis provides biosynthetic precursor molecules needed for cellular biomass production pathways, including glycerol and citrate for lipids; ribose sugars for nucleotides; nonessential amino acids

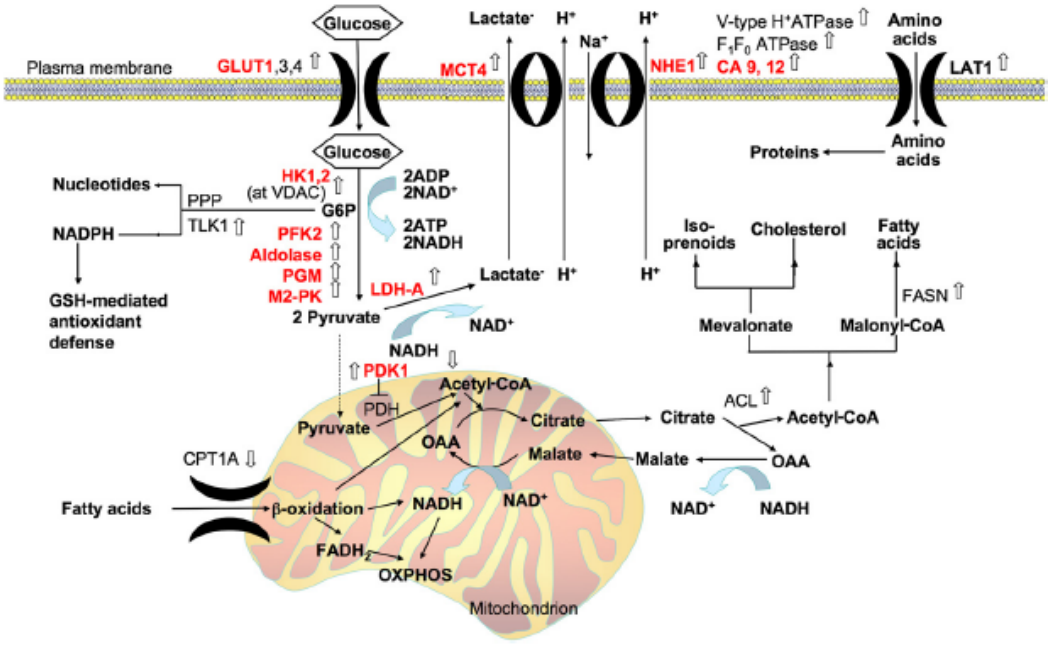


Figure 1.5 Metabolic reprogramming in cancer cells.

Small arrows pointing up or down indicate cancer-associated upregulation/activation or downregulation/inhibition of enzymes, respectively. Alterations indicated in red can be caused by the activation of HIF-1. GLUT, glucose transporter; GSH, glutathione; HIF, hypoxia-inducible factor; IDO, indoleamine 2,3-dioxygenase; HK, hexokinase; OXPHOS, oxidative phosphorylation; LAT1, L-type amino acid transporter 1; LDHA, lactate dehydrogenase isoform A; MCT, monocarboxylate transporter; PDH, pyruvate dehydrogenase; PDK, pyruvate dehydrogenase kinase; PFK, phosphofructokinase; PI3K, phosphatidylinositol 3-kinase; PGM, phosphoglycerate mutase; PKM2, pyruvate kinase isoform M2; PPP, pentose phosphate pathway; SCO2, synthesis of cytochrome c oxidase 2; TLK, transketolase; VDAC, voltage-dependent anion channel. Reproduced with permission from (Kroemer and Pouyssegur 2008).

and NADPH (DeBerardinis, Lum et al. 2008; Vander Heiden, Cantley et al. 2009). Moreover, aerobic glycolysis may protect cells from reactive oxygen species and apoptosis as well as premature senescence (Brand 1997; Brand and Hermfisse 1997; Hsu and Sabatini 2008; Kondoh

2008). Taken together, the Warburg effect seems to provide growth advantages to proliferating cancer cells.

1.2.2 Possible origins of the Warburg effect

Even though the growth advantages of the Warburg effect for cancer cells, and the mode by which cells switch from full oxidative phosphorylation to partial aerobic glycolysis are becoming clearer, whether the aberrant glucose metabolism is the cause or the consequence of the oncogenic transformation in cancers is still debated.

Originally, Warburg proposed that such aerobic glycolysis switch was due to permanent mitochondrial defects, preventing normal oxidative phosphorylation (Warburg 1956). While this may be true for some cancer cells, the generality of this hypothesis was challenged by the finding that some tumor cell lines maintain an intact capacity for oxidative metabolism (Guppy, Leedman et al. 2002; Rodriguez-Enriquez, Vital-Gonzalez et al. 2006). Gatenby and Gillies then proposed the hypoxia-induced metabolic reprogramming during early carcinogenesis within a hypoxic microenvironment, considering that partial oxygen pressure drops to very low values in the inner mass of a growing tumor (Helmlinger, Yuan et al. 1997; Gatenby and Gillies 2004). However, by now it is evident that the glycolytic switch is acquired very early in carcinogenesis, even before tumors encounter hypoxia (Vander Heiden, Cantley et al. 2009). For example, lung cancers and leukemic cells, which are surrounded by abundant oxygen, still metabolize glucose in part through the aerobic glycolysis pathway (Nolop, Rhodes et al. 1987; Gottschalk, Anderson et al. 2004), indicating that aerobic glycolysis cannot be solely attributed to hypoxia adaptation mediated through the key oxygen sensor, HIF-1.

A third possible cause is thought to be a consequence of oncogenic alterations, such as mutations/deletions in p53, MYC, and/or AKT (Semenza 2009; Feng and Levine 2010; Locasale and Cantley 2010). For example, the activation of hypoxia-inducible factor (HIF), a transcription factor that is activated by hypoxic stress but also by oncogenic, metabolic, and oxidative stress, often leads to the overexpression of the glucose transporters and various glycolysis pathway enzymes or isozyme subtypes explaining the increased glucose uptake and altered utilization (Sutphin, Giaccia et al. 2007; Semenza 2009). p53, one of the most frequently mutated genes in cancers, is both a positive regulator of mitochondrial respiration and a negative regulator of glycolysis (Feng and Levine 2010). In agreement with these findings, it has recently been shown that germline mutation of metabolic enzymes of the TCA cycle (e.g., isocitrate dehydrogenase (IDH), succinate dehydrogenase (SDH) and fumarate hydratase (FH)) in themselves can also lead to the aberrant glucose metabolism (Favier, Briere et al. 2009; Kaelin 2009; Sudarshan, Sourbier et al. 2009; Ward, Patel et al. 2010).

While the reasons listed above together can potentially explain the emergence of aerobic glycolysis in cancer cells, they cannot explain why proliferating primary lymphocytes also utilize similar aerobic glycolysis pathways (Wang, Marquardt et al. 1976). Indeed, proliferating lymphocytes are not alone. Human endothelial cells (Peters, Kamp et al. 2009), hepatocytes (Barbini, Rodriguez et al. 2007), vascular endothelial cells (Werle, Jahn et al. 2005), thymocytes (Brand 1997) and fibroblasts (Kittlick and Babin 1985) also exhibit a Warburg-like effect when they are proliferating rapidly in normoxic conditions. Taken together, these observations indicate that aerobic glycolysis is an inherent feature of all rapidly proliferating cells (Vander Heiden, Cantley et al. 2009).

1.2.3 Potential therapeutic significances of the Warburg effect

As mentioned in Part 1.2.1, the distinctiveness of cancer cell metabolism has been successfully exploited for cancer diagnostics through the metabolic imaging technique, PET. With renewed research interest in the Warburg effect an emerging field aims to explore cancer cell metabolism for therapeutic purposes: for identifying metabolic tumor markers and for targeted therapies.

There have been a number of glycolysis-related proteins investigated with regard to their potential as a prognostic tumor marker. For example, the over-expression of GLUT-1, the predominant glucose transporter in most tissues, is found to be correlated with ^{18}F-FDG uptake and poor prognosis, thus acting as a potential marker of radioresistance and poor response to chemotherapy in various human malignancies (Brophy, Sheehan et al. 2009; Legan, Luzar et al. 2009; Ayala, Rocha et al. 2010; Ohba, Fujii et al. 2010). Lactate dehydrogenase (LDH), the key enzyme converting pyruvate to lactate in the anaerobic glycolysis pathway, has also been shown to represent a marker of malignancy in gastric and lung cancers and tumors of other origins (Hann and Bombardieri 2000; Koukourakis, Giatromanolaki et al. 2003; Kimura, Yamaguchi et al. 2005; Kolev, Uetake et al. 2008). So called, oncometabolites (Ward, Patel et al. 2010), such as lactate or 2-hydroxyglutarate may also serve as serum markers for tumor diagnosis and monitoring therapeutic responses (Fischer, Hoffmann et al. 2007; Dang, White et al. 2010).

Besides tumor diagnostic markers, the aberrant tumor cell metabolism is also attractive as cancer's "Achilles' heel" for targeted therapy (Kroemer and Pouyssegur 2008). Recently, Cuezva and coworkers have reviewed various small compounds targeting the metabolism of cancer cells with potential clinical utility (Table 1.1) (Cuezva, Ortega et al. 2009). In brief, there are two major directions explored for metabolic targeting. The first is to inhibit glycolysis, leading to ATP depletion and cell death. For example, the glucose analog, 2-deoxy-D-glucose (2-

DG), can inhibit hexokinase, inhibit cell proliferation, and induce apoptosis in a variety of cancer cells *in vitro*; It can also sensitize the cancer cells *in vivo* and *in vitro* to radiation and chemotherapy treatments, and without affecting normal cells (Dwarakanath 2009; Farooque, Afrin et al. 2009). Similar to 2-DG, 3-bromopyruvate (3-BrPA) also targets hexokinase and abolishes the ATP production with potential antitumor activity and specificity (Buijs, Vossen et al. 2009; Ganapathy-Kanniappan, Vali et al. 2010; Qin, Xin et al. 2010).

Table 1.1 Small compounds targeting the metabolism of the cancer cell with potential clinical utility

#	Compound	Metabolic process	Target	Clinical trials
1	2-deoxyglucose	Glycolysis	Hexokinase	Phase I/II
2	3-bromopyruvate	Glycolysis	Hexokinase	Pre-clinical
3	Lonidamine	Glycolysis	Hexokinase	Phase II/III
4	5-thioglucose	Glycolysis	Hexokinase	Pre-clinical
5	Mannoheptulose	Glycolysis	Glucokinase	Pre-clinical
6	Iodoacetic acid	Glycolysis	GAPDH	Pre-clinical
7	α-chlorohydrin	Glycolysis	GAPDH	–
8	Ornidazole	Glycolysis	GAPDH	–
9	Oxalate	Glycolysis	PK	Pre-clinical
10	Somatostatin	Glycolysis	PK	Phase I–III
11	Dichloroacetate	Pyruvate oxidation	PDK1	Phase I/II
12	Glufosfamide	Glycolysis/PPP	Glucose transporter	Phase I/II
13	6-aminonicotinamide	PPP	G6PD	Pre-clinical
14	Oxythiamine	PPP	TK	Pre-clinical
15	Arsenite	apoptosis	MOMP	Pre-clinical
16	Betulinic acid	apoptosis	MOMP	Phase I/II
17	Jasmonates	apoptosis	MOMP	Pre-clinical
18	Soraphen A	Fatty acid synthesis	ACC	Pre-clinical
19	Cerulenin	Fatty acid synthesis	FAS	Pre-clinical
20	Etomoxir	β-oxidation	CPT1	Pre-clinical

GAPDH, glyceraldehyde 3-P-dehydrogenase; PK, pyruvate kinase; PDK1, pyruvate dehydrogenase kinase 1; G6PD, glucose-6-phosphate dehydrogenase; TK, transketolase; MOMP, mitochondrial outer membrane permeabilization; ACC, acetyl-CoA carboxylase; FAS, fatty acid synthase; CPT1, carnitine palmitoyl transferase 1; PPP, pentose phosphate pathway. Reproduced with permission (Cuezva, Ortega et al. 2009).

The second major function of potential cancer therapeutic compounds is to shift glycolysis towards mitochondrial respiration or glucose oxidation, e.g., by inhibiting LDH activity through the use of oxamate or siRNA (Coe and Strunk 1970; Fantin, St-Pierre et al. 2006). An intriguing and promising example is dichloroacetate (DCA), a clinically approved

drug for hereditary lactic acidosis disorders and a known inhibitor of pyruvate dehydrogenase kinase (PDK), which catalyzes the phosphorylation of pyruvate dehydrogenase (PDH) and therefore limits the amount of pyruvate entering the mitochondrial oxidative phosphorylation (Stacpoole 1989; Stacpoole, Henderson et al. 1998). It has been found recently that DCA can selectively kill cancer cells and established tumors by shifting metabolism from glycolysis to glucose oxidation and oxidative phosphorylation (Bonnet, Archer et al. 2007). DCA has also been shown to sensitize prostate cancer cells to irradiation (Cao, Yacoub et al. 2008). Currently, DCA is in phase I/II clinical trials (Cuezva, Ortega et al. 2009).

Although the above examples offer hope for specifically killing cancer cells through metabolic targeting, the real-world situation is more complicated than expected. Cancer cells may switch to alternative energy sources and non-glycolytic pathways for tumor cell survival and growth (Buzzai, Bauer et al. 2005). For example, drug and radiation resistant cancer cells use fatty acid oxidation to support mitochondrial oxidative phosphorylation when glucose becomes limited (Harper, Antoniou et al. 2002). This indicates that tumor cell metabolism has intrinsic plasticity and flexibility to maintain its robustness against external perturbations. This suggests that for the identification of the vulnerable points of cancer metabolism a systems-oriented metabolic network analysis approach may prove helpful for drug target discovery.

1.3 *S. AUREAS* ANTIBIOTICS RESISTANCE

As we will describe in **Chapter 3**, the full metabolic network reconstruction of mammalian cells in a tissue and cell type-specific fashion is not yet possible. Therefore, we turn to prokaryotic cells in which full metabolic network reconstruction is now routine. Specifically, to provide a

proof-of-principle of the potential of a systems-oriented metabolic network analysis for drug discovery targeting cell metabolism, we focus on the bacterium, *Staphylococcus aureus,* because of its virulence, its adaptation ability to different environmental conditions (Lowy 1998), and because its mortality now surpasses that of human immunodeficiency virus/AIDS in the United States (Klevens, Morrison et al. 2007).

S. aureus is a major hospital/community-acquired opportunistic pathogen, which can cause bacteremia, pneumonia, endocarditis, skin lesions, meningitis and other metastatic complications (Mongkolrattanothai, Aldag et al. 2009; Hulten, Kaplan et al. 2010; Wu, Wang et al. 2010). Although the use of existing antimicrobial agents, such as β-lactam antibiotics, sulfa drugs, tetracycline and clindamycin, has controlled some of these infections (Lowy 1998; David and Daum 2010; Herigon, Hersh et al. 2010; Hulten, Kaplan et al. 2010; Wu, Wang et al. 2010), the emergence of methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant *S. aureus* (VRSA) isolates underscores the need to identify new targets and molecules effective against multidrug-resistant strains of *S. aureus*.

1.3.1 Possible mechanisms of antibiotics resistance.

Resistant *Staphylococcus aureus*, especially methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Staphylococcus aureus* (VRSA)—just as deadly as MRSA but more challenging to treat, have been one of the three classes of emerging antibiotic-resistant pathogens as major threats to public health (Lam and Wunderink 2006; Moran, Krishnadasan et al. 2006; Fischbach and Walsh 2009). The evolution of increasingly antibiotic resistant species results from a multitude of factors that include wide-spread and inappropriate use of antibiotics, increase of international travels, and the extensive use of antibiotics as growth enhancers in

animal feed (Butaye, Devriese et al. 2003; Adaleti, Nakipoglu et al. 2010; Bosso, Mauldin et al. 2010; English and Gaur 2010; Martinez and Silley 2010).

Although the antimicrobial resistance mechanisms of *S. aureus* have not been completely clarified, there are at least three major forms of it. First, *S. aureus* may acquire through e.g., lateral transfer genes encoding enzymes that destroy the antibacterial agent before it can exert its effect. For example, when *S. aureus* is exposed to β-lactam antibiotics, such as penicillin, an extracellular enzyme, β-lactamase, is frequently synthesized to hydrolyze the β-lactam ring leading to its inactivation (Tenover 2006). Secondly, susceptible *S. aureus* can also acquire resistance through stepwise chromosomal mutations to modify the target site, or to produce an alternative metabolic pathway to bypass the drug action. For example, in the critical regions of DNA gyrase, the enzyme that serves as the target for quinolones, amino acids changes can reduce the affinity of quinolone to its targets (Moudgal and Kaatz 2009; Martinez and Silley 2010). In addition, bacteria can acquire resistance by mutation or transposon facilitated gene transfer from other organisms or strains to express efflux pumps that prevent the drug from reaching its intracellular targets (e.g., efflux of fluoroquinolones in *S. aureus*) (Martinezbeltran and Canton 1994; Lowy 2003; Moudgal and Kaatz 2009). Indeed, the fact that the efflux systems can mediate resistance to a variety of drug classes makes them very effective mechanisms of drug resistance.

Indeed, antibiotic resistance genes from pathogens only comprise a tiny fraction of the antibiotics resistome, which is the collection of all the antibiotic resistance genes and their precursors in pathogenic and non-pathogenic bacteria (Wright 2007; Wright 2010). Genomic approaches have revealed that resistance genes are present in all bacteria, including bacterial pathogens (D'Costa, McGrann et al. 2006; Fajardo, Martinez-Martin et al. 2008). The intrinsic

resistome may be a potential source of antibiotic resistance, given the appropriate selection pressure (Wright 2007). Thus, as rapidly as new antibiotics are introduced, *staphylococci* can very often develop efficient mechanisms to neutralize them, inevitably leading to fewer effective bactericidal antibiotics to treat these often life-threatening infections.

1.3.2 Current approaches to search for novel anti-Staphylococcus agents.

As noted above, the emergence of MRSA and VRSA argues for the development of new antibiotics that act on novel targets through new mode of actions. Equally important is the quick and cost effective approach to find unique vulnerable targets in *Staphylococcus aureus*, including its resistant strains.

One potential approach is to modify the existing antibiotics and scaffolds through synthetic tailoring to combat the mechanism of drug resistance. Indeed, most chemical scaffolds from which today's antibiotics have been derived were discovered between the mid-1930s and the early 1960s (e.g., β-lactams, glycopeptides, tetracyclines) (Newman and Cragg 2009). Just four such scaffolds -cephalosporins, penicillins, quinolones, and macrolides- account for ~73% of the new antimicrobials developed between 1981 and 2005 (Newman and Cragg 2007). During synthetic tailoring, the core of the antibiotics is left intact while the chemical groups at its periphery are modified to improve the drug's property to bypass the resistance mechanisms. In this sense, the analysis of the crystal structure of drug targets (e.g., modifications of β-lactams that attack the active site of its target) and synthesis of carbohydrate modified compounds (glycopeptides analogues with altered carbohydrates) are increasingly used to develop alternative agents (Tenover 2006). However, new antibacterial agents with different mechanisms of actions

are also needed, since it may be very difficult to outsmart organisms that have evolved for several billion years to be able to adapt to hostile environments.

Since the introduction of penicillin in the 1940s, significant effort has been invested in whole-cell screening for natural products, in which antibacterial activities are identified before the targets or modes of action are determined. Most natural product antibiotics usually come from soil actinomycetes (Fischbach and Walsh 2009). Recently, more underexplored ecological niches, such as terrestrial and marine symbioses, and bacterial taxa have started to be searched for new antibacterial agents (Newman and Cragg 2007; Wang, Kodali et al. 2007). For example, a deep-sea sediment sample yielded an actinomycete that produces a new antifolate scaffold, abyssomicins (Bister, Bischoff et al. 2004). Recently, a novel antibiotic, platensimycin, was also discovered from natural product screening to target the fatty acid biosynthesis pathways. More importantly, this promising antibiotic has displayed activity against various drug-resistant *S. aureus* strains (Wang, Soisson et al. 2006; Wang, Kodali et al. 2007).

An additional approach is to search for synthetic antibiotics with novel targets by novel strategies. For example, the synthetic molecules, oxazolidinones, are of increasing importance in the treatment of Gram-positive pathogens, including MRSA (Fischbach and Walsh 2009). As bacterial genomics emerged in the late 1990s, genomics-based technologies and synthetic libraries in the discovery of synthetic antibiotics have been of great interest with the promise to accelerate the discovery of new antibiotics. However, they have not lived up to their promise for at least two reasons: First, access to the target through the bacterial cell wall/ membrane may be limited. Secondly, pre-existing efflux system may limit target accessibility (Payne, Gwynn et al. 2007). Therefore, an unbiased whole-cell screen and novel strategies should be implemented for the success of synthetic antibiotics.

1.4 THESIS OUTLINE

Compared to normal resting cells, cancer cells exhibit distinct metabolisms, in which glycolysis with lactic acid secretion is utilized concomitantly with mitochondrial oxidative phosphorylation even in the presence of oxygen, a phenomena known as the Warburg effect or aerobic glycolysis. The understanding of the Warburg effect may provide insights that can lead to the development of alternative therapeutic options against human cancers. However, the origin of the Warburg effect is not fully understood.

In **Chapter 2** of this thesis, we introduce one potential origin for the Warburg effect -, molecular crowding-, and provide background information of its relevance for the metabolism of fast growing *E. coli* cells. As cancer cells also proliferate rapidly and exhibit similar aerobic glycolysis phenotypes to fast growing *E. coli* cells, in **Chapter 3** we aim to examine whether the MC is a relevant origin of cancer metabolism. Since the complete human metabolic network has not been fully reconstructed, we developed a reduced flux balance model of the human central metabolism that incorporates the molecular crowding (MC) constraint and the maximum glucose uptake constraint. We found that the model successfully captured the main characteristics of cancer metabolism (aerobic glycolysis), indicating that MC constraint may contribute to the origin of the Warburg effect. Next, we experimentally tested the model's predictions using mammalian cells from low to high growth rates as a proxy for MC alteration. We found that, consistent with the model, faster growing cells (i.e. cells with high MC) indeed have increased aerobic glycolysis, thus indicating MC as an important contributor to cancer metabolism.

Living organisms are generally robust against various perturbations, such as mutations, toxins, environmental alterations, and even drugs, the fact which necessitates the exploitation of systems approaches to identify the fragile points of robust microbial systems. In turn, the

systems-oriented approach may prove helpful for drug discovery, targeting e.g., the Warburg effect in tumor cells for therapeutic utilization. However, the reliable identification of enzyme targets within metabolic networks requires a completely reconstructed, high quality human metabolic network to be cell type specific, which is not yet attainable. Therefore, to provide a proof-of-concept we investigate this potential instead on the single cell pathogen, *Staphylococcus aureus,* where the metabolic network of the strain N315 have previously been reconstructed (Becker and Palsson 2005; Heinemann, Kummel et al. 2005).

As there are various *S. aureus* strains in terms of pathogenesis, antibiotic susceptibility, virulence and genome sequences, in **Chapter 4**, we first asked if it is sufficient to reconstruct the metabolic network on only one *S. aureus* strain to identify generic antimicrobial targets. We performed metabolic reconstructions and subsequent flux balance analyses of those 13 multidrug-resistant and sensitive *S. aureus* strains, whose genome sequences were available. We identified 70 single enzymes and 54 pairs of enzymes whose corresponding metabolic reactions are predicted to be unconditionally essential for growth. Of these, 44 single enzymes and 10 synthetic lethal enzyme pairs proved to be common to all 13 *S. aureus* strains, including many that had not been previously identified as being essential for growth by gene deletion experiments in *S. aureus.* We thus conclude that metabolic reconstruction and *in silico* analyses of *multiple* strains of the same bacterial species are needed for species-specific antibiotic target identification.

Potential target identification is the very first step toward antibiotic drug discovery. To validate the essentiality and druggability of the identified potential targets, we studied the common targets of 13 *S. aureus* strains in fatty acid biosynthesis and histidine biosynthesis pathways as a proof-of-principle study in **Chapter 5** and **Chapter 6**, respectively.

In **Chapter 5**, we focus on the predicted enzyme targets of the type-II bacterial fatty-acid biosynthesis (FAS II) pathway since most enzymes of this pathway are essential in the 13 *S. aureus* strains. We first identified a list of potential small molecule inhibitors through virtual screening targeting the individual enzyme targets in the FAS II pathways. We then showed experimentally that a subset of potential inhibitors could inhibit both enzyme activities *in vitro* and bacterial cell viability *in vivo*, which indicates the potential and accuracy of the systems oriented metabolic analysis in identifying drug targets.

In **Chapter 6**, we have further extended the investigation to the histidine biosynthesis pathway where we predicted the potential inhibitors through virtual screening, which were then validated directly through whole-cell assays, avoiding the time-consuming and often difficult enzyme assays. Of 18 representative compounds tested on 3 *S. aureus* strains and 2 *E. coli* strains, we showed that 13 compounds are inhibitors of some or all of the *S. aureus* strains, while 14 compounds are able to weakly inhibit growth in one or both *E. coli* strains. The high hit rate obtained from a fast virtual screen demonstrates the applicability of this novel strategy to the histidine biosynthesis pathway.

In **Chapter 7**, we conclude the four parts of the thesis with a perspective on how each part contributes to the current understanding of cellular metabolism and possible therapeutic opportunities.

The completion of this thesis has resulted in the following publications:

1. Vazquez A., **Liu J.**, Zhou Y., and Oltvai ZN. (2010) Catabolic efficiency of aerobic glycolysis: The Warburg effect revisited. BMC Systems Biology, 4(1):58. PMID: 20459610

2. Henriksen ST., **Liu J.**, Estiu G., Oltvai ZN. and Wiest O. (2010) Identification of novel bacterial histidine biosynthesis inhibitors using docking, ensemble rescoring and whole cell assays. Bioorg. & Med. Chem., 18(14):5148-56. PMID: 20573514

3. Shen Y*, **Liu J***, Estiu G, Isin B, Ahn YY, Lee DS, Barabási AL, Kapatral V, Wiest O, Oltvai ZN.(2010) Blueprint for antimicrobial hit discovery targeting metabolic networks. Proc Natl Acad Sci U S A. 107(3):1082-7. PMID: 20080587 (Shen Y. and Liu J. contributed equally to the paper.)

4. Lee DS, Burd H, **Liu J**, Almaas E, Wiest O, Barabási AL, Oltvai ZN, Kapatral V. (2009) Comparative genome-scale metabolic reconstruction and flux balance analysis of multiple Staphylococcus aureus genomes identify novel antimicrobial drug targets. J. Bacteriol. 191(12):4015-24. PMID: 19376871

2.0 MOLECULAR CROWDING REPRESENTS AN IMPORTANT BOUND IN RAPIDLY DIVIDING *E. COLI* CELLS

Understanding a cell's metabolism at the system level and obtaining quantitative predictions for cellular metabolic activities require the identification and modeling of the physicochemical constraints that limit the cell's metabolic capabilities at different growth conditions (Price, Reed et al. 2004; Vazquez, Beg et al. 2008). Molecular crowding (MC), also known as the excluded volume effect, is one such potential constraint in all cells' interior, since the high concentration of intracellular macromolecules (DNA, RNA, polysaccharides and protein complexes) can exert profound influences on many cellular processes. These include diffusion, reaction equilibrium and reaction rates, protein folding, and protein-protein- and protein-DNA interactions (Ellis 2001; Acerenza and Graña 2006; Zhou, Rivas et al. 2008), each of which can have an impact on the regulation of metabolic pathways. In this chapter, we will briefly summarize our previous results for developing an understanding the potential effects of MC on *E. coli* metabolism and explain why MC could be a potential metabolic constraint to the formation of the Warburg effect in cancer cells.

To investigate the potential effect of MC on system-level cell metabolism, we incorporated the MC constraint to the flux balance analysis framework (FBAwMC) (Beg, Vazquez et al. 2007; Vazquez, Beg et al. 2008;). Briefly, as introduced in Chapter 1, in FBA a cell's reconstructed metabolic network is mathematically represented by the stoichiometric matrix $\boldsymbol{S}$ with its element S_{mi} indicating the stoichiometric coefficient of metabolite m in reaction

i. The FBA states that in the stationary state, the fluxes {Schilling, , 1998} of the metabolic reactions consuming and producing a metabolite should be in balance,

$$\sum_{i=1}^{N} S_{mi} f_i = 0, \qquad [1]$$

where f_i denotes the flux of reaction *i*.

Due to the MC constraint, the volume occupied by one macromolecule will not be available to others, therefore there is only a limited solvent capacity for the allocation of metabolic enzymes. Thus, given that the enzyme molecules have a finite molar volume v_i only a finite number of them fit in a given cell volume *V*. Indeed, if n_i is the number of moles of the *i*th enzymes, then

$$\sum_{i=1}^{N} v_i n_i \leq V \qquad [2]$$

where the inequality sign accounts for the volume of other cell components and the free volume necessary for cellular transport as well. Eq. 2 represents a constraint on the enzyme levels n_i, potentially affecting their maximum attainable values and relative abundance. Dividing by cell mass *M*, we can reformulate this constraint in terms of the enzyme concentrations $E_i = n_i/M$ (moles per unit mass), resulting in

$$\sum_{i=1}^{N} v_i E_i \leq \frac{1}{C} \qquad [3]$$

where $C=M/V\approx 0.34\ g/ml$ is the cytoplasmic density of *E. coli* (Zimmerman and Trach 1991). As the reaction flux f_i can be directly affected or determined by the enzyme concentration E_i, the MC constraint (Eq. 2), and the enzyme concentration constraint (Eq. 3), thus, can be considered as a metabolic flux constraint

$$\sum_{i=1}^{N} a_i f_i \leq 1 \qquad [4]$$

Where $a_i=Cv_i/b_i$ quantifies the contribution of reaction i to the overall crowding, thus referred as "crowding coefficient" (Beg, Vazquez et al. 2007). The parameter $b_i=f_i/E_i$ is the coefficient of proportionality between reaction rate and enzyme concentration, thus determined by the reaction mechanism, kinetic parameters and metabolite concentrations.

To understand the relevance of the metabolic flux constraint (Eq. 4) at physiological conditions we need to estimate the crowding coefficient a_i from experimental data. At a first approximation we estimate b_i from the enzyme's turnover numbers in the BRENDA database (Schomburg, Chang et al. 2004) for about one hundred *E. coli* enzymes. The turnover numbers were found to vary over five orders of magnitude, from 10^{-2} to 10^{2} l/s. As C=0.34 g/ml and the molar volume of proteins are proportional to their molar masses (v=0.73M) (Lee 1983), the crowding coefficient for these one hundred enzymes can be calculated, resulting in 0.014±0.009 l/[mmol/g/h]. Because of the large enzyme turnover variations, the crowding coefficients are also distributed over a wide range, from 10^{-6} to 10^{0} l/[mmol/g/h] (Vazquez, Beg et al. 2008).

By superimposing this MC constraint as an additional metabolic flux constraint under conditions of aerobic growth in a glucose-limited medium, the application of FBAwMC on the *E. coli* MG1655 metabolic network is able to predict internal metabolic fluxes as a function of the growth rate. A subset of the FBAwMC-derived flux predictions in the central carbon metabolism are shown in Figure 2.1. At low growth rates, the model predicts (red line) higher fluxes in the TCA cycle and lower activities in the glycolysis pathway. In contrast, at high growth rates, the model predicts slightly decreased fluxes in the TCA cycles and increased fluxes in the glycolysis pathway. To be noted, in most cases the predicted fluxes (red line) are within

the experimentally determined range (black line), *especially* with high growth rate, indicating that MC constraint is most relevant in *fast* proliferating *E. coli* cells. These results indicate that even without knowing the precise value of the crowding coefficients, FBAwMC was able to capture the metabolic switch taking place from the TCA cycle to the glycolysis and acetate excretion pathway when shifting *E. coli* cells from slow to fast growth rate. Therefore, MC seems to represent an important, physiologically relevant bound to constrain the metabolism of fast growing *E. coli* cells.

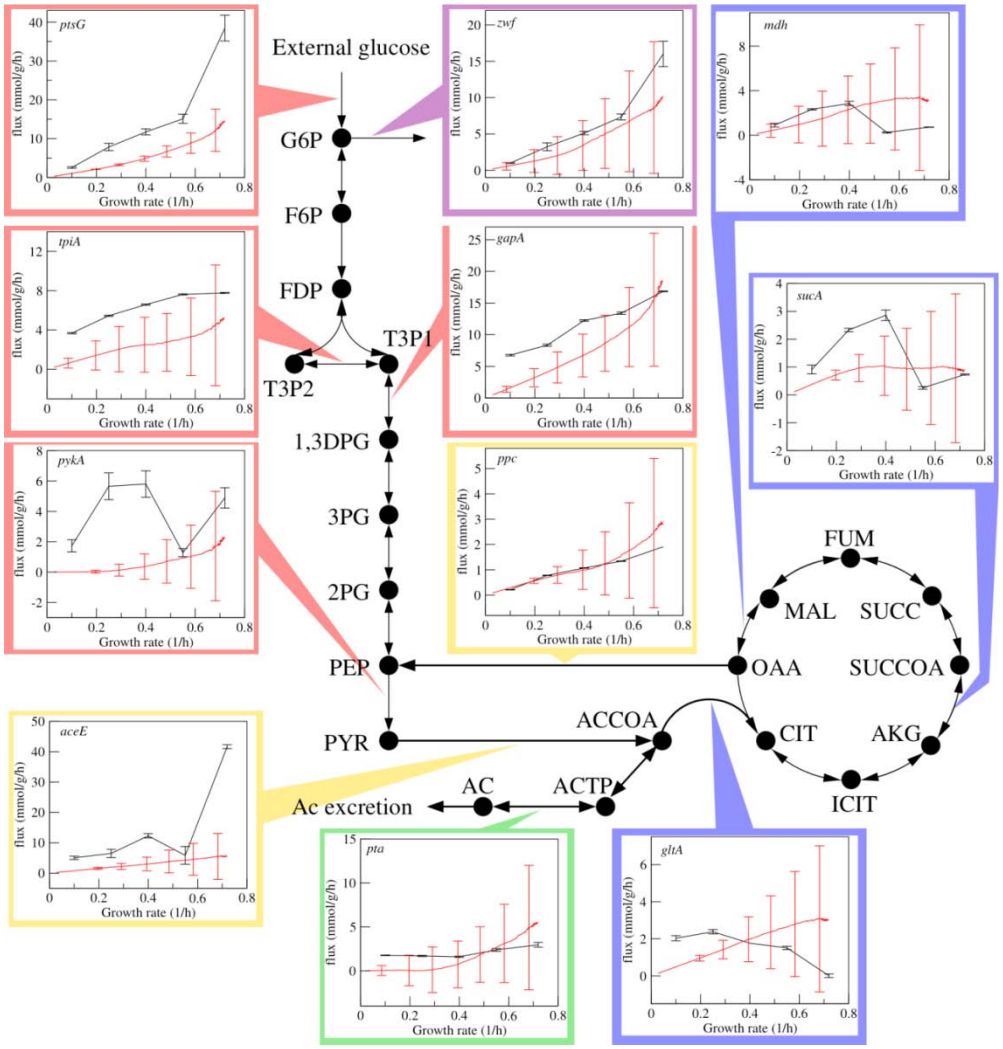


Figure 2.1: Predicted vs. measured metabolic fluxes in the *E. coli* central metabolism

Comparisons between the FBAwMC-predicted (red plots) and measured (black plots) fluxes as a function of growth/dilution rates for selected reactions in the central carbon metabolism of *E. coli*. The experimental flux measurements were performed at dilution rates 0.1, 0.25, 0.4, 0.55 and 0.72 h^{-1}. Selected reactions of glycolysis (red boxes), the first reaction of the pentose phosphate pathway (*zwf*) (magenta box), the TCA cycle (blue boxes), acetate excretion pathway (green boxes) and the reactions catalyzed by *ppc* and *aceE* connecting the glycolytic- and TCA pathways (yellow boxes) are shown. Reproduced with permission from (Vazquez et al., 2008).

3.0 MOLECULAR CROWDING POTENTIALLY CONTRIBUTES TO THE WARBURG EFFECT

3.1 INTRODUCTION

The Warburg effect has been found in most, if not all, types of cancers (Kim, Gardner et al. 2005; Hsu and Sabatini 2008; Ortega AD 2009; Zhivotovsky and Orrenius 2009). However, numerous studies have found that this distinct metabolic phenotype is not unique to cancer cells. For example, proliferating primary lymphocytes (Wang, Marquardt et al. 1976), human endothelial cells (Peters, Kamp et al. 2009), hepatocytes (Barbini, Rodriguez et al. 2007), smooth vascular muscle cells (Werle, Jahn et al. 2005), thymocytes (Brand 1997) and fibroblasts (Kittlick and Babin 1985) can also exhibit a Warburg-like effect when they are proliferating rapidly in normoxic conditions, indicating that there may be a direct correlation between aerobic glycolysis and rapid proliferation (Vander Heiden, Cantley et al. 2009).

As of yet, a systems-level interpretation of why all rapidly proliferating mammalian cells exhibit aerobic glycolysis has not been provided. As we detailed in Chapter 2, our previous work strongly imply that MC represents an important bound for fast growing *E. coli* and *S. cerevisiae* cellular metabolism.

Considering that cancer cells are also fast proliferating, the MC constraint is likely to affect cancer metabolism, as well. Therefore, we aim to assess the potential role of MC in cancer cell metabolism. To this end, we introduce a reduced flux balance model of ATP generation that

incorporates a glucose uptake capacity constraint and the MC constraint. The model indicates that the Warburg effect is a favorable catabolic state for rapidly proliferating mammalian cells with high glucose uptake capacity. We also experimentally examined human BJ fibroblast cell lines with serial oncogenic changes and stepwise increase in growth rates as a proxy for MC, as MC is only physiologically relevant in cells with high growth rates (Vazquez, Beg et al. 2008). We find that the fastest growing cell line exhibits the highest lactate excretion rate and the lowest mitochondrial respiration rate, which are consistent with the model predictions, thus inferring MC as a potential contributor to the Warburg effect.

3.2 RESULTS

3.2.1 Reduced flux balance model of cell metabolism

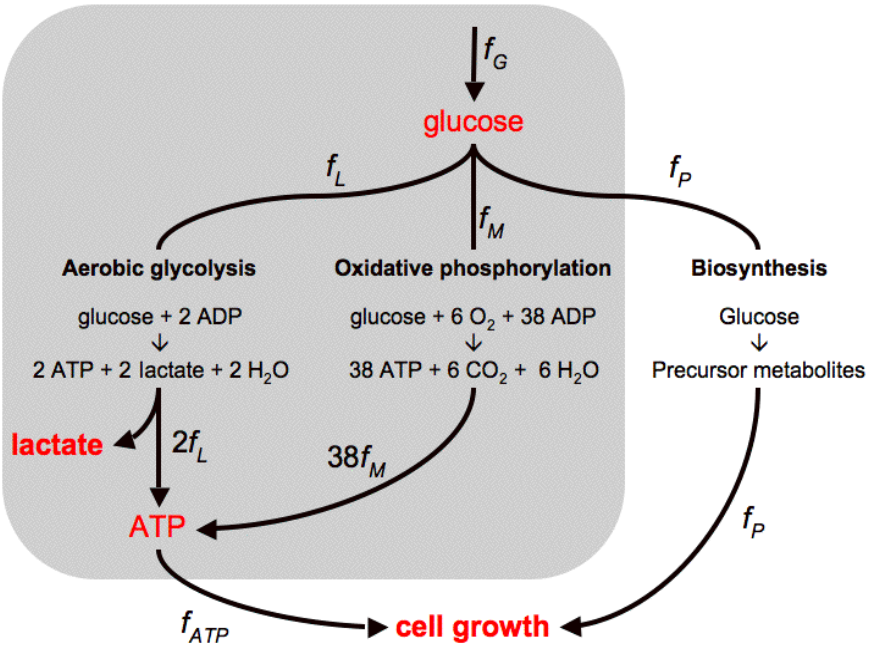


Figure 3. 1 Reduced model of cell metabolism.

Schematic representation of ATP generation pathways via aerobic glycolysis (glycolysis + pyruvate reduction to lactate) and oxidative phosphorylation (glycolysis + mitochondrial respiration). The variable f_G, denotes the glucose uptake rate, f_L and f_M, the components of the glucose uptake routed towards lactate excretion and respiration, respectively, and f_{ATP} the ATP production rate. Intermediate substrates of glucose catabolism are also used in a third process accounting for the production of precursor metabolites needed in anabolic processes (f_P). The pathways considered in our model are shown in the gray-shaded area, while f_P is treated as constant. Reproduced with permission from (Vazquez, Liu et al. 2010)

Figure 3.1 depicts a schematic model of glycolysis and mitochondrial respiration, the main pathways for ATP generation in cells (Vazquez, Liu et al. 2010). The glucose uptake flux (f_G) is partitioned into the flux of aerobic glycolysis (f_L), to oxidative phosphorylation (f_M), and to biomass synthesis pathways (f_P) (e.g., intermediate metabolites of the pentose phosphate pathway and the TCA cycle (Vander Heiden, Cantley et al. 2009)). Aerobic glycolysis represents glycolysis, converting glucose into pyruvate, and then pyruvate reduction by lactate dehydrogenase (LDH) in the cytosol, resulting in 2 moles of the end product lactate and 2 moles of ATP per mole of glucose. ATP generation through oxidative phosphorylation is decomposed into the generation of pyruvate through glycolysis, followed by the oxidation of pyruvate in the TCA cycle and the respiratory chain in mitochondria, yielding 38 moles of ATP per mole of glucose. In this reduced model, we assume that f_P has to be overall proportional with the available energy currencies (phosphate donors, mainly ATP) to enable cell growth. Thus, in the absence of a full scale kinetic model, in our modeling we do not take f_P into account and use f_{ATP} as a surrogate for the overall cell metabolic rate (Fig. 3.1).

Summing up the ATP production rate from aerobic glycolysis and mitochondrial respiration, we can obtain

(1) $$f_{ATP} = 2\, f_G + 36\, f_M = 2\, f_L + 38\, f_M ,$$

Our aim is to determine the optimal flux distribution (f_L, f_M) that provides maximum ATP production rate under the cell's metabolic constraints. Of these, the first metabolic constraint is associated with the maximum glucose uptake rate (or glucose uptake capacity)

(2) $f_L + f_M = f_G \leq F_G$,

where F_G is the maximum glucose uptake rate.

The second constraint, MC, quantified below, reflects on the high concentration of macromolecules within the cell's cytoplasm (Ellis 2001; Minton 2006), resulting in a limited solvent capacity for the allocation of metabolic enzymes. Enzyme molecules have a finite volume and the total sum of their volumes cannot exceed the cell volume. In our case, this constraint applies to the volume occupied by glycolytic enzymes, LDH and mitochondria. Enzymes associated with other pathways occupy a fraction of the intracellular volume as well. Nevertheless, this fraction simply restricts the amount of the cytoplasmic space available to glycolytic enzymes, LDH and mitochondria. More precisely, if V_G, V_L and V_M are the cell volume occupied by glycolytic enzymes, LDH and mitochondria, respectively, then

(3) $V_G + V_L + V_M \leq V_{ATP}$,

where V_{ATP} is the total cell volume reserved for the allocation of components of the ATP producing pathways. The occupied volumes V_G, V_L and V_M are proportional to the enzyme masses M_G, M_L and M_M, with $V_G=v_G M_G$, $V_L=v_L M_L$ and $V_M=v_M M_M$, where v_G, v_L and v_M are the specific volumes of glycolytic enzymes, LDH and mitochondria, respectively. In turn, the glycolytic rate (f_G), the lactate excretion rate ($2f_L$) and the mitochondrial ATP production rate ($36f_M$) are proportional to the mass of glycolytic enzymes, lactate dehydrogenase and mitochondria respectively, with $f_G=r_G M_G/V$, $2f_L=r_L M_L/V$ and $36f_M=r_M M_M/V$, where r_G is the glycolysis rate per unit of glycolytic enzyme mass, r_L is the rate of lactate production per mass of

LDH, r_M is the mitochondrial ATP production rate per unit of mitochondrial mass, and V is the cell volume. In these equations the product by the mass and the division by the cell volume takes into account that the rates r are commonly reported in the literature per unit of dry weight, while the pathway rates f are reported per unit of cell volume. Because of the interdependency of volume, mass and reaction rate, the volume constraint (3) can be translated to the metabolic constraint

$$(4) \qquad (a_L + a_G)\, f_L + (a_M + a_G)\, f_M \leq \phi_{ATP} \,,$$

where $\phi_{ATP} = V_{ATP}/V$ is the total volume fraction of the cell cytoplasm occupied by glycolytic enzymes, LDH and mitochondria, and the *crowding coefficients* $a_G = v_G/ r_G$, $a_L = 2\, v_L/ r_L$ and $a_M = 36\, v_M/ r_M$ quantify the occupied volume fractions per unit of glycolytic, lactate excretion and mitochondrial respiration rate, respectively. Using empirical data reported in the literature we have estimated the crowding coefficients. We obtain $a_G \approx 0.0027$ (min/mM), $a_L \approx 0.00023$ (min/mM) and $a_M \approx 0.10$ (min/mM), indicating that the mitochondria contributes about 5 and 50 times more to molecular crowding than glycolytic enzymes and lactate dehydrogenase, respectively. The unexpected consequences derived from this fact will be uncovered below.

In our modeling we assume that V_{ATP}, r_G, r_L and r_M are constant parameters that can be obtained from experimental estimates. Note though, that this is an approximation, as there may be regulatory mechanisms that under certain environmental or developmental conditions are capable of altering the amount of intracellular space allocated to ATP producing pathways and the activity of glycolytic enzymes, LDH and mitochondria.

Taken together the relevant metabolic optimization problem is as follows: maximize the ATP production rate (1) under the metabolic capacity constraints (2) and (4).

The optimal solution is graphically illustrated in Figure 3.2a and b (lines). At low glucose uptake rates ATP is entirely produced by the oxidative phosphorylation pathway and there is no lactate production (and excretion). This trend continues up to a threshold glucose uptake rate, f_1, when mitochondria occupy the entire cell volume fraction (V_{ATP}) available for ATP production pathways (Fig. 3.2a, blue dashed line, white background, *glucose-limited regime*).

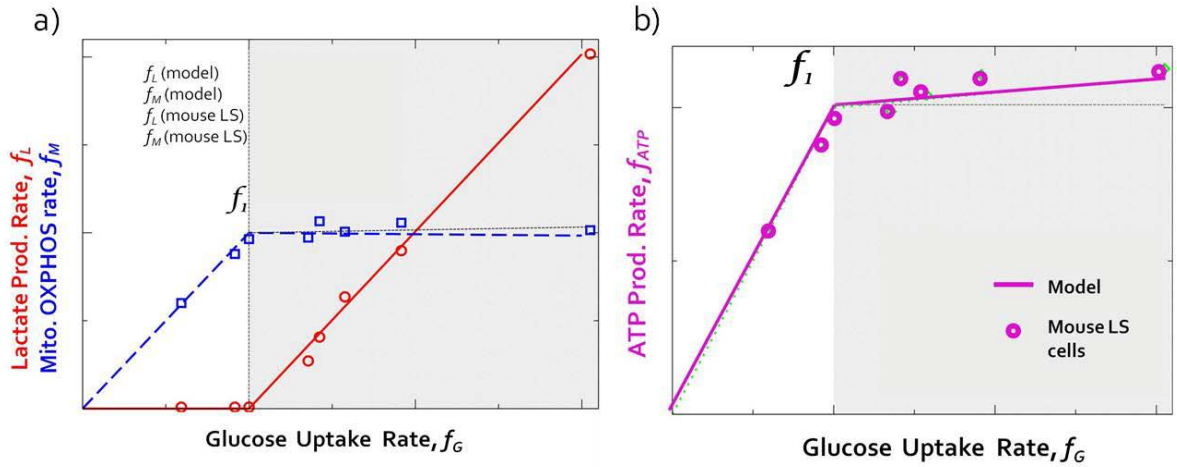


Figure 3.2. Catabolic regimes within proliferating mammalian cells.

a) Model-predicted (blue dashed line: mitochondrial OXPHOS flux; red solid line: lactate production flux) and measured (blue open square: mitochondrial OXPHOS; red open circle: lactate production flux in mouse LS cells (Sinclair 1974)) glycolysis and mitochondrial respiration fluxes as a function of the glucose uptake rate; gray dashed line: horizontal control line.

b) Model-predicted (purple solid line: ATP production flux) and measured (purple open circle: ATP production flux in mouse LS cells (Sinclair 1974)) ATP production fluxes as a function of the glucose uptake rate; the two different metabolic regimes are indicated by the white and gray backgrounds, respectively. All fluxes are reported in units of f_1, the glucose uptake rate threshold. Reproduced with permission from (Vazquez, Liu et al. 2010).

Beyond this threshold value the concentration of mitochondria, and therefore the rate of ATP production through oxidative phosphorylation, cannot be increased further. However, additional glucose uptake may be diverted toward pathway(s) less efficient in terms of ATP yield per mole of glucose. Indeed, the optimal solution predicts a metabolic switch, in which glucose uptake rates larger than f_1 leads to a linearly increasing lactate production (Fig. 3.2a, red solid line, gray

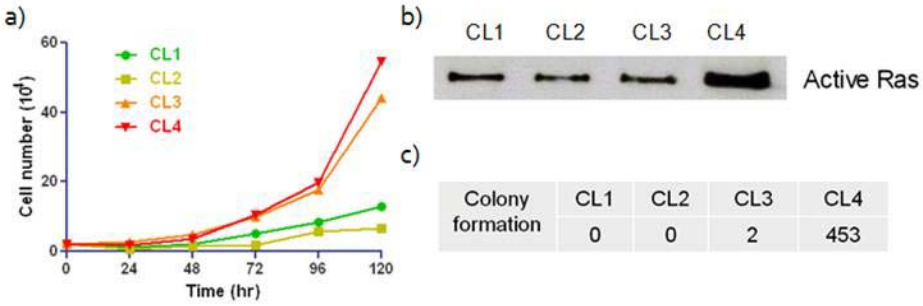
background, *space-limited regime*), and therefore a component of the ATP production is now derived from aerobic glycolysis. Furthermore, the increase in lactate production is accompanied by a slight, gradual reduction of the oxidative phosphorylation flux (Fig. 3.2a, blue dashed line, gray background). Therefore, due to the increased aerobic glycolysis rate ATP production can still be increased beyond f_1 (Fig. 3.2b, purple line). However, because aerobic glycolysis has a lower yield than oxidative phosphorylation, the ATP production rate increases with a lower slope beyond f_1 (Vazquez, Liu et al. 2010).

To address the validity of these predictions we first utilized previously reported experimental data for mouse LS cells (Sinclair 1974) (Fig. 3.2a,b), the growth rate of which were manipulated by their growth in chemostat cultures at different dilution rates, a protocol that is known to substantially vary the cells' glucose uptake rate. Surprisingly, the empirical lactate excretion plots follow the model predictions: at low glucose uptake rates there is no lactate excretion, but once a threshold is reached the lactate excretion rate increases linearly with the further increase in glucose uptake rate. Their oxidative phosphorylation (Fig. 3.2a, blue squares) and ATP production rate (Fig. 3.2b, diamonds) from the published experimental data (Sinclair 1974) can also be calculated (Vazquez, Liu et al. 2010) and also exhibit a good agreement with the model predictions (lines). Not only the model fits for the LS cells, but also does it fit glucose uptake and lactate secretion data for hybridoma cells (Gambhir, Korke et al. 2003) and a mixture of cancer and normal cells (Zu and Guppy 2004), indicating that the MC constraint is potentially a contributor of aerobic glycolysis at high glucose uptake rate (Vazquez, Liu et al. 2010).

3.2.2 MC potentially contributes to aerobic glycolysis in human fibroblasts with serial oncogenic changes and stepwise increase in growth rates

Based on the model predictions under the MC constraint, we hypothesized that the molecular crowding (MC) constraint may, at least partially, contribute to the Warburg effect when the glucose uptake rate is high. To confirm and establish the role of MC in cancer metabolism, we used mammalian cells with stepwise oncogenic changes.

To this end, first we examined human BJ fibroblast cells serially transduced with the catalytic domain of human telomerase (hTERT), SV40 large T (LT) and small T (ST) antigens, and an oncogenic allele of H-ras (H-RasV^{12}) (Hahn, Counter et al. 1999). Similar to previous studies using these cells (Hahn, Counter et al. 1999), CL1-[hTERT], CL2-[hTERT+LT], CL3-[hTERT+LT+ST], and CL4-[hTERT+LT+ST+H-Ras] cells displayed various growth rates, accompanied with increasing active Ras level and tumorigenic potential (Figure 3.3).



Colony formation	CL1	CL2	CL3	CL4
	0	0	2	453

Figure 3.3. H-Ras expression, colony formation and growth rate of normal and tumorigenic fibroblasts

(**a**.) Growth curves of CL1, CL2, CL3 and CL4 cells. For each measurement, continuously cultured fibroblasts, which were cultured in DMEM supplemented with 10% FBS, 25mM glucose and 4mM glutamine in a 37 °C humidified CO_2 incubator, were seeded at 2×10^4 cells in 12-well plates, and cell count data were determined by blinded counting of cells on a standard grid-patterned hemocytometer every 24h.

(**b**.) Western blot of active H-Ras in CL1-4 cells. CL1-4 cell lysates (1mg) were subjected to active Ras pull-down assay (Thermo Scientific) and fifty percent of the eluted samples (25µl, active Ras) were separated by 15% Tris SDS-PAGE gel and immunostained with H-Ras antibody (Thermo Scientific).

(**c**.) Soft agar assay. 10,000 cells of each cell type were seeded on 0.35% Top Agar and colonies were stained with 0.005% crystal violet and counted. All experiments were performed in duplicates and the mean numbers are shown in the table.

After confirming that CL1-4 did have different growth rates, we then characterized the metabolic features of these four cell lines. The glycolysis pathway utilizes the glucose that cells take up from the extracellular medium and converts it to pyruvate, which is then catalyzed by lactate dehydrogenase (LDH) to lactate, secreted to the extracellular medium resulting in the acidification. As it has been already shown by Ramanathan *et al.* that CL3 and CL4 have increased glucose uptake rates, compared to CL1 and CL2 (Ramanathan A 2005), we measured the lactate level in the extracellular medium and calculated the lactate production per total protein level of each cell line. It is evident that compared to the other cell types CL4 cells excrete a significantly higher levels of lactate than CL1 and CL2 cells at most time points tested (Fig. 3.4a). Of note, starting from 96h CL3 cells also display significantly higher extracellular lactate levels, a finding that is likely related to their increased growth rate that is similar to that of CL4 cells (Fig. 3.3a), and that has been shown to be independent of active H-Ras level.

Correspondingly, the LDH enzyme activities of CL3 and CL4 are significantly higher than that of CL1 and CL2 cells at most time points (Fig. 3.4b). However, LDH protein expression levels and sizes are not significantly different among the four cell types (Fig. 3.4c and 3.4d). These data confirm that lactate production increases with the rising glucose uptake rates in these cells (Ramanathan A 2005), and indicate that the increase in LDH activity is primarily achieved by posttranslational activation of the enzyme. We also compared the mitochondrial mass and membrane potential level of CL1-4 cells as a function of time at 48h, 72h, 96h and 120h after seeding. As mitochondrial membrane potential is usually generated by the activities of respiratory chain complex, it can be used as a proxy for mitochondrial respiratory activity. The

mitochondrial membrane potential of CL1-4 cells was assessed by staining with JC-1, a dye that is able to monitor mitochondrial membrane potential (largely) independent of mitochondrial mass (Cossarizza, Baccaranicontri et al. 1993). The mitochondrial content of the cell was assessed using MitoTracker Green (MTG) staining (Agnello, Morici et al. 2008).

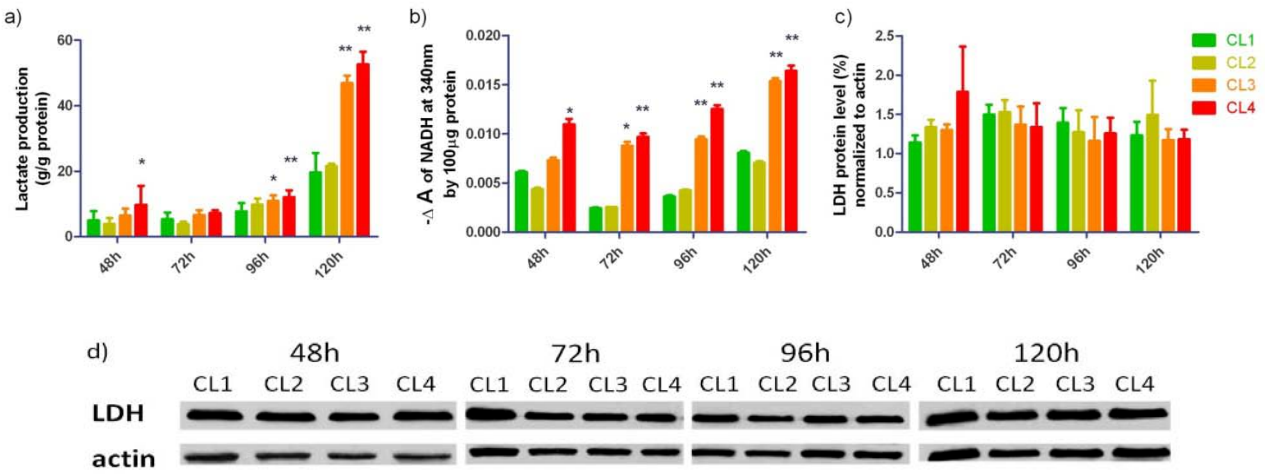


Figure 3.4 Glycolytic characteristics of normal and tumorigenic fibroblasts

(**a.**) Extracellular lactate levels normalized by the total protein concentration of CL1-4 cells;

(**b.**) Relative LDH activity in 100µg protein lysates of CL1-4 cells expressed as the signal drop of NADH fluorescence at 340nm in the first 30s. Mean ± SD of a minimum of three independent experiments are shown. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$ compared to CL1 at the corresponding time point, in a Student's t-test.

(**c.**) Relative LDH protein levels of fibroblasts as assessed by ImageJ.

(**d**.) LDH protein expression of fibroblasts. 50µg protein of each cell line lysates were subjected to 15% Tris SDS-PAGE gel and immunostained with LDH (Santa Cruz Biotech) and actin-specific monoclonal antibodies (Sigma).

After normalizing all the data to the value of CL1 at 48h, it is apparent that CL4 cells have higher membrane potential than CL1 at 48h (Fig. 3.5a), followed by a progressive drop to values lower than CL1 afterwards. Interestingly, the mitochondrial mass/content in these fibroblasts displayed similar alterations over the time (Fig. 3.5b). These results demonstrate that 48 h after seeding CL4 cells display both a decreased mitochondrial content and activity, followed by similar but lesser trends in the other cell types.

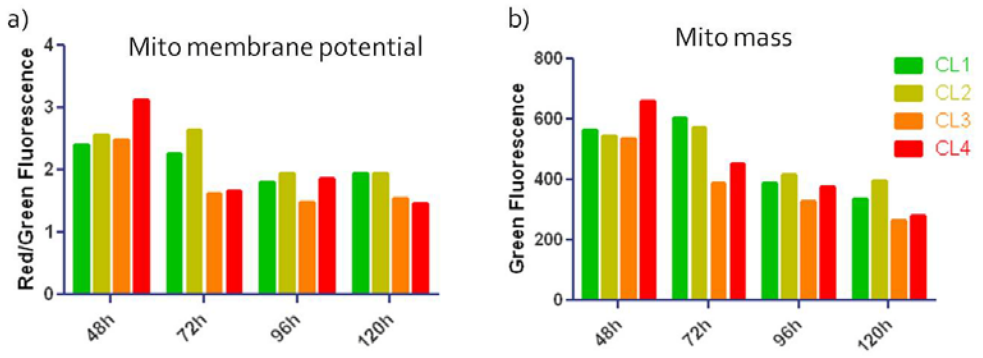


Figure 3.5 mitochondrial characteristics of normal and tumorigenic fibroblasts

(**a**.) Mitochondrial membrane potentials of the indicated cell types are shown, as determined by JC1 staining and subsequent flow cytometry analyses;

(**b**) Mitochondrial content/ mass of the indicated cell types are shown, as determined by MitoTracker Green staining and subsequent flow cytometry analyses.

As the model predictions were built on the flux of glycolysis and mitochondrial respiration, which is the rate of reactions or pathways, we next employed the Seahorse bioenergetic flux analyzer to indirectly measure the glycolysis flux (extracellular acidification rate, ECAR) and mitochondrial respiration flux (oxygen consumption rate) (Qian and Van Houten 2010). As shown in Figure 3.6, CL4 cells had displayed significantly higher glycolysis rate/ flux (Fig 3.6a) while lower mitochondrial respiration rate (Fig 3.6b) than slower growing CL1-3 cells, consistent with the results from previous glycolytic parameters and model predictions.

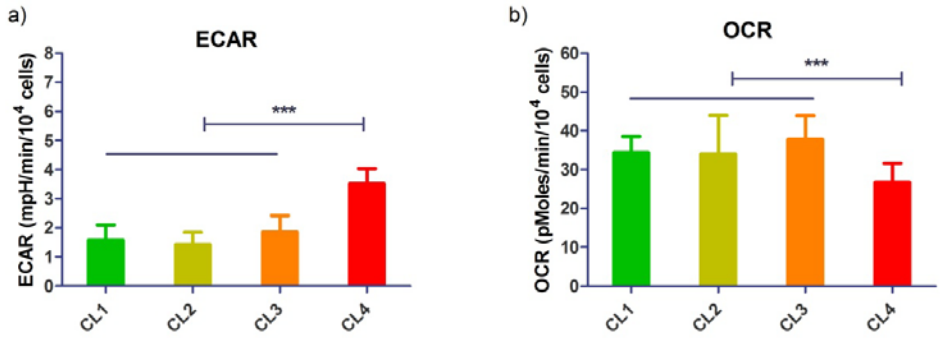


Figure 3.6: The glycolysis (ECAR) and mitochondrial respiration (OCR) flux of CL1-4

a) The extracellular acidification rate (ECAR) per cell measured by the Seahorse flux analyzer in quadruplicates;
b) The oxygen consumption rate (OCR) per cell measured by the Seahorse flux analyzer in quadruplicates. Star represents statistic significance in paired student's t-tests. ***: $p<0.001$.

3.2.3 MC potentially contributes to aerobic glycolysis in human fibroblasts different growth rates due to passage numbers

Compared to CL1-3, CL4 has an additional oncogenic H-Ras, which is known to be able to alter energy metabolism, both glycolysis and oxidative phosphorylation, in multiple ways (Biaglow, Cerniglia et al. 1997; de Groof, Lindert et al. 2009; Yang, Wang et al. 2010). To eliminate the possible contribution from oncogenes other than the growth rate itself, next we elected to utilize CL4 early vs. late passage cells as another experimental system, since the growth rate of transformed cells are shown to increase gradually during prolonged culture *in vitro* (de Groof, Lindert et al. 2009). Indeed, CL4 late passage cells grew faster than CL4 early passage cells (Fig. 3.7a) even though they have relatively similar levels of active H-Ras (Fig. 3.7b) and tumorigenic potentials as determined by colony formation on the soft agar (Fig. 3.7c). This indicates that additional genetic/ epigenetic changes on the cell line play a role in the observed growth rate difference, but they are not sufficient to alter the tumorigenecity of the cell line.

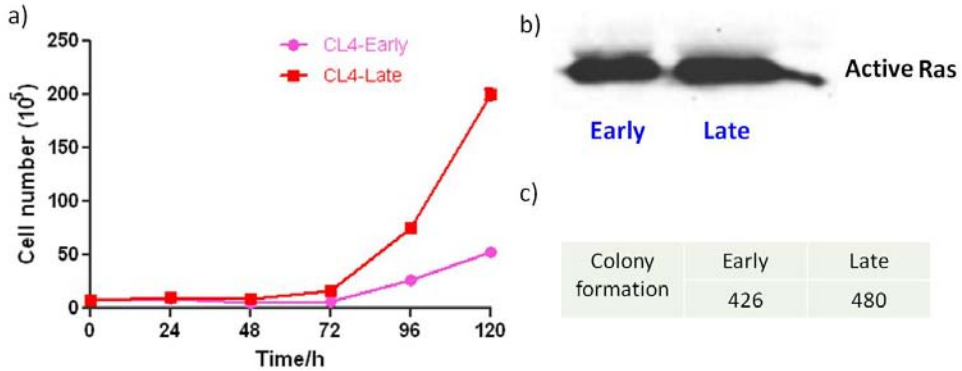


Figure 3.7. Growth rate, H-Ras expression and colony formation of CL4 early and late

passage cells.

(**a**.) Growth curves of CL4 early and late passage cells. For each measurement, continuously cultured fibroblasts, which were cultured in DMEM supplemented with 10% FBS, 25mM glucose and 4mM glutamine in a 37 °C humidified CO_2 incubator, were seeded at 2×10^4 cells in 12-well plates, and cell count data were determined by blinded counting of cells on a standard grid-patterned hemocytometer every 24h.

(**b**.) Western blot of active H-Ras in CL4 early and late cells. CL4 early and late cell lysates (1mg) were subjected to active Ras pull-down assay (Thermo Scientific) and fifty percent of the eluted samples (25µl, active Ras) were separated by 15% Tris SDS-PAGE gel and immunostained with H-Ras antibody (Thermo Scientific).

(**c**.) Soft agar assay. 10,000 cells from each cell lines were resuspended in 2 ml of 0.35% top agar in complete DMEM growth medium and seeded in six-well plates previously filled with 2 ml of 0.7% basal agar in complete growth medium. After 2 weeks, 1ml 0.005% crystal violet was added to each well to stain the colonies for 2h and colonies in each well were counted.

Similar to the first system, after normalized to total protein level, the faster growing CL4 late passage cells still took up glucose and secreted lactate significantly more than the slower CL4 early cells at almost all time points examined (Fig. 3.8a and b). The LDH enzyme activities of

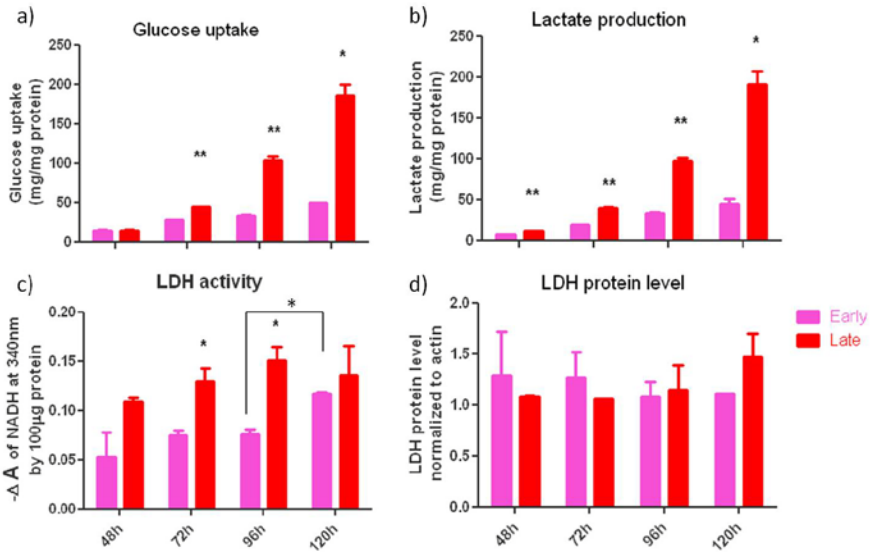


Figure 3.8 Glycolytic characteristics of CL4 early and late cells

(**a**.) Glucose uptake normalized by the total protein concentration of CL4 early and late cells at different time points;

(**b.**) Extracellular lactate levels normalized by the total protein concentration of CL4 early and late cells;

(**c**.) Relative LDH activity in 100µg protein lysates of CL4 early and late cells expressed as the signal drop of NADH fluorescence at 340nm in the first 30s. Mean ± SD of a minimum of three independent experiments are shown. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$

(**d**.) Relative LDH protein levels of CL4 cells. 50µg protein of each cell line lysates were subjected to 15% Tris SDS-PAGE gel and immunostained with LDH (Santa Cruz Biotech) and actin-specific monoclonal antibodies (Sigma).

CL4 late passage cells are significantly higher than that of CL4 early cells at most time points (Fig. 3.8c). Of note, at 120h CL4 early cells also display a significantly higher LDH activity, a finding that is likely related to their increased growth rate at 120h (Fig. 3.7a). However, LDH protein expression levels are not significantly different in the two cell lines (Fig. 3.8d). These data confirm that the increase in LDH activity is primarily achieved by posttranslational activation of the enzyme, and that faster cells tend to be more glycolytic.

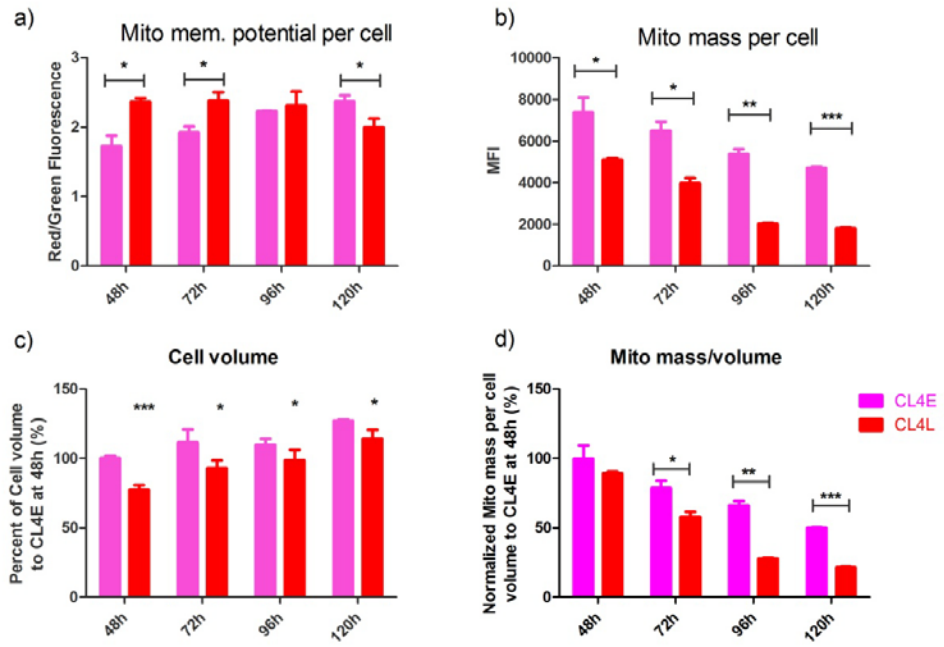


Figure 3.9 Mitochondrial characteristics of CL4 early and late cells

(**a**.) Mitochondrial membrane potentials of CL4 early and late are shown, as determined by JC1 staining and subsequent flow cytometry analyses;

(b.) Mitochondrial content/ mass of CL4 are shown, as determined by MitoTracker Green staining and subsequent flow cytometry analyses;

(c.) Cell volume as estimated from forward scatter in the flow cytometry analyses;

(d.) Mito mass per cell volume as calculated from dividing mitochondrial mass in (b.) by cell volume (c.). Star represents statistic significance in paired student's t-tests. *: $p<0.05$; **: $p<0.01$; ***: $p<0.001$.

In terms of the mitochondrial mass and membrane potential level, CL4 late passage cells have higher membrane potential per cell than CL4 early cells initially at 48h and 72h (Fig. 3.9a), followed by a progressive drop to values lower than CL4 early cells after 96h. Interestingly, the mitochondrial mass/content of CL4 late cells proved lower than that of CL4 early cells all the time (Fig. 3.9b). The decreased mitochondrial respiration activity in CL4L cells is confirmed by this significantly lower oxygen consumption rate (Fig. 3.10b) and the diminished cell volume occupied by the mitochondria mass/ content (Fig. 3.9d). Of note, all the glycolytic parameters at different time points are normalized to the total protein level, while the glycolysis rate is quantified by the Seahorse flux analyzer normalized to total cell number. Thus, non-significant difference between CL4E and CL4L in the glycolysis rate (Fig. 3.10a), may be attributed to the smaller volume of CL4L (Fig. 3.9c); that is, given the same cell volume, CL4L will exhibit higher glycolysis rate per cell. Taken together, the observed alterations in glycolysis activities and mitochondrial parameters support the model predictions: an increase of lactate excretion rate and decline of mitochondrial respiration rate with an increasing growth rate, a proxy of MC levels.

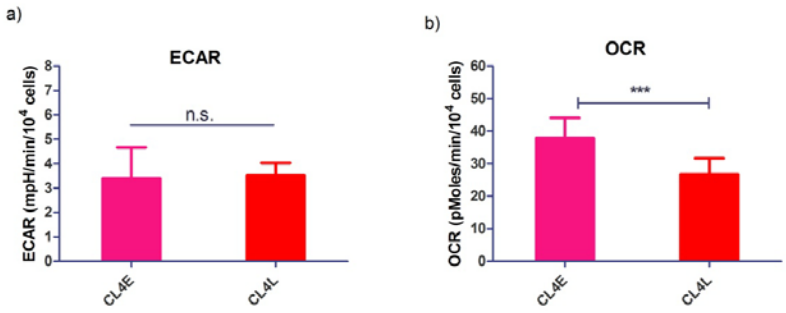


Figure 3.10 The glycolysis (ECAR) and mitochondrial respiration (OCR) rate of CL4 early and late cells

a) The extracellular acidification rate (ECAR) per cell measured by the Seahorse flux analyzer in quadruplicates;
b) The oxygen consumption rate (OCR) per cell measured by the Seahorse flux analyzer in quadruplicates. Star represents statistic significance in paired student's t-tests. *: $p<0.05$; **: $p<0.01$; ***: $p<0.001$.

3.3 DISCUSSION

To provide a systematic view of why and how proliferation itself is associated with an increase of aerobic glycolysis and a decline of mitochondrial respiration, is fundamentally important to understand the origin of the Warburg effect and may suggest novel insights for cancer therapy.

In our study, we first developed a reduced flux balance model of ATP production (Fig. 3.1) and predicted the existence of two different metabolic regimes in mammalian cells, depending on whether the cell metabolic rate is limited by the glucose uptake or by the MC constraint. In conditions where glucose transport limits the cell's metabolic activity the model predicts that the optimal solution is to produce ATP entirely via mitochondrial respiration; however, at high glucose uptake rates the optimal solution is characterized by an abrupt increase of the flux towards lactate production (and excretion) and a gradual decrease of mitochondrial respiration (Fig. 3.2). This second regime is determined by the existence of the MC constraint, a limited cytoplasmic solvent capacity for allocating the components of the ATP generating pathways. For glucose uptake rates exceeding the threshold level, f_1, the cell cannot further increase the concentration of mitochondria and thus switch to aerobic glycolysis lacate production pathway in order to match the increased glucose uptake rate. The model predictions have been indirectly supported by the experimental data from mouse LS cells (Sinclair 1974),

hybridoma cells (Gambhir, Korke et al. 2003) and a mixture of cancer and normal cells (Zu and Guppy 2004).

Based on this model prediction, we proposed that MC is a potential contributor to the Warburg effect. To provide additional experimental support to this hypothesis, we first used human foreskin fibroblasts with stepwise oncogenic transformation to mimic cells with increasing growth rates as an indirect measure of MC levels (Vazquez, Beg et al. 2008). We found that the fastest growing cell line exhibited increased lactate production, lactate dehydrogenase activity, and the glycolysis rate (ECAR), while it displayed a declined mitochondrial respiration activity (Fig. 3.3-7), consistent with what others have found using the same model system (Ramanathan A 2005). In addition, we showed that such metabolic alterations were not only due to the expression of the oncogenic Ras, since CL4E and CL4L, having the same active Ras level but different cell passage numbers, still followed a similar upregulation of glycolytic activities and downregulation of mitochondrial respiration rate as their proliferation capabilities differed (Fig. 3.8-10). The results are further supported from another similar experiments, where oncogenic Ras/E1A transformed low passage murine embryonic fibroblast (MEFs) (Ras-LP) also displayed higher mitochondrial respiration activity and lower lactate secretion compared to more rapidly dividing high passage MEFs (Ras-HP) (de Groof, Lindert et al. 2009). Yet, it is noteworthy, that the glycolytic rate of CL4L per cell is not different from that of CL4E cells, which disagrees with the model predictions and measured metabolic phenotypes, that faster growing cell have higher glycolytic rate. Indeed, the model assumes that the volume fractions occupied by ATP production pathway enzymes, that is, glycolytic enzymes and mitochondria, are constant. Given that each mitochondria alone hosts metabolic enzymes that function in the TCA cycle, heme biosynthesis, β-oxidation of fatty acids, steroidogenesis,

certain amino acids metabolisms and Fe/S cluster formations (Reichert and Neupert 2004), the total mass and thus occupied volume of glycolytic enzymes could be neglected, compared to mitochondria. In contrast to the modeling assumption, the mitochondrial mass and cell volume estimated from flow cytometry indicated that faster growing CL4L exhibited less mitochondria mass and smaller cell volume as well as lower volume fraction occupied by mitochondria (Fig. 3.9), and that the assumed constant mitochondria volume fraction is one of the model's limitations. Thus, cells may possess regulatory mechanisms to tune the volume fraction of mitochondria at different growth rates and whether such mechanisms are regulated by the MC remains to be investigated.

This study, together with our previous efforts (Beg, Vazquez et al. 2007; Vazquez, Beg et al. 2008; Vazquez, de Menezes et al. 2008; Vazquez, Liu et al. 2010), indicates that MC is not only important for constraining the *E. coli* and *S. cerevisiae* metabolism, but also a potential contributor to the metabolism of rapidly proliferating mammalian cells, through controlling the optimal distribution of enzymes responsible for glycolysis or mitochondrial respiration. With the help of systems-oriented simulation of a reduced human cell metabolism model, here, we find that MC, as a biophysical constraint, could potentially contribute to aerobic glycolysis (the Warburg effect), for all rapidly proliferating cells. The identified biophysical constraint MC together with this systems-oriented approach may prove useful for future target identification for cancer therapy, once the tissue and cell type specific human metabolic network is fully reconstructed.

These results may also have therapeutic indications on cancer therapy. It has been previously proposed that the Warburg effect may represent a unique property of cancer cell metabolism as a metabolic target (Chen, Lu et al. 2007; Gatenby and Gillies 2007; Vander

Heiden 2010). Pharmacological reduction of lactate production and excretion may indeed reduce the invasive potential of tumor cells or makes them more susceptible to apoptosis (Cao, Yacoub et al. 2008; Madhok, Yeluri et al. 2010). However, the metabolic state of tumor cells alone is not a unique target for cancer chemotherapy. The combination of the metabolic targeting with individual, tumor specific isozymes (Christofk, Vander Heiden et al. 2008; Christofk, Vander Heiden et al. 2008) or tumor-penetrating peptides (Feron 2010; Sugahara, Teesalu et al. 2010) may improve its specificity and efficacy.

3.4 MATERIALS AND METHODS

Cell culture and growth rate determination

Genetically-manipulated derivatives of human BJ fibroblasts, transfected with the catalytic subunit of human telomerase (hTERT) (CL1), hTERT+ SV40 large T (LT) (CL2), hTERT+LT+ SV40 small T (ST)(CL3), and hTERT+LT+ST+ V^{12} mutant human Ras (Ras) (CL4) (Hahn, Counter et al. 1999) were used. The cells were maintained in DMEM supplemented with 25mM glucose, 4mM glutamine, 1mM pyruvate and 10% FBS at 37° in a CO_2 incubator with 5% CO_2 and 95% air. For assessing their growth rates cells were seeded at a density of 2×10^4 cells in 12-well plates and at every 24 h a subset were collected following trypsinization (in 0.25% trypsin, 0.53 mM EDTA). Cell numbers were determined by blinded counting of cells on a standard grid-patterned hemocytometer.

Western blot

CL1-4 cells were harvested by trypsinization at 48h, 72h, 96h and 120h after seeding, and washed with PBS. The cell pellets were dissolved in 100µl RIPA lysis buffer supplemented with Halt complete protease inhibitor cocktail and phosphatase inhibitor cocktail (ThermoFisher) on ice for 1h. After centrifugation for 15min at 4°C, and the protein concentration of the supernatants were determined by Bradford assay (Bio-Rad). 50 µg of each sample was then applied to 15% Tris SDS-PAGE gel and, following electrophoresis, transferred to PVDF membranes at 6v for 16 hr at 4°C. The membranes were then immunostained using actin-specific monoclonal antibody (Sigma, 1:8000), LDH monoclonal antibody (Santa Cruz Biotech, 1:4000) and VDAC/porin monoclonal antibody (Abcam, 1:5000), respectively; All experiments were performed in triplicates.

Ras activity measurements

Ras activity was performed using active Ras pull-down and detection kit (Thermo Scientific), as instructed. Briefly, 1mg fresh lysates from CL1-4 were subjected to affinity precipitation of activated Ras with the GST-fused Ras-binding domain (RBD) of Raf1 through the glutathione agarose resin. Half of the eluted samples (25µl, active Ras) from CL1-4 lysates were separated with 15% Tris SDS-PAGE gel and immunoblotted with anti-Ras Ab.

Soft agar assay

Soft agar assays were performed as described (Stausbol-Gron, Havsteen et al. 1998). Briefly, 10,000 cells from each cell lines were resuspended in 2 ml of 0.35% top agar in complete DMEM growth medium and seeded in six-well plates previously filled with 2 ml of 0.7% basal agar in complete growth medium. The assay was performed in duplicate. After 2

weeks, 1ml 0.005% crystal violet was added to each well to stain the colonies for 2h and colonies in each well were counted.

LDH activity measurements

LDH activity was determined by the decreasing absorbance of β-NADH at 340nm as described (Wu, Daugulis et al. 1992) with slight modifications. Sodium pyruvate and β-NADH were dissolved separately in 0.2M Tris-HCl, pH 7.4. The mixture solution of 2mM sodium pyruvate and 0.2mM β-NADH was prepared fresh before being used in the LDH activity assay. A total cell lysate sample (100μg) was pipetted into a 1.5ml disposable semi-micro cuvette (1cm light path), followed by the addition of 1ml pyruvate-NADH mixture solution. After mixing the cuvette was immediately placed in the spectrophotometer and the absorbance was recorded for 30s at 340nm. The slope of the absorbance over this time period (-ΔA/min) was proportional to the rate of NADH consumption, which represents the LDH activity.

Measuring mitochondrial content and membrane potential

Cells were trypsinized and washed with PBS in suspensions of 1 x 10^6 cells/ml at the indicated time points, followed by incubating them with 2μg/ml JC-1 or 100nM MitoTracker Green (Invitrogen) at 37 °C for 15min, respectively. After rinsing with PBS, cells were subjected to flow cytometry to read the fluorescence intensity at 590nm (red channel) and 520nm (green channel). All fluorescence intensities were normalized to that of CL1 fibroblasts at 48h.

Measurement of cellular oxygen consumption and acidification rate:

All cell lines were cultured in one T175 and T75 flasks prior to time-course experiments when monolayer cells were trypsinized and suspended in the complete growth medium DMEM. After cell counting, $7.4X10^5$ cells were seeded in each $145cm^2$ petri-dish (6 dishes, one for each day) for each cell line. The final volume was brought to 30ml in each dish. 24h after seeding, all petri-dish cultures were replenished with fresh complete DMEM. At each time point, each dish was trypsinized and counted for growth curve. Also, at 48h and 120h after seeding, cells were detached and counted for experiments using the Seahorse instrument platform, which allows the simultaneous measurement of oxygen consumption rate and extracellular acidification rate by the use of pO_2 and pH probe in the formed microchambers (Qian and Van Houten 2010). The whole experiment was repeated once with the same cell lines (one passage older, only 2 dishes per cell line, 48h and 120h only). 48h and 120h after seeding, all 6 cell lines were seeded at 30,000 viable cells per well in triplicates or quadruplicates into XF24 tissue culture plates in complete DMEM medium. Cell medium was changed after 20h to unbuffered DMEM (8.3g/L DMEM, 25mM glucose, 4mM glutamine, 1mM pyruvate, 31.6mM NaCl, and 15mg Phenol Red, adjusted pH to 7.4 with NaOH) and kept in CO_2-free incubator for 1h. Measurement of cellular oxygen consumption and extracellular acidification rate was performed during 4-min intervals. After the experiment, all the measurements were normalized to the total cell number.

Statistical analyses

Paired two-tail student's *t*-tests were used to assess the significance between two groups. For more than two groups, one-way ANOVA was performed to determine the differences between groups. *: $p<0.05$; **: $p<0.01$; ***: $p<0.001$.

3.5 ACKNOWLEGMENT

AV was sponsored by the RWJ Foundation, Project title: “Strengthening the Cancer Institute of New Jersey in Cancer Prevention, Control and Population Science to Improve Cancer Care”, Cancer Informatics Category. This work has been supported in part by NIH grants NIAID U01 AI070499 and NIGMS R01 GM085022 to ZNO.

4.0 COMPARATIVE GENOME-SCALE METABOLIC NETWORK ANALYSIS OF MULTIPLE *STAPHYLOCOCCUS AUREUS* GENOMES IDENTIFY NOVEL ANTI-MICROBIAL DRUG TARGETS

4.1 INTRODUCTION

As we described in **Chapter 3**, the full metabolic network reconstruction of mammalian cells in a tissue and cell type specific fashion is not yet possible. Therefore, we turn to prokaryotic cells in which full metabolic network reconstruction is now routine. Specifically, to provide a proof-of-principle of the potential of a systems-oriented metabolic network analysis on drug discovery targeting cell metabolism, we focus on *Staphylococcus aureus* because of its virulence, and its adaptation ability to different environmental conditions (Lowy 1998), and because its mortality now surpasses that of human immunodeficiency virus/AIDS in the United States (Klevens, Morrison et al. 2007). In addition, the emergence of methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant *S. aureus* (VRSA) isolates, underscores the need to identify new targets and molecules effective against multidrug-resistant strains of *S. aureus.*

In order to address the differences in the metabolic capabilities of various *S. aureus* strains and further refine a generic antimicrobial drug target identification scheme (Almaas, Oltvai et al. 2005) from a systems-biology perspective, we performed metabolic reconstructions on multidrug-resistant and sensitive strains of all 13 *S. aureus,* with available whole genome sequences. They include strain N315 (a MRSA strain), Mu50 (a VRSA strain), JH9 (a vancomycin-nonsusceptible MRSA strain), JH1 (a vancomycin-susceptible, hospital-acquired

MRSA strain), COL (a hospital-acquired MRSA strain), 252 (a hospital-acquired MRSA strain), USA 300 (a community-acquired MRSA strain), MW2 (a community-acquired MRSA strain), and RF122 (a bovine mastitis strain). This approach enabled us to identify the functional pathways, metabolic reactions, and transport reactions of several sequenced strains of *S. aureus*. The identified metabolic pathways and their individual reactions were systematically compared with those archived in the KEGG ligand database (Kanehisa and Goto 2000). After metabolic reconstruction, we also performed flux balance analysis (FBA) on such reconstructed metabolic networks to identify single enzymes and synthetic enzyme pairs that are unconditionally required for the growth (biomass production) of all *S. aureus* strains as potential anti-Staphylococcus targets.

4.2 RESULTS

4.2.1 Analysis of metabolic reconstruction of *S. aureus*.

We initially studied the MRSA N315 strain as a model for reconstruction and FBA studies, and then studied the other 12 strains. From the ERGO genomic database, we identified 2,593 ORFs in the *S. aureus* N135 chromosome and 31 ORFs on its plasmid, of which 2,505 ORFs were assigned a functional role. A total of 906 ORFs were identified to have an enzymatic activity (i.e., they had EC numbers), and 668 ORFs had unique complete EC numbers (including two full EC numbers on the plasmid). We identified a list of 61 functions with incomplete or partial EC numbers. For all the complete and incomplete EC number annotations, the associated reactions were identified from the ERGO pathway collections and the KEGG pathway database. A total of

1,493 reactions were generated from the two databases, after reactions with substrates that were not likely to be metabolized by *S. aureus* and those that did not have a corresponding transport system were removed.

We compared this genome-scale reconstruction with two previous reconstructions of *S. aureus* N315 (Becker and Palsson 2005; Heinemann, Kummel et al. 2005). We found that 75 reactions in Becker and Palsson's reconstruction did not have a corresponding ORF in the *S. aureus* N315 genome, and several reactions were catalyzed strictly by eukaryotic enzymes. In a similar study by Heinemann et al. (Heinemann, Kummel et al. 2005), only 774 reactions representing only 23% of the ORFs with 394 unique enzymes were considered. In contrast, we identified a total of 546 enzymes representing 1,497 reactions. The number of identified transport reactions in our study was 164 compared to 84 in the Becker and Palsson (Becker and Palsson 2005) and 59 in the Heinemann et al. (Heinemann, Kummel et al. 2005) studies. We also found that 51 of the 59 transport reactions had an associated ORF, and 47 of them (92%) were also found in our reconstruction. Using annotations based on protein similarities, comparative genomics, and pathways for all the strains, we obtained a comprehensive list of enzymes, reactions, and metabolites for all the *S. aureus* strains (Table 4.1). We found that over 90% of the enzymes, metabolic reactions, and metabolites were common to all strains, suggesting the presence of a common "core of metabolic reactions" among all the strains studied (Almaas, Oltvai et al. 2005).

Table 4.1. Comparative statistics for enzymes and metabolites that were used for the metabolic reconstructions of 13 S. aureus strains

Parameter	No. for strain:												
	Newman	TCH1516	USA300	Mu3	JH9	JH1	COL	RF122	Mu50	MW2	MSSA476	MRSA252	N315
ORFs	2,641	2,683	2,604	2,698	2,726	2,681	2,618	2,515	2,714	2,632	2,597	2,655	2,593
ORFs with functions	1,915	1,896	2,002	2,031	2,087	2,096	1,953	1,915	2,091	2,021	1,975	2,024	2,031
unique full EC no.	650	646	671	663	670	663	666	666	667	665	671	675	668
Dual-fusion enzymes	51	49	50	51	51	51	50	50	50	50	49	49	51
Triple-fusion enzymes	6	6	6	6	6	6	6	6	6	6	6	6	6
Quadruple-fusion enzymes	4	4	4	4	4	4	4	4	4	4	4	4	4
Very likely reactions	1,231	1,213	1,255	1,247	1,254	1,247	1,254	1,248	1,251	1,251	1,258	1,258	1,253
Less likely reactions	215	220	218	218	220	219	214	225	219	215	216	221	220
Likely reactions	20	19	20	20	20	19	20	16	20	20	20	20	20
EC no. with reactions	534	522	544	543	547	541	538	537	545	540	540	544	546
Cellular reactions	1,208	1,186	1,224	1,222	1,227	1,220	1,214	1,213	1,225	1,219	1,217	1,223	1,227
Transport reactions	141	146	148	144	146	145	150	151	145	147	152	152	146
Exchange reactions	117	120	121	119	121	120	124	125	120	120	125	125	120
Metabolites	1,412	1,399	1,428	1,422	1,431	1,422	1,425	1,414	1,427	1,421	1,432	1,437	1,431

4.2.2 Identification of single unconditionally essential enzymes and synthetic-lethal pairs.

In an effort to identify antibiotic drug targets, we performed *in silico* enzyme deletion studies for each *S. aureus* strain (Almaas, Oltvai et al. 2005; Becker and Palsson 2005; Thiele, Vo et al. 2005; Deutscher, Meilijson et al. 2006). With a metabolic reconstruction in hand, we built a flux balance-based model of *S. aureus*. The vulnerability of the reconstructed metabolic network models can be utilized to derive the candidates for antibiotic targets. In the case of *S. aureus* N315, of the 1,497 identified metabolic reactions, only 23% were active, i.e., carrying nonzero fluxes (Almaas, Oltvai et al. 2005). The FBA, as described in Materials and Methods, identified only 337 reactions that were active in the rich growth medium, in which all the included transport reactions can occur without limitation. Furthermore, to validate the identified potential antibiotic drug targets, we initially performed an *in silico* single enzyme deletion study for growth in (simulated) rich medium for each *S. aureus* strain. In various natural environments, in which uptake reactions are limited, more enzymes are essential than in the rich medium. In contrast, the unconditionally essential enzymes are at the intersection of the sets of such essential enzymes in all possible environments. We found that 70 enzymes proved essential for growth in one or more of the 13 *S. aureus* strains, while 44 of these were essential in all the strains (Table

4.2). Figure 4.1a shows the network structure of the common unconditionally essential enzymes. For example, two enzymes, EC 2.5.1.7 and EC 1.1.1.158, that are involved in the conversion of UDP-*N*-acetyl-D-glucosamine to UDP-*N*-acetylmuramate through UDP-*N*-acetyl-3-*O*-(1-carboxyvinyl)-D-glucosamine are predicted to be unconditionally essential. Deletion (inactivation) of either of these enzymes leads to failure in generating the biomass component UDP-*N*-acetylmuramate, since these two consecutive reactions constitute the only pathway generating UDP-*N*-acetylmuramate. While there is no comprehensive list of experimentally confirmed essential genes for any strain of *S. aureus*, we used the list of available ORFs that were identified to be essential for *S. aureus* strains RN450 and RN4220 (Ji, Zhang et al. 2001; Forsyth, Haselbeck et al. 2002) to determine if our prediction matched any experimentally verified genes.

Of the 44 common single essential enzymes that were identified from FBA, we found at least six enzymes, including transketolase (EC 2.2.1.1), hydroxyl-methylbilane synthase (EC 2.5.1.61), methionine adenosyltransferase (EC 2.5.1.6), UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7), protein N (pi)-phospho-histidine-sugar phosphortransferase (EC 2.7.1.69), and acetyl-CoA carboxylase (EC 6.4.1.2), to be experimentally verified in *S. aureus* (Forsyth, Haselbeck et al. 2002). In addition, orthologs for 17 enzymes were found experimentally to be essential in other bacteria. Many enzymes were predicted to be dispensable for growth, i.e., biomass could be generated in the presence of alternate biosynthetic routes, along with their corresponding ORFs. We hypothesized that if pairs of enzymes independently play an imporant role in metabolism, then deleting both enzymes could lead to lethality (Lehar, Stockwell et al. 2008; Torella, Chait et al. 2010). Such pairs were identifiable by monitoring the biomass production reactions on inactivation of two enzymes

simultaneously in the FBA computations. We identified 54 synthetic lethal pairs of enzymes in all 13 strains whose deletions could completely abolish biomass generation. Interestingly, only 10 of these 54 synthetic-lethal pairs were common to all 13 strains (Table 4.3).

Table 4.2. Single essential enzymes in all strains of *S. aureus*

Pathway	EC no.	Gene	Enzyme name	Published experimental evidence
Amino sugar metabolism	1.1.1.158	*murB*	UDP-*N*-acetylmuramate dehydrogenase	*Salmonella enterica* serovar Typhimurium
	2.5.1.7	*murA*	UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase	*Streptococcus pneumoniae*, *Mycobacterium tuberculosis* H37Rv, *S. aureus*
Fatty acid biosynthesis	1.1.1.100	*fabG*	3-Oxoacyl-[acyl carrier protein] reductase	
	2.3.1.39	*fabD*	[Acyl carrier protein] *S*-malonyltransferase	
	2.3.1.41	*fabB/fabH*	Beta-ketoacyl-acyl-carrier-protein synthase I	
	3.1.2.14		Oleoyl-[acyl carrier protein] hydrolase	
	4.2.1.58	*fabZ*	Crotonoyl-[acyl carrier protein] hydratase	
	4.2.1.60	*fabA*	3-Hydroxydecanoyl-[acyl carrier protein] dehydratase	*E. coli*
	4.2.1.61		3-Hydroxypalmitoyl-[acyl carrier protein] dehydratase	
	6.4.1.2	*accA/accD*	Acetyl-CoA carboxylase	*B. subtilis*, *Haemophilus influenzae*, *S. pneumoniae*, *S. aureus*
Histidine metabolism	1.1.1.23	*hisD*	Histidinol dehydrogenase	*M. tuberculosis* H37Rv
	2.4.2.17	*hisG*	ATP phosphoribosyltransferase	*M. tuberculosis* H37Rv, *H. influenzae*
	3.1.3.15	*hisB*	Histidinol phosphatase	
	3.5.4.19	*hisI*	Phosphoribosyl-AMP cyclohydrolase	*M. tuberculosis* H37Rv
	3.6.1.31	*hisI*	Phosphoribosyl-ATP diphosphatase	
	4.2.1.19	*hisB*	Imidazoleglycerol-phosphate dehydratase	*M. tuberculosis* H37Rv, *H. influenzae*
	5.3.1.16	*hisA*	1-(5-Phosphoribosyl)-5-[(5-phosphoribosylamino) methylideneamino] imidazole-4-carboxamide isomerase	
	2.6.1.9	*hisC*	Histidinol-phosphate transaminase	
Amino acid (phenylalanine, tyrosine and tryptophan) biosynthesis	1.1.1.25	*aroE*	Shikimate dehydrogenase	
	2.4.2.18	*trpD*	Anthranilate phosphoribosyltransferase	*M. tuberculosis* H37Rv
	2.5.1.19	*aroA*	3-Phosphoshikimate 1-carboxyvinyltransferase	
	2.5.1.54	*aroP/aroH*	3-Deoxy-7-phosphoheptulonate synthase	
	2.7.1.71	*aroL/aroK*	Shikimate kinase	*M. tuberculosis* H37Rv
	4.1.1.48	*trpC*	Indole-3-glycerol-phosphate synthase	*M. tuberculosis* H37Rv
	4.1.3.27	*trpE*	Anthranilate synthase	*M. tuberculosis* H37Rv, *H. influenzae*
	4.2.1.10	*aroQ*	3-Dehydroquinate dehydratase	*M. tuberculosis* H37Rv
	4.2.1.20	*trpA/trpB*	Tryptophan synthase	*M. tuberculosis* H37Rv, *H. influenzae*
	4.2.1.51	*pheC*	Prephenate dehydratase	*H. influenzae*
	4.2.3.4	*aroB*	3-Dehydroquinate synthase	*M. tuberculosis* H37Rv, *H. influenzae*
	4.2.3.5	*aroC*	Chorismate synthase	*M. tuberculosis* H37Rv, *H. influenzae*, *Helicobacter pylori*
	5.3.1.24	*trpF*	Phosphoribosylanthranilate isomerase	
	5.4.99.5	*aroQ*	Chorismate mutase	*H. influenzae*
Porphyrin metabolism	2.1.1.107	*cysG*	*C*-Methyltransferase, uroporphyrinogen-III	
	2.5.1.61	*hemC*	Hydroxymethylbilane synthase	*S. aureus*
	4.2.1.24	*hemB*	Porphobilinogen synthase	
	4.2.1.75	*hemD*	Uroporphyrinogen-III synthase	
	4.99.1.4	*cysG*	Sirohydrochlorin ferrochelatase	
Methionine metabolism	2.5.1.6	*metK*	Methionine adenosyltransferase	*S. aureus*
	3.2.2.9		Adenosylhomocysteine nucleosidase	
	4.4.1.21	*luxS*	*S*-Ribosylhomocysteine lyase	
Pentose phoshpate pathway	2.2.1.1	*tktA*	Transketolase	*B. subtilis*, *S. aureus*, *H. influenzae*, *M. tuberculosis* H37Rv
Multiple	2.7.1.69	*yadI*	Protein-N(pi)-phosphohistidine-sugar phosphotransferase	*S. aureus*
Unknown	3.5.1.-	*pcnA*	Pyrazinamidase	*M. tuberculosis* H37Rv
	4.1.3.-	*hisF*	Imidazole glycerol phosphate synthase, cyclase subunit	

Reproduced with permission from (Lee et al., 2009).

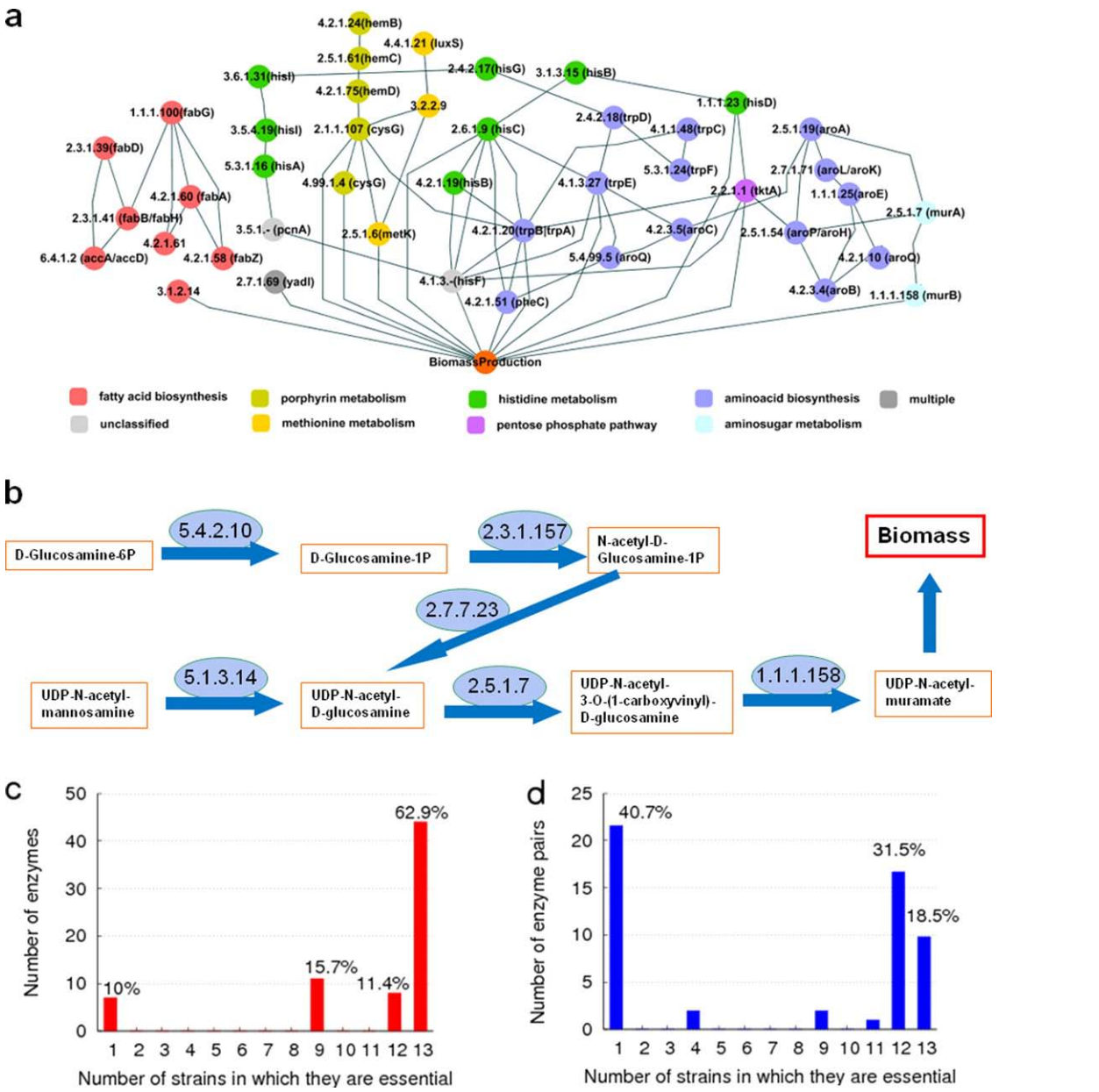


Figure 4.1. Single essential and synthetic-lethal enzymes of the *S. aureus* metabolic network and their strain dependence.

(**a**) Sub-networks of 44 strain-independent single essential enzymes of multiple *S. aureus* strains. Two nodes (enzymes) are connected if their reactions involve common metabolites. The EC number and the associated genes are presented for each node. The node colors correspond to the metabolic pathways the enzymes belong to.

(**b**) Subset of the amino sugar pathway involving two single essential enzymes, EC 2.5.1.7 and EC1.1.1.158. Three pairs of enzymes, EC 2.3.1.157-EC 5.1.3.14, EC 2.7.7.23-EC 5.1.3.14, and EC 5.4.2.10-EC 5.1.3.14, are synthetic-lethal pairs in all strains. The enzymes in each pair individually do not affect growth (biomass production), but their simultaneous inactivation/deletion causes lethality.

(**c**) Distribution of the numbers of strains in which an enzyme is singly essential. Altogether, 70 enzymes were identified as singly essential in at least one strain. Of these, more than 60% are found in all the S. aureus strains.
(**d**) Distribution of the numbers of strains in which the deletion of a pair of enzymes is synthetic lethal. A total of 54 synthetic-lethal pairs were identified in at least one strain. About 40% of them are found only in one strain, reflecting the strain specificity of synthetic-lethal pairs, with less than 20% being present in all the strains. Reproduced with permission from (Lee et al., 2009).

Six lethal pairs belong to amino sugars/peptidoglycan pathways that can be grouped under the cell wall metabolic subsystem. Two lethal pairs belong to the amino acid subsystem, and a single pair of enzymes belongs to amino acid and 1-carbon metabolic pathways. Only one pair of enzymes is grouped in the nucleotide and carbohydrate subsystems. An example of a common lethal pair is shown in amino sugar metabolism (Fig. 4.1b). In *S. aureus*, there are two pathways that can generate UDP-*N*-acetylglucosamine, the precursor of UDP-*N*-acetylmuramate. One is from D-glucosamine-1P through *N*-acetyl-D-glucosamine-1P, and the other is from *N*acetyl- D- mannosamine. The first pathway consists of three enzymatic reactions catalyzed by EC 5.4.2.10, EC 2.3.1.157, and EC 2.7.7.23, and the second has EC 5.1.3.14. Therefore, we found three synthetic-lethal pairs (EC 5.3.1.14-EC 5.4.2.10, EC 5.1.3.14-EC 2.3.1.157, and EC 5.1.3.14-EC 2.7.7.23). The subset of the amino sugar biosynthetic pathway involving these enzymes is shown in Fig. 4.1b. The number of single essential enzymes varies by individual strain. As seen in Fig. 4.1c, more than 90% of the identified single essential enzymes are essential in nine strains. The strain dependence of the single essential enzymes is, however, stronger than that of the metabolic reconstructions, as 63% of the single essential enzymes are found in all the strains, while more than 90% of the metabolic reconstructions overlap. Synthetic-lethal pairs exhibit even stronger strain dependence (Fig. 4.1d); less than 20% are found in all the strains, and 40% are found in only one specific strain.

Table 4.3. Common synthetic-lethal pairs of enzymes required for growth identified in all 13 *S. aureus* strains

Sl	Enzyme pair	Gene	Reaction	Pathway	Subsystem
1	Prephenate dehydrogenase EC 1.3.1.12	*tyrA*	Prephenate + NAD^+ ↔ 3-(4 hydroxyphenyl) pyruvate + CO_2 + NADH + H^+	Tyrosine biosynthesis	Amino acid biosynthesis
	Arogenate dehydrogenase EC 1.3.1.43	*adH*	L-Arogenate + NAD^+ ↔ L-tyrosine + CO_2 + NADH	Tyrosine biosynthesis via arogenate	Amino acid biosynthesis
2	Methylenetetrahydrofolate reductase EC1.5.1.20	*mthfR*	5-Methyltetrahydrofolate + $NADP^+$ ↔ 5,10-methylenetetrahydrofolate + NADPH + H^+	Tetrahydrofolate synthesis	One-carbon metabolism
	Cystathionine gamma-synthase EC 4.4.1.8	*metF*	L-Cysteine + H_2O ↔ hydrogen sulfide + pyruvate + NH_3	Cysteine metabolism	Amino acid biosynthesis
3	5-Methyltetrahydrofolate-homocysteine methyl-transferase EC 2.1.1.13	*metH*	5-Methyltetrahydrofolate + L-homocysteine ↔ tetrahydrofolate + L-methionine	L-Methionine biosynthesis	Amino acid biosynthesis
	Cystathionine gamma-synthase EC 4.4.1.8	*metF*	L-Cysteine + H2O ↔ hydrogen sulfide + pyruvate + NH3	Cysteine metabolism	Amino acid biosynthesis
4	UDP-*N*-acetylglucosamine pyrophosphorylase EC 2.3.1.157	*gcaD/*	Acetyl-CoA + D-glucosamine 1-phosphate ↔ CoA + *N*-acetyl-D-glucosamine 1-phosphate	Peptidoglycan and lipopolysaccharide biosynthesis	Cell wall metabolism
	UDP-*N*-acetylglucosamine pyrophosphorylase EC 2.7.7.23	*glmU*	UTP + *N*-acetyl-D-glucosamine 1-phosphate ↔ pyrophosphate + UDP-*N*-acetyl-D-glucosamine UTP + *N*-acetyl-alpha-D-glucosamine 1-phosphate ↔ pyrophosphate + UDP-*N*-acetyl-D-glucosamine	Amino sugar biosynthesis	Cell wall metabolism
5	UDP-*N*-acetylglucosamine pyrophosphorylase EC 2.3.1.157	*gcaD/*	Acetyl-CoA + D-glucosamine 1-phosphate ↔ CoA + *N*-acetyl-D-glucosamine 1-phosphate	Peptidoglycan and lipopolysaccharide biosynthesis	Cell wall metabolism
	UDP-*N*-acetylglucosemine 2-epimerase EC 5.1.3.14	*rffE*	UDP-*N*-acetyl-D-glucosamine ↔ UDP-*N*-acetyl-D-mannosamine UDP-*N*-acetyl-D-glucosamine + H_2O ↔ *N*-acetyl-D-mannosamine + UDP	Amino sugar biosynthesis	Cell wall metabolism
6	UDP-*N*-acetylglucosamine pyrophosphorylase EC 2.3.1.157	*gcaD*	Acetyl-CoA + D-glucosamine 1-phosphate ↔ CoA + *N*-acetyl-D-glucosamine 1-phosphate	Peptidoglycan and lipopolysaccharide biosynthesis	Cell wall metabolism
	Phosphoglucosamine mutase EC 5.4.2.10	*glmM*	D-Glucosamine 1-phosphate ↔ D-glucosamine 6-phosphate	Amino sugar biosynthesis	Cell wall metabolism
7	UDP-*N*-acetylglucosamine pyrophosphorylase EC 2.7.7.23	*glmU*	UTP + *N*-acetyl-D-glucosamine 1-phosphate ↔ pyrophosphate + UDP-*N*-acetyl-D-glucosamine UTP + *N*-acetyl-alpha-D-glucosamine 1-phosphate ↔ pyrophosphate + UDP-*N*-acetyl-D-glucosamine	Amino sugar biosynthesis	Cell wall metabolism
	UDP-*N*-acetylglucosemine 2-epimerase EC 5.1.3.14	*rffE*	UDP-*N*-acetyl-D-glucosamine ↔ UDP-*N*-acetyl-D-mannosamine UDP-*N*-acetyl-D-glucosamine + H_2O ↔ *N*-acetyl-D-mannosamine + UDP	Amino sugar biosynthesis	Cell wall metabolism
8	UDP-*N*-acetylglucosamine pyrophosphorylase EC 2.7.7.23	*glmU*	UTP + *N*-acetyl-D-glucosamine 1-phosphate ↔ pyrophosphate + UDP-*N*-acetyl-D-glucosamine UTP + *N*-acetyl-alpha-D-glucosamine 1-phosphate ↔ pyrophosphate + UDP-*N*-acetyl-D-glucosamine	Amino sugar biosynthesis	Cell wall metabolism
	Phosphoglucosamine mutase EC 5.4.2.10	*glmM*	D-Glucosamine 1-phosphate ↔ D-glucosamine 6-phosphate	Amino sugar biosynthesis	Cell wall metabolism
9	UDP-*N*-acetylglucos-amine 2-epimerase EC 5.1.3.14	*rffE*	UDP-*N*-acetyl-D-glucosamine ↔ UDP-*N*-acetyl-D-mannosamine UDP-*N*-acetyl-D-glucosamine + H_2O ↔ *N*-acetyl-D-mannosamine + UDP	Amino sugar biosynthesis	Cell wall metabolism
	Phosphoglucosamine mutase EC 5.4.2.10	*glmM*	D-Glucosamine 1-phosphate ↔ D-glucosamine 6-phosphate	Amino sugar biosynthesis	Cell wall metabolism
10	Carbamoyl-phosphate synthase EC 6.3.5.5	*carA*	2 ATP + L-glutamine + HCO_3- + H_2O ↔ 2 ADP + orthophosphate + L-glutamate + carbamoyl phosphate	Pyrimidine and glutamate metabolism	Nucleotide amino acid metabolism
	Pyruvate carboxylase EC 6.4.1.1	*pycA/pycB*	ATP + pyruvate + HCO_3^- ↔ ADP + orthophosphate + oxaloacetate	Gluconeogenesis	Carbohydrate metabolism

4.2.3 Analysis of the components of minimal medium for *S. aureus*.

The minimal medium for the growth of the laboratory strain *S. aureus* NCTC 8385, requiring a combination of salts, carbon source, amino acids, and vitamins, was deduced by trial

and error in the early 1970s (Rudin, Sjostrom et al. 1974). However, with the availability of complete genome sequences, annotations, pathway analyses, and metabolic reconstructions, one can rationally design an optimum growth medium (Bhattacharyya, Stilwagen et al. 2002; Becker and Palsson 2005). We compared the published minimal-medium components (Rudin, Sjostrom et al. 1974) with the reconstructions and calculated that six components could be synthesized by *S. aureus,* and therefore, these components are not required for growth of *S. aureus* (Table 4.4). Specifically, for several salt components of the medium, a dedicated transport system(s) has been identified. For amino acids, such as L-arginine, L-glutamic acid, L-valine, L-threonine, and L-phenylalanine, both biosynthetic pathways and transport systems are present in all genomes. We also identified the ORFs coding for synthesizing and transporting enzymes for cofactors, such as biotin and pantothenate. Interestingly, the ORF for pantothenate uptake (Na/pantothenate symporter) is not found in any of the *S. aureus* genomes (Table 4.4). We conclude that *S. aureus* can synthesize biotin and does not require its exogenous addition to the medium.

To experimentally confirm the minimal-medium prediction, we compared the growth of three different *S. aureus* strains. These included 2 of the 13 reconstructed *S. aureus* strains (Mu50 and USA300), as well as a patient *S. aureus* isolate. We tested the growth abilities of these strains in rich LB medium (control), in the standard minimal medium (AAM+) (Rudin, Sjostrom et al. 1974), and in minimal medium (AAM-) predicted by our study. The growth patterns of all three strains were the highest in LB medium compared to growth in either minimal medium and followed a typical sigmoid growth curve (Fig. 4.2). However, the growth rates in the predicted minimal growth medium (AAM-) without the addition of seven compounds were similar to those in the standard minimal medium (AAM+). These results thus support the minimal-medium predictions, and hence, the validity of our metabolic-pathway reconstructions.

Table 4.4. Analysis of minimal growth medium for *S. aureus*

Sl	Ingredient	Amt/liter	Reconstruction-based analysis	Comments
1	KCl	3.0 g	Potassium uptake sodium symport *ktrAB*	K is transported by a dedicated transport system
2	NaCl	9.5 g	Potassium uptake sodium symport *ktrAB* *N*-Acetylneuraminate uptake sodium symport L-Proline uptake sodium symport *putP*	At least three symport systems are identified
3	$MgSO_4 \cdot 7H_2O$	1.3 g	Mg phosphate uptake proton symport pit Magnesium uptake protein *mgtE* Magnesium uptake via *corA*	At least three symport systems are identified
4	$(NH4)_2 SO_4$	4.0 g	Ammonium uptake uniport via *amtB*	One dedicated uniport system has been identified
5	$CaCl_2 \cdot 2H_2O$	22 mg	Ca phosphate uptake proton symport pit	One dedicated uptake system has been identified
6	KH_2PO_4	140 mg	Potassium uptake sodium symport *ktrAB*	K is transported by a dedicated transport system
	$FeSO_4 \cdot 7H_2O$	6 mg	Ferrous iron uptake *feoB* Ferrous iron uptake FTR1 family	One dedicated and a generic "family" transport system have been identified
7	$MnSO_4 \cdot H_2O$	10 mg	Mn phosphate uptake proton symport pit	One dedicated uptake system has been identified
8	Citiric acid	6 mg	Taken in by citrate-Fe_3 dicitrate ABC transport system; transports iron, as well	One dedicated uptake system has been identified
9	Tris	12.1 g	Buffer	
10	Glucose	5 g	Glucose uptake *glcBC* PTS forming glucose-6-phosphate D-Glucose uptake glucose transporter *glcU* Glucose uptake *glcABC* PTS forming glucose-6-phosphate	At least three uptake systems are identified
11	L-Arginine	125 mg	3 enzymes in the synthesis of L-arginine from carobmyl phosphate and ornithine are present (EC 2.1.3.3, EC 6.3.4.5, EC 4.3.2.1). Ornithine is made from 5-glutmate semialdehyde and 1-pyrroline 5-carboxylate from proline	Arginine need not be added in the medium, as the biosynthetic pathway is present
			Arginine uptake proton symport Arg permease	Two uptake systems are identified; one is a permease, and the second is an ABC transporter
			L-Arginine uptake arginine ABC transporter	
12	L-Proline	200 mg	The organism is missing two enzymes (EC 2.7.2.11 and EC 1.2.1.41) involved in L-proline synthesis from L-glutamate. However, it possesses a bypass producing L-proline from L-ornithine using two other enzymes (EC 2.6.1.13 and EC 1.5.1.2) L-proline uptake proton symport *proP* L-proline uptake sodium symport *putP*	Proline is to be added in the medium. In addition, two protein symport systems, *proP* and *putP*, have been identified
13	L-Glutamic acid	250 mg	3 enzymes producing L-glutamate from 2-oxoglutarate are present	Glutamic acid need not be added in the medium, as the biosynthetic pathway is present
			L-glutamate uptake sodium symport gltS	Two uptake systems using Na^+ and proton symport have been identified
			L-glutamate uptake proton sodium symport *gltT*	
14	L-Valine	150 mg	Synthetic pathway present	Valine need not be added in the medium, as the biosynthetic pathway is present
			L-Valine uptake proton symport (Leu-Ile-Val) transporter	A dedicated transporter has been identified
15	L-Threonine	150 mg	Synthetic pathway present	Threonine need not be added in the medium, as the biosynthetic pathway is present
			L-Threonine export proton antiport Thr-Ser exporter	A dedicated transporter has been identified
16	L-Phenylalanine	150 mg	Synthetic pathway present	Phenlyalanine need not be added in the medium, as the biosynthetic pathway is present
			No uptake transporter present	A dedicated phenylalanine transporter has been not identified
17	L-Leucine	150 mg	The organism is missing an ORF for leucine transaminase (EC 2.6.1.6) and leucine dehydrogenase (EC 1.4.1.9). However, it is able to synthesize L-leucine due to the presence of an ORF for branched-chain amino acid transaminase (EC 2.6.1.42) L-Isoleucine export proton antiport AzlCD L-Leucine uptake proton symport (Leu-Ile-Val) transporter	One specific and one general leucine transporter have been identified
18	L-Cysteine	80 mg	Two enzymes (EC 2.3.1.30 and EC 2.5.1.47) involved in the L-cysteine biosynthesis pathway from L-serine are present. However, the L-serine pathway is broken, so L-cysteine will be synthesized in the presence of exogenous L-serine. In addition, the L-cysteine biosynthesis pathway from L-homocysteine is also broken due to the absence of cystatione beta-synthase (EC 4.2.1.22). Cysteine uptake by ABC transporter present	One ABC transporter system has been identified
19	Biotin	0.1 mg	Biotin synthetic pathway present, primarily from pimelate. Pimelate is converted to pimeloyl CoA by 6.2.1.14. In the presence of alanine and pimeloyl-CoA, biotin is made with the four enzymes (EC 2.3.1 0.47, EC 2.6.1.62, EC 6.3.3.3, and EC 2.8.16) that are present in the N315 genome	Biotin need not be added in the medium, as the biosynthetic pathway is present. However, ORFs for biotin accessory proteins, such as BioH, BioC, and BioY, are absent
20	Thiamine	2 mg	Four enzymes out of six required for de novo thiamine biosynthesis are absent. Therefore, thiamine should be added in the medium for growth	
21	Nicotinic acid	2 mg	De novo synthesis from aspartate is absent; however, salvage reactions from NAD or nicotinic acid are complete. Salvage pathway enzymes that convert nicotinamide to nicotinase (EC. 3.5.1.19) and nicotinase to β-nicotinate D-ribonucletoide (EC. 2.4.2.11) are present. Further conversion to amido NAD (by EC 2.7.7.18) and to NAD by DNA synthase (EC. 6.3.5.1) is present	Addition of either NAD or nicotinic acid is essential for growth on minimal medium
22	Calcium pantothenate	2 mg	D-Pantothenate can be synthesized from keto-alaine and THF with the use of enzymes (EC.2.1.2.11, EC.1.1.1.69 and EC 6.3.2.1). The genes for these enzymes are present, as well	Exogenously provided pantothenate may not be taken in due to the absence of Na/pantothenate symporter

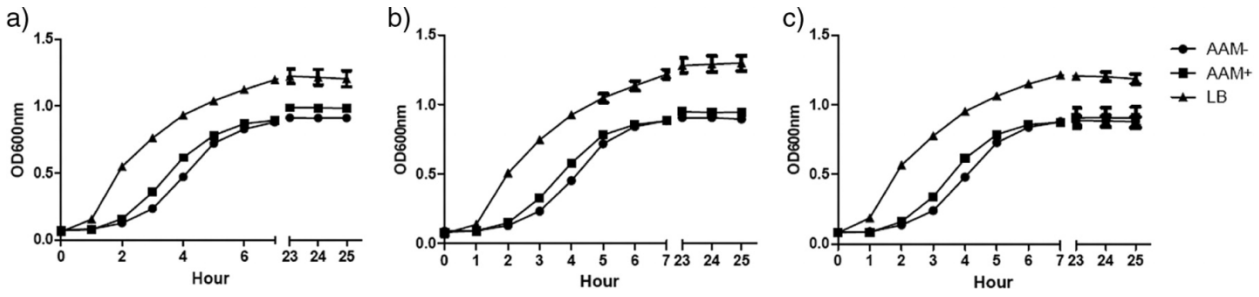


Figure 4.2. Experimental validation of *S.aureus* growth in metabolic-reconstruction-based minimal medium.

Shown are the growth curves of *S. aureus* Mu50 (a), *S. aureus* USA300 (b), and an *S. aureus* patient isolate (c). AAM– is a modified minimal medium derived from metabolic reconstruction; AAM+ is a published minimal medium (Rudin, Sjostrom et al. 1974). Reproduced with permission from (Lee et al., 2009).

4.3 DISCUSSION

The availability of whole-genome sequences not only provides the possibility of identifying gene variability and pathogenic features, but is also valuable for identifying *in silico* metabolic networks and their modeling. Although several virulence and pathogenic genes in *S. aureus* have been identified, they tend not to be feasible drug target candidates, as they are usually strain or species specific or prophage borne. On the other hand, metabolic enzymes tend to be common to all strains and offer broad-spectrum target candidates irrespective of the bacteria or their host specificities.

In this study, we performed metabolic-network reconstructions of fully sequenced *S. aureus* strains via protein similarity calculation, automated annotation, and manual curation. Compared to the previous reconstructions made for a single strain of *S. aureus* (N315), our reconstructions contain a larger data set of both biochemical reactions and transport compounds. Moreover, based on the functional annotation of several *S. aureus* strains, we completed strain-

dependent metabolic reconstructions, which help us to understand the wide range of virulence and drug resistance among the *S. aureus* strains. As the first step, we carried out FBA on the metabolic reconstructions and derived unconditionally essential single enzymes and pairs of enzymes in each strain.

The variation of the reconstructed metabolic networks among the strains does not appear to be great, i.e., over 90% of the metabolic enzymes and reactions are similar in the 13 strains. This was expected, owing to the clonality of the *S. aureus* genomes, with more than 80% overall DNA sequence similarity (Baba, Bae et al. 2008; Feng, Chen et al. 2008). However, in order to describe the metabolic diversity of various drug-resistant phenotypes of *S. aureus*, we performed the FBA for these metabolic networks. In particular, we identified unconditionally essential metabolic enzymes by *in silico* enzyme deletion studies, which in turn represent potential antimicrobial drug targets. These essential enzymes partly overlap among all the strains to about 60% for single essentiality and about 20% for synthetic lethality, suggesting strong strain specificity of the functional states of the metabolic networks. We therefore conclude that development of antimicrobial drugs should take into consideration the strain dependency of the vulnerability of metabolic networks in *S. aureus*. We presented a list of the single essential and synthetic-lethal enzymes for each drug-resistant *S. aureus* strain. Previously published experimental results show that more than half of the single essential enzymes were identified experimentally as being essential in various bacteria, including *S. aureus*.

Two interesting findings of this study are the identification of new single essential enzymes and paired synthetic-lethal enzymes as antibiotic targets common to all strains of *S. aureus*. A surprising finding is that a majority (six pairs) of synthetic-lethal pairs of enzymes are in the amino sugar biosynthesis pathways, which are involved in cell wall metabolism. Three

lethal enzyme pairs were identified in amino acid metabolic pathways. Among the single essential enzymes, eight were identified as playing roles in fatty acid biosynthesis.

It should be noted that the metabolic reconstructions performed in this work are not complete. First, incorrect ORF calling led to inevitable errors in the reconstructions for several reasons (Feist, Herrgard et al. 2009). Second, in contrast to well-studied model prokaryotic organisms, such as *E. coli* and *B. subtilis*, there are limited biochemical and physiological data for *S. aureus*, which hinders testing and improving the metabolic networks reconstructed in this study, in spite of the accessibility of literature-based curation (Feist, Herrgard et al. 2009). For instance, protein complexes and isoenzymes in the metabolic network of *S. aureus* have not been identified as extensively as in *E. coli* (Feist and Palsson 2008). Given this limitation, we assumed that each of the genes associated with a reaction was needed to catalyze the reaction, and this might have created false positives in our essentiality predictions. Moreover, the biomass components of *S. aureus* have not yet been fully described, forcing us to use those of the related gram-positive bacterium *B. subtilis* or those derived from literature searches. Future biochemical experiments on S. *aureus* will improve and confirm the results obtained in this study. It would also be informative to perform experimental gene deletion studies to test our findings.

4.4 MATERIALS AND METHODS

Metabolic reconstruction of multidrug-resistant *S. aureus* strains.

In this study, we carried out the metabolic reconstruction of multiple *S. aureus* genomes using the ERGO bioinformatics suite (Overbeek, Larsen et al. 2003) and the KEGG ligand/reaction database as of 2007 (http://www.genome.jp/ligand) (Kanehisa and Goto 2000).

Initially, we focused on the MRSA strain N315 as a model organism for metabolic reconstruction and FBA studies. Its whole-genome sequence (NC_002745) and the sequence for plasmid pN315 (NC_003140) were obtained from NCBI (GenBank) and downloaded into ERGO, a genome informatics platform that is developed and maintained by Integrated Genomics, Inc. Similar processes were performed on several other sequenced genomes, including *S. aureus* Mu50 (NC_002758), *S. aureus* MW2 (NC_003923), *S. aureus* subsp. *aureus* COL (NC_002951), *S. aureus* EMRSA-16 strain 252 (NC_002952), *S. aureus* methicillin-sensitive strain 476 (NC_002953), *S. aureus* subsp. *aureus* JH1 (NC_009632), *S. aureus* subsp. *aureus* JH9 (NC_009487), *S. aureus* RF122 (NC_007622), *S. aureus* subsp. *aureus* COL (NC_002951), *S. aureus* subsp. *aureus* USA300 (NC_007793), *S. aureus* subsp. *aureus* USA300_ TCH1516 (NC_010079), and *S. aureus* subsp. *aureus* Newman (NC_009641).

In general, the whole-genome DNA sequences, along with associated open reading frames (ORFs)/protein sequences, were integrated into the ERGO genome database. Automated and manual annotations of the ORFs were carried out as described previously (Kapatral, Anderson et al. 2002). Briefly, automatic ORF calling was performed using a combination of programs, including GLIMMER and Critica. The protein similarities were computed by BLAST "all against all," with over 3.5 million protein sequences present in the non-redundant database from over 1,362 genomes. This was followed by functional assignments, relationship computation, pathway analyses, and automated metabolic reconstruction (metabolic and nonmetabolic subsystems). These automated annotations are largely based on the existence of ortholog and protein family clusters. Manual annotation included a review of every gene in the genome. Once the functional annotations were assigned, we performed additional manual pathway curations of the functions in order to identify their roles in appropriate metabolic

subsystems. The automatically assigned functions were validated using both proprietary and publicly available tools to make new functional assignments. During manual annotation, we also considered and reconciled motif/domain database results (for example, from COGS, Pfam, and INTERPRO) to make functional assignments. Missing steps in pathways were identified, and genes that could fulfill these functions were sought in order to determine if the pathways should be included in the organism's metabolic catalog. This process resulted in the identification of functions for un-annotated or erroneously annotated proteins. Metabolic/cellular reconstructions were finally derived by interconnecting the entire set of pathways identified in a given strain. These pathways in ERGO were grouped into metabolic and non-metabolic systems that were interconnected into a metabolic network between subsystems, such as amino acids, carbohydrates, lipids, secondary metabolism, and sulfur and phosphorus metabolism. The non-metabolic pathways contained lists of genes/functions that represented subsystems, such as protein secretion, virulence, secretion, and phages. All 13 *S. aureus* genomes were simultaneously annotated and analyzed. Gap-filling reactions were not performed because the ORFs corresponding to the missing enzymes were not present. Unlike in the work of Becker and Palsson , where the gap-filling reactions were performed, the aim of this study was to identify the genes that could be experimentally tested for essentiality. In cases of missing steps within a given pathway, we searched for orthologs or published biochemical activities (for example, for the histidine biosynthesis pathway). ORFs with enzymes as functions with complete or incomplete Enzyme Commission (EC) numbers were identified from the functional categories of the ERGO genome analysis suite. The associated biochemical reactions for each of the enzymes were selected in the KEGG reaction database. The biochemical reactions with substrates that did not have an identified transport or uptake system and those that were not supported by experimental

studies of *S. aureus* were considered "unlikely reactions" and were not included in the FBA. Individual transport reactions were added from the ERGO pathway collection. Universal (SwissProt) names were used for enzymes to avoid nomenclature conflict for the FBA studies. All the reactions and their corresponding KEGG identifiers (reaction IDs and compound IDs) were used in the FBA computations.

FBA of *S. aureus* metabolic networks.

The metabolic networks were constructed individually for all 13 *S. aureus* strains, and the reaction data from each of the species were used for FBA computations. The total numbers of enzymes of all the species ranged from 522 to 547. The biochemical reactions varied from 1,444 to 1,497, with total metabolites ranging from 1,399 to 1,437. The metabolic reactions were classified into three categories: cellular (reactions in the cytoplasm), transport (reactions involving both the intra- and extracellular metabolites), and exchange (reactions involving either uptake or excretion of the extracellular metabolites). The stoichiometries of these metabolic reactions were adapted from the KEGG ligand database and the ERGO bioinformatics suite, yielding the stoichiometry matrix S, with its element S_{ij} indicating the stoichiometric coefficient of the metabolite i in the reaction j. The FBA states that in the stationary state, the fluxes $\{v_j\}$ of the metabolic reactions are those that maximize v_{biomass} subject to:

$$\begin{gathered} S \cdot v = 0 \\ v_j \geq 0 \text{ for all irreversible reactions j} \\ v_j \leq v_{\max,\text{j}} \text{ for all uptake reactions j} \end{gathered} \tag{1}$$

where v_{biomass} indicates the flux of the reaction that produces biomass, acting as a drain of those biomass compounds. The biomass-generating reaction is given by $\sum_i c_i m_i \rightarrow \text{biomass}$ in which m_i is

each of the biomass compounds and c_i is the coefficient. Currently, there are two predicted biomass compositions proposed for *S. aureus* strain N315 (Becker and Palsson 2005; Heinemann, Kummel et al. 2005): one is based on and similar to that of a related gram-positive bacterium, *Bacillus subtilis* (Becker and Palsson 2005), with 58 biomass compounds, and the second was established from an extensive literature search on *S. aureus* resulting in 50 biomass compounds (Heinemann, Kummel et al. 2005). For each set of biomass components, we first checked whether each of the biomass compounds, m_i, was synthesized in our metabolic-network reconstructions by performing the FBA assuming the biomass-generating reaction to be $m_i \rightarrow$ biomass for each m_i. If this reaction had a nonzero flux, we considered the biomass compound to be synthesized in the given strain of *S. aureus*. For the biomass composition from (Becker and Palsson 2005), we found that for nine *S. aureus* strains (TCH1516, USA300, Mu3, JH9, JH1, RF122, Mu50, MRSA252, and N315), 47 of the 58 biomass components could be synthesized, that is, the flux of the modified biomass-generating reaction for each of these compounds was nonzero. In contrast, the other 11 components could not be synthesized, suggesting absence of the biosynthetic pathways to generate them or absence of the uptake transporters for them. Three *S. aureus* strains, Newman, COL, and MW2, could synthesize only 44 biomass components, whereas 43 components could be synthesized by the methicillin-sensitive strain 476. For the biomass composition of (Heinemann, Kummel et al. 2005), each of the 13 strains synthesized 41 compounds of the proposed 51. For the missing components, one could assume that unknown gene products catalyzed the corresponding biosynthetic pathways and therefore add them to the metabolic network. Those added pathways would be likely to be essential for the growth of *S. aureus*. However, we elected to retain only the biomass components synthesized in the current form of the metabolic networks of *S. aureus* strains to obtain definitely and unconditionally

essential enzymes. These computational studies predicted the synthesis of 47 and 41 biomass compounds, respectively, with 25 common compounds (see Dataset 1 in the supplemental material). We considered both biomass compositions from Becker and Palsson and Heinemann in our study for comparison. Therefore, we used two metabolic-reconstruction versions for each *S. aureus* strain. The two versions were identical except for the biomass generation reaction involving biomass compounds from either Becker and Palsson or Heinemann et al. this study, we present the results obtained from the metabolic reconstructions with the biomass composition based on the studies of Becker and Palsson. The results using the biomass composition from Heinemann et al. can be found in the supplemental material (Lee, Burd et al. 2009). For the FBA computations, we used ideally rich medium in which all the uptake reactions could occur without limitation on their fluxes, since we aimed to identify unconditionally essential enzymes, which we define below. The FBA is a linear programming problem and was carried out using the GNU Linear Programming Kit.

Identification of single unconditionally essential and synthetic-lethal enzymes of *S. aureus* strains.

In the FBA scheme, an enzyme is considered to be essential (Almaas, Oltvai et al. 2005) if an additional constraint,

$$v_j = 0 \qquad 2)$$

for all reactions *j* catalyzed by the selected enzyme lead to no growth ($v_{\text{biomass}} = 0$). If inactivation of an enzyme leads to no growth in an ideally rich environment (medium) in which all the uptake reactions are allowed to have unlimited fluxes, it would also lead to no growth in any other environment. Therefore, the unconditionally essential enzymes, i.e., 3)

essential under all growth conditions, can be obtained from the FBA equation 1 combined with equation 2 and setting all upper limits of uptake reactions as follows:

$$v_{\max,j}=\infty$$

for all uptake reactions *j*. Similarly, an unconditionally synthetic-lethal pair of enzymes can be identified by monitoring the biomass generation reaction while blocking the reactions catalyzed by the considered pair of enzymes in an ideally rich environment. That is, the removal of two enzymes, E1 and E2, is considered to be unconditionally lethal if the removal of just one of them is not lethal and the FBA (equation 1) in the ideally rich environment (equation 3) leads to no growth ($v_{biomass}$ = 0) (Deutscher, Meilijson et al. 2006) under the following constraints:

$$v_j=0 \qquad 4)$$

for all uptake reactions *j* catalyzed by the enzyme E1 or E2.

Design and analysis of growth media.

Using the complete metabolic analysis of all the strains of *S. aureus* and the reference minimal medium, we derived a medium that supports growth on a minimal medium, using metabolic-reconstruction methods, as described for *Xylella fastidiosa* (Bhattacharyya, Stilwagen et al. 2002). The compounds from the published minimal medium (AAM^+) for *S. aureus* (Rudin, Sjostrom et al. 1974) were verified with the metabolic reconstruction of *S. aureus* genomes. A modified synthetic minimum medium was predicted that did not contain L-arginine, L-glutamic acid, L-valine, L-threonine, L-phenlyalanine, and biotin (AAM^-). The AAM^+ and AAM^- media were calibrated to pH 7.4 to prevent salt precipitation, and LB was prepared using standard-grade chemicals. Three strains, including two MRSA strains (*S. aureus* Mu50 and USA300) and a

methicillin-sensitive patient isolate (from the University of Pittsburg Medical Center), were tested for growth in the minimal media under identical medium and growth conditions. These strains were inoculated into a sterile glass tube containing 3 ml LB and were grown at 37°C at 200 rpm for 16 h. For each strain, 900 µl of the overnight culture was pipetted into three separate 1.5-ml microcentrifuge tubes, centrifuged, and washed with LB, AAM^+, or AAM^- medium, respectively. The individual pellets were then resuspended in 200 µl of LB, AAM^+, or AAM^- medium, and 10 µl of each suspension was then inoculated into 2 ml of the corresponding medium for culture in a 24-well plate with uninoculated LB, AAM^+, or AAM^- medium as a control. The cultures were placed on a shaker rotating at 225 rpm at 37°C. Each experiment was done in triplicate and was repeated three times, and the growth was measured for 25 h at 600 nm using a microplate reader.

4.5 ACKNOWLEDGEMENT

Support for this work was provided by NIH grant 5U01AI070499-02.

5.0 FASII AS A VALID ANTIMICROBIAL TARGET AGAINST S. AUREUS METABOLIC NETWORKS

5.1 INTRODUCTION

Antibiotic drug resistance significantly eroded the effectiveness of currently available antimicrobial drugs towards disease-causing bacteria (Fischbach and Walsh 2009). As a consequence, today the yearly mortality rate in the United States due to multi-drug resistant *S. aureus* infections is higher than due to AIDS (Klevens, Morrison et al. 2007). Therefore, the emergence of superbugs, methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant *S. aureus* (VRSA) isolates, underscores the need to identify new targets and molecules effective against multidrug-resistant strains of *S. aureus.*

The challenge to find "druggable" targets (Russ and Lampel 2005) that will have an effect in the complex interaction network of a living organism has only been partially met by genomics and functional genomics approaches (Orth, Batalov et al. 2004). For example, systematic gene deletion studies have been widely used to identify essential genes whose protein product could serve as potential antibiotic targets in a given bacterium. This approach yields only growth-condition specific results, as a molecular target that is found to be essential in one specific environment may not be essential in others. Our previous study (Lee, Burd et al. 2009) utilized comparative systematic analysis on the reconstructed metabolic networks of 13 *S. aureus* strains to predict a list of single essential enzymes and synthetic lethal pairs as essential under all conditions and thus representing potential targets for antimicrobial drug discovery. Here, to

provide a proof-of-principle demonstration that the predicted single essential enzymes or synthetic lethal pairs through metabolic network analysis can yield potential "druggable" targets, we focus on the initiation step and elongation cycle of the fatty acid biosynthesis pathway (FAS II) from the predicted list as a case study. We elected to focus on this pathway because the FASII elongation cycle is catalyzed by enzymes fundamentally different from their eukaryotic counterpart (Campbell and Cronan 2001), several antimicrobials have already successfully targeted its enzymes in various bacteria (Heath, White et al. 2001; Wang, Kodali et al. 2007; Abbanat, Morrow et al. 2008; Chan and Vogel 2010), and because many of its enzymes were shown to be essential in *E. coli* under select growth conditions (Gerdes, Scholle et al. 2003; Kang, Durfee et al. 2004; Baba, Ara et al. 2006).

In conjuction with computer-aided virtual screening of small-molecule chemical libraries against the enzymes' active sites, we discovered forty potential inhibitors of FabD, FabH, FabB/F, FabI, and FabA/Z, of which a small subset inhibited both enzyme activities *in-vitro* and bacterial cell viability *in vivo*. Therefore, the systems-analysis of metabolic networks identified FAS II as a valid anti-Staphylococcus target, representing an effective approach for antimicrobial target identification.

5.2 RESULTS

5.2.1 Identification of antimicrobial targets by the analysis of bacterial metabolic networks

Our previous study (Lee, Burd et al. 2009) performed comparative systematic analyses on the reconstructed metabolic networks of 13 *S. aureus* strains to predict a list of single essential enzymes and synthetic lethal pairs under all growth conditions. To identify generic targets against both Gram-positive (*S. aureus*) and Gram-negative (*E. coli*) bacteria, we performed a similar Flux Balance Analysis (FBA) to identify the essential metabolic reactions of *E. coli* MG1655 using a recent metabolic network reconstruction of this Gram-negative model organism (Feist, Henry et al. 2007). This analysis predicted 38 metabolic reactions as having nonzero flux under all growth conditions and being indispensable for the synthesis of the full set of biomass components in this bacterium. In the absence of isozymes or compensatory (non-specific) enzymatic activities, the enzymes catalyzing these reactions are expected to be essential under all conditions and thus represent potential targets for antimicrobial drug discovery (Almaas, Oltvai et al. 2005; Lee, Burd et al. 2009). A high fraction of the enzymes catalyzing the 38 predicted indispensable reactions were found to be essential in three previous genome-scale gene deletion studies (Gerdes, Scholle et al. 2003; Kang, Durfee et al. 2004; Baba, Ara et al. 2006). Seven of these indispensable reactions are shared among *E. coli* MG1655 and 13 *S. aureus* strains (Lee, Burd et al. 2009), with five of them being experimentally validated as essential in at least one of the *E. coli* gene deletion studies (Table 5.1). These reactions are not distributed randomly within the metabolism but cluster into distinct metabolic pathways that are known to play key roles in bacterial cell wall-, amino acid-, and porphyrin biosyntheses.

Table 5.1 Enzymes catalyzing essential metabolic reactions conserved between *E. coli* and *S. aureus* strains may represent novel antibiotic targets.

	Gerdes	Baba	Kang
aminosugars metabolism			
murB			
amino acid biosynthesis			
aroA			
aroC			
porphyrin metabolism			
hemB			
hemD			
methionine metabolism			
metK			
mtn			
fatty acid biosynthesis			
fabD			
fabH			
fabF			
fabG			
accA			
accD			

Top: the list of enzymes of shared unconditional essential metabolic reactions of *E. coli* MG1655 and 13 *S. aureus* strains predicted from FBA analysis. Bottom: enzymes of fatty acid biosynthesis that are unconditionally essential in both bacteria in the absence of exogenous unsaturated fatty acids. Predicted enzyme targets with uniform agreement with *E. coli* viability studies (red color), with one or two mismatch (blue color), or in disagreement with all 3 *E. coli* viability studies (green color). Background color of orange (grey) indicates the enzyme-encoding gene is identified as essential (non-essential) in the indicated *E. coli* deletion phenotype study (Gerdes, Scholle et al. 2003; Kang, Durfee et al. 2004; Baba, Ara et al. 2006).

Some reactions are indispensable only in one type of bacterium. For example, most enzymes of the type II bacterial fatty acid biosynthesis (FAS II) pathways (Zhang and Rock 2008) are predicted as unconditionally essential in 13 *S. aureus* species (Lee, Burd et al. 2009), but not in *E. coli* MG1655 that has several transporters that enable it to take up fatty acids directly from the environment (Feist, Henry et al. 2007). Yet, in standard (fatty acid-free) growth media the inactivation of most FAS II enzymes renders *E. coli* cells inviable (Table 5.1). This can be countered by the addition of unsaturated fatty acids to the growth media (Harder, Ladenson et al. 1974; Campbell and Cronan 2001). In a recent intraperitoneal infection model

with gram positive *S. agalactiae* inactivation of FAS II enzymes also proved ineffective (Brinster, Lamberet et al. 2009), suggesting that pharmaceutical inhibition of FAS II enzymes may not be a feasible strategy for the treatment of systemic bacterial infections. However, these enzymes may be suitable targets for the eradication of cutaneous antibiotic-resistant bacterial colonizations or for treating cutaneous infections caused by e.g., methicillin-resistant *S. aureus* (MRSA) (Jones, Sura et al. 2003; Kazakova, Hageman et al. 2005).

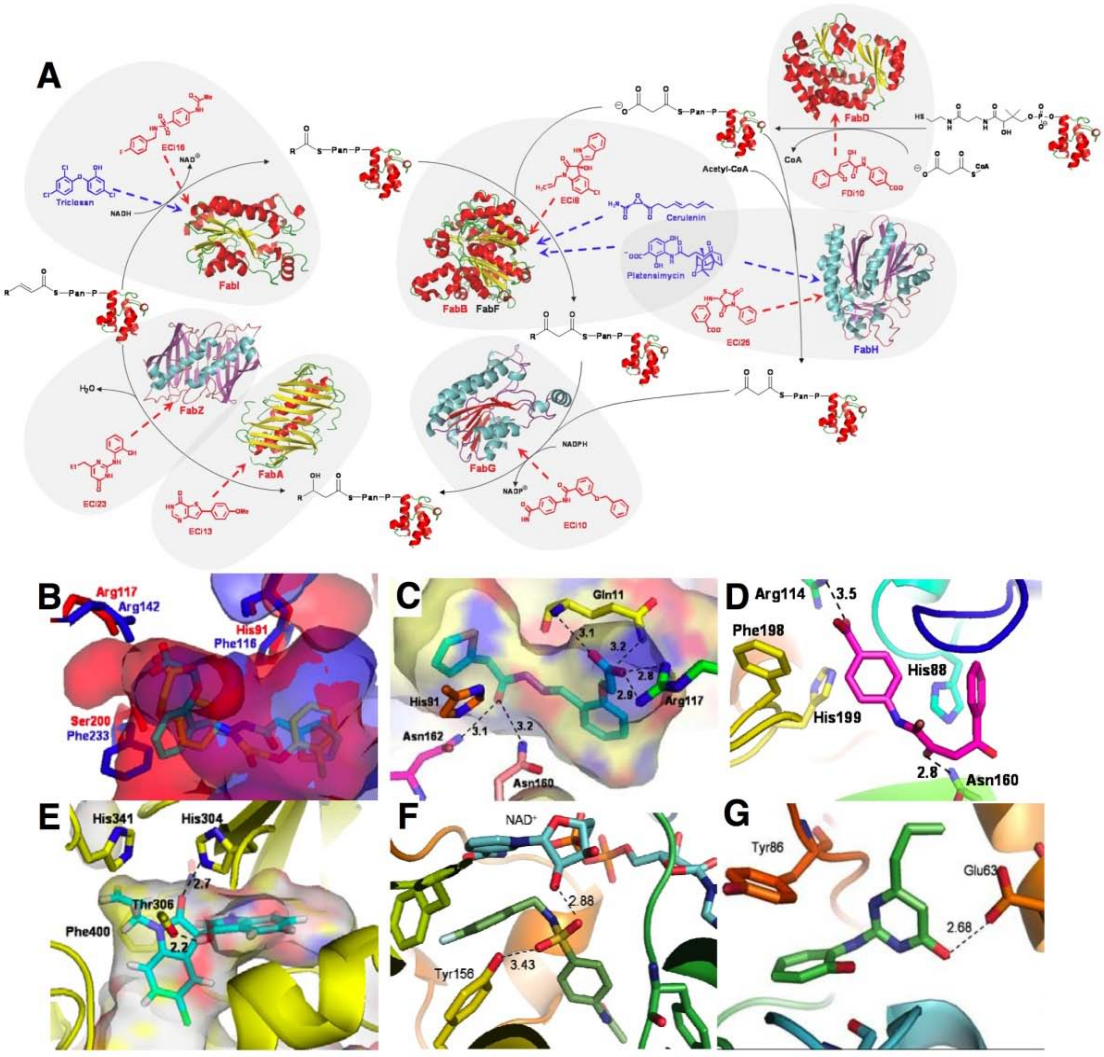


Figure 5.1 FAS II pathway and Small molecule inhibitors against FAS II enzymes

a. FAS II pathway and structures of select, existing antimicrobials (blue) and the newly identified small molecule inhibitors (red) against their target enzymes. The structures used for the virtual screenings are

also shown. *E. coli* viability data for FAS II pathway enzymes from three independent gene deletion studies are indicated: uniformly essential (red), dispensable (black) or conflicting deletion phenotype (blue) enzymes are shown.
b. The active site of *E. coli* FabD (PDB#2g2z, in red) superimposed on human mitochondrial MCAT (PDB#2c2n, in blue) with two docked ligands, FDi2 (cyan) and FDi14 (yellow); **c**-**g**. Interactions of selected inhibitors in the active sites of their respective targets: **c**. FDi2 in *E. coli* FabD after 8ns MD; **d.** FDi10 docked to *S. aureus* FabD homology model built from *E. coli* FabD; **e**. ECi8 docked to *S. aureus* FabF; **f**. ECi16 docked to *E. coli* FabI; and **g**. ECi23 in a homlogy model of *E. coli* FabZ after 8ns MD. Reproduced from (Shen and Liu et al., 2010).

5.2.2 Virtual screening of small-molecule compounds against enzymes of bacterial fatty acid biosynthesis

Enzymes of the FAS II pathway with uniformly lethal deletion phenotype in *E. coli* are considered as potential high-confidence targets, while those with at least one mismatch in the *E. coli* deletion phenotype data, or having isozymes, were considered lower-confidence targets (Table 5.1). Previous analyses demonstrated that the essential metabolic enzymes of *E. coli*, and particularly those in the metabolic core, tend to be conserved among bacteria on the amino acid sequence level (Gerdes, 2003; Almaas, 2005). Therefore, we next examined the amino-acid level conservation of those enzymes of bacterial fatty acid biosynthesis that have at least one experimentally-determined crystal structure among the fifteen most common human bacterial pathogens found in hospital patient populations at the University of Pittsburgh Medical Center. Of these enzymes, we first focused on the Malonyl-CoA-acyl carrier protein transacylase (MCAT, FabD, EC 2.3.1.39) which catalyzes the first committed step in the initiation step of the fatty-acid biosynthesis pathway (Fig. 5.1a).

Structural conservation of an enzyme target among prokaryotes, and especially the conservation of its active site, is one prerequisite for broad-spectrum antibiotic development. A second prerequisite is that either no human ortholog exist for that enzyme, or that it is

sufficiently different from its bacterial counterparts, in order to minimize potential drug toxicity. In humans, the elongation cycle enzymes are fundamentally different from their bacterial counterparts (Campbell and Cronan 2001; Maier, Leibundgut et al. 2008). However, FabD exists in a cytosolic and mitochondrial form, the former playing the major role for fatty acid biosynthesis in human cells (Campbell and Cronan 2001; Heath and Rock 2004; Maier, Leibundgut et al. 2008). For human mitochondrial FabD, a crystal structure is also available (PDB ID 2c2n). Hence, we first performed a quantitative sequence homology search between *E. coli* FabD and the fifteen most common hospital isolates from the University of Pittsburgh Medical Center. We found that the bacterial enzymes' active sites show high amino acid sequence conservation (Fig. 5.2a), while the human mitochondrial enzyme has some critically different residues. Additionally, human mitochondrial FabD active site showed the largest evolutionary distance compared to all examined bacterial species (Fig. 5.2b). Figure 5.2c displays the *E. coli* FabD structure (right panel) and its active site (left panel), including catalytic and substrate binding residues and bound substrate of the enzyme (malonate).

On a structural level, overall secondary structure of human FabDs showed similarity to its bacterial counterparts (see main text). However, a detailed structural comparison of the *E. coli* and human mitochondrial and cytosolic FabD active sites reveal sufficient differences between the human and prokaryotic enzymes to suggest that bacterial FabD may indeed serve as a useful molecular target for novel antibiotics.

Subsequently, we have identified 15 potential inhibitors (FDi1-15) of bacterial FabDs (Serre, Verbree et al. 1995; Keatinge-Clay, Shelat et al. 2003; Oefner, Schulz et al. 2006; Li, Huang et al. 2007). Briefly, the ZINC lead library (Irwin and Shoichet 2005) containing ~10^6 small molecules pre-filtered for drug-like properties was docked either to the crystal structures of

E. coli FabD or to a FabD homology model derived for *S. aureus* N315 using successively more accurate scoring functions, followed by manual inspection of poses and rescoring by MM-PBSA calculations from an ensemble of molecular dynamics (MD) simulations (see Supporting Information for details) (Thompson, Humblet et al. 2008). Figure 5.1b shows two representative potential inhibitors, FDi2 and FDi14, bound to the active site of *E. coli* FabD. They exemplify the pharmacophore that is further confirmed by other docked structures and the putative binding mode of the natural substrate (Oefner, Schulz et al. 2006). The close-up of FDi2 in the active site of *E. coli* FabD illustrates the relevant interactions (Fig. 5.1c). The carboxylate acts as a phosphate isostere and interacts with the catalytic Arg^{117}, Leu^{93} and Gln^{11}, respectively. A polar linker to a thiophene ring enables stacking against His^{91} and occupation of a hydrophobic pocket by a benzene moiety. Similar interactions are found for FDi10 in the active site of *S. aureus* FabD (Fig. 5.1d).

A particular strength of the approach used here is that it has an inherently systems-wide view rather than targeting a single enzyme, as is done in traditional drug design. If an enzyme is essential by itself (e.g., FabD) its specific inhibitors should be able to kill bacteria. Similarly, isozyme pairs, such as FabB and FabF in *E. coli*, are also promising targets if both of them are targeted simultaneously. For example, FAS II inhibitors thiolactomycin, cerulenin or platensimycin are known to simultaneously inhibit both FabB and FabF (Fig. 5.1a) and have broad-spectrum antibacterial activity (Wang, Soisson et al. 2006).

We therefore screened all remaining enzymes of the FAS II pathway (FabH, FabB/F, FabG, FabA/Z, and FabI) (Fig. 5.1a) using ligand bound crystal structures or homology models. This yielded 26 small molecules that can potentially inhibit the FAS II elongation cycle enzymes

(ECi1-26). As several of the enzymes are functionally redundant in some organisms and the product of one step is the substrate of the next, the binding sites share many similarities. As a

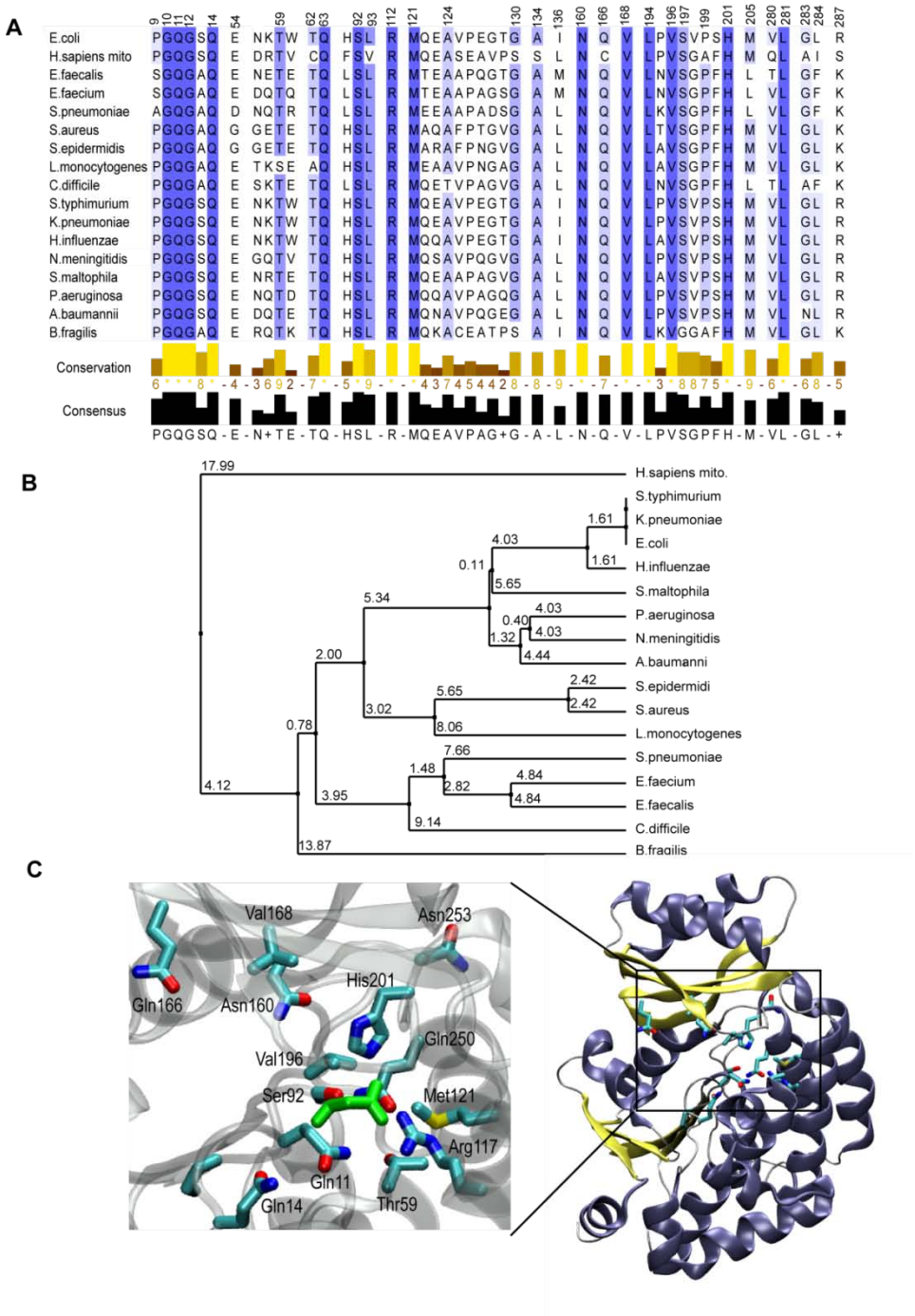


Figure 5.2. Comparison of FabD sequence and structure between human mitochondrial and the 16 most common bacterial pathogens (among hospital isolates).

a. Sequence similarities of the FabD active sites. The residues lining the active site of FabD of human mitochondria and the bacterial species show some distinct features. The residues are labeled with their corresponding residue numbers in E. coli FabD. The middle panel shows the conservation score for each residue. Stars indicate full conservation among all species. The bottom panel shows the most common residues in the active sites.

b. Evolutionary distance of FabDs. The phylogenetic tree shows the evolutionary relationships among pathogenic bacteria with existing FabD sequence and human mito-chondrial FabD. The numbers on the tree branches indicate the average distance measured in sequence identity.

c. The structure of *E. coli* FabD. The right panel shows the ribbon diagram of *E. coli* FabD, while the left panel is a closer look at its active site. Catalytic and substrate binding residues are displayed and labeled. A bound substrate of the enzyme (malonate) is shown in green. PDB IDs that are used for the figure are 1mla, 2g2y, 2g2z. Reproduced from (Shen and Liu et al., 2010).

result, similar common binding moieties, such as a phosphate isostere, a hydrogen binding polar group, and a nonpolar ring system in similar position emerge from the screens. This suggests the possibility of inhibiting several targets simultaneously. This concept of poly-pharmacology has been successfully applied in a number of cases before (Lehár, Stockwell et al. 2008) and is part of the inherent design in our case.

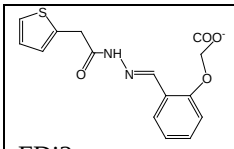
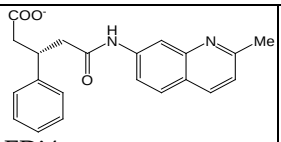
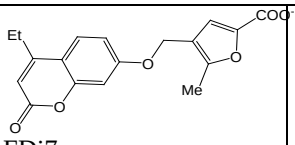
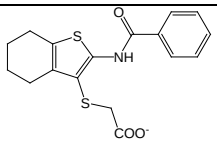
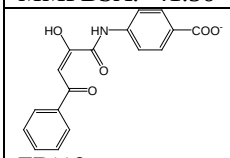
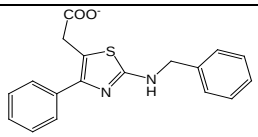
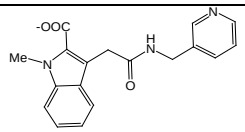
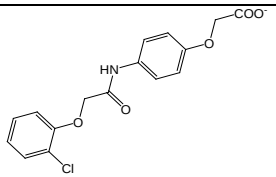
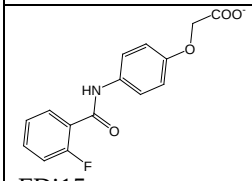
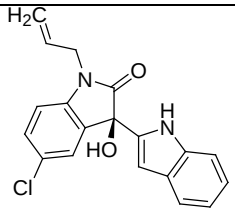
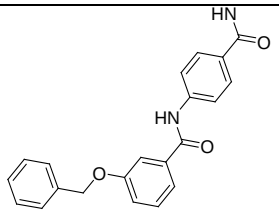
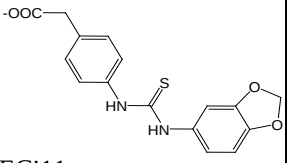
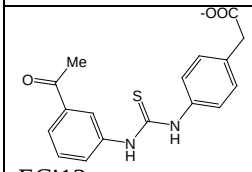
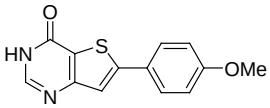
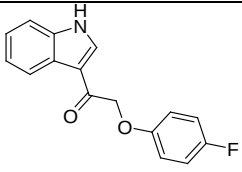
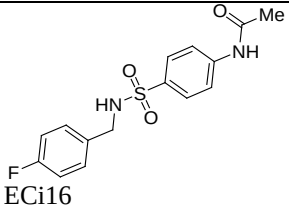
Representative results from these studies are shown in Figures 5.1e-g. ECi8 is predicted to inhibit both FabB and FabF (Fig. 5.1e), shares the binding characteristics of known inhibitors (Wang, Soisson et al. 2006; Wang, Kodali et al. 2007), and has a good shape complementary to the active sites of both enzymes. In agreement with the hypothesis outlined above, cross-docking shows that ECi8 also has good poses and scores for binding to FabI and FabG, the adjacent enzymes in the elongation cycle.

The NADPH and NADH dependent reductases FabG and FabI have similar active sites and several inhibitors exist for the latter, including triclosan (Fig. 5.1a). The most promising inhibitor identified by virtual screening, ECi16, shows similar interactions as triclosan: the fluorophenyl ring π-stacks with the oxidized nicotinamide while the oxygen atoms of the sulfamide group participate in hydrogen bond interactions with the conserved active site tyrosine (Y^{156} in *E. coli* FabI) and the 2'hydroxyl of NAD^+ (Fig. 5.1f).

The active sites of the dehydratases FabA and FabZ are formed along the dimer interface, with the critical His^{70} and Asp^{84} (FabA)/Glu^{63}(FabZ) contributed by different monomers. The virtual screening identifies the interaction with Asp^{84}/Glu^{63} as the most relevant for inhibitor

binding. Figure 18g depicts this interaction in the binding of ECi23 in the FabZ active site, where the inhibitor is also stabilized by π-stacking with Tyr[86]. A complete list of the identified potential FabD and elongation cycle inhibitors are presented in Table 5.2.

Table 5.2. Structures of compounds showing activity in in vitro experiments

 FDi2 ZINC0072513 2g2y GScore: -6.09 MMPBSA: -41.56	 FDi4 ZINC00152793 2g2y GScore: -9.94 MMPBSA: -42.61	 FDi7 ZINC00200929 2g2y GScore: -8.18 MMPBSA: -52.72	 FDi8 ZINC00284275 2g2y GScore: -12.27 MMPBSA: -33.51
 FDi10 ZINC02968391 2g2y GScore: -10.5 MMPBSA: -34.22	 FDi11 ZINC03350762 2g2y GScore: -7.28 MMPBSA: -40.94	 FDi12 ZINC03660299 2g2y GScore: -8.54 MMPBSA: -31.81	 FDi14 ZINC04819675 2g2y GScore: -7.94 MMPBSA: -18.89
 FDi15 ZINC06702282 2g2y GScore: -7.35 MMPBSA: -31.07	 ECi8 ZINC02214825 2gfv GScore: -11.67 MMPBSA: -32.52	 ECi10 ZINC06697642 1q7b GScore -9.10 MMPBSA -25.67	 ECi11 ZINC04618181 1hnj GScore: -10.52 MMPBSA: -34.24
 ECi12 ZINC00621461 1hnj GScore: -11.36 MMPBSA: -31.84	 ECi13 ZINC04716354 1mka GScore -12.98 MMPBSA -23.7	 ECi14 ZINC04487283 1qg6 GScore -13.8 MMPBSA -28.3	 ECi16 ZINC00433116 1qg6 GScore -13.42 MMPBSA -36.71

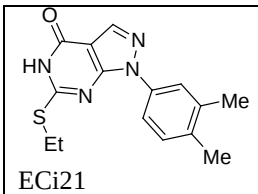
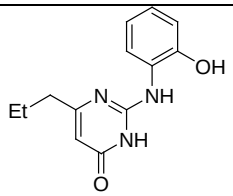
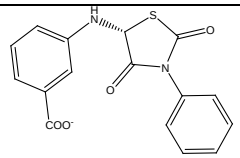
 ECi21 ZINC04928462 2gfx GScore: -11.03 MMPBSA: -37.35	 ECi23 ZINC01281455 Homolog built on 1z6b GScore -15.07 MMPBSA -42.84	 ECi26 ZINC01317935 1hnj GScore: -8.38 MMPBSA: -40.79	

Table 5.2 includes the numbering of the compounds used in the study, their corresponding ZINC numbers, the PDB # of the crystal structures used to screen against in the computational studies, Gscore (kcal/mol) from docking using Glide XP and MMPBSA score (kcal/mol)from MMPBSA calculation based on 8ns molecular dynamics simulation using Amber9. Reproduced from (Shen and Liu et al., 2010).

5.2.3 Testing the effect of predicted inhibitors by *in vitro* enzyme assays

To experimentally validate the inhibitory effect of the 41 compounds selected from the 10^6 library members on bacterial fatty acid enzymes, we set up two separate published enzyme screening assays, one for testing FabD activity (Molnos, Gardiner et al. 2003) and the other for testing the inhibition of the elongation cycle (Kodali, Galgoci et al. 2006). For the FabD screening assay, we cloned the *E. coli* FabD, and its protein substrate, AcpP, into a prokaryotic expression vector, induced their expression, and purified them to homogeneity. We then tested the purified FabD's activity in the absence or presence of 400μg/ml of the predicted inhibitors (see Table 5.3). Five inhibitors completely block the enzymatic activity of FabD (yellow shading) and generate negatively changing fluorescence signals (NADH), while the three weak inhibitors (orange shading) can significantly decrease the rate of fluorescence signals catalyzed by the purified FabD's activity compared to that in the presence of DMSO or the absence of inhibitors (see Table 5.3). Since the established FabD enzyme assay (Molnos, Gardiner et al. 2003) was based on the FabD-KDH coupled reactions, we also performed KDH counter-

screening assay to eliminate the possible inhibition of potential inhibitors against KDH. FDi11, 12, 14, and 15 did not show any inhibition against KDH compared to DMSO control, while FDi2, 7, 8, 10 can partially inhibit the KDH activity.

Table 5.3 Summary of the in vitro enzyme assays of FabD inhibitors.

	FabD assay			KDH assay
Slopes	1	2	3	
Control	40.856	34.51	40.1925	-
DMSO	30.972667	31.410333	35.731	-
FDi1	32.251	29.645	30.805	**N/A**
FDi2	-13.557	-10.892	-17.063	±
FDi3	23.044	26.97	23.515	**N/A**
FDi5	25.257	26.247	21.728	**N/A**
FDi6	23.515	26.285	22.732	**N/A**
FDi7	-575.525	-401.568	-623.088	±
FDi8	-4.537	-8.743	-7.943	±
FDi9	22.597	25.545	24.113	**N/A**
FDi10	1.492	2.375	1.615	±
FDi11	2.292	-8.643	-17.61	-
FDi12	18.667	15.572	19.612	-
FDi13	23.272	31.678	22.952	**N/A**
FDi14	21.112	24.532	19.297	-
FDi15	16.681	18.493	14.22	-

Strong inhibitors (yellow shading) and weak inhibitors (orange shading) are shown. +++: strong inhibition; ±: partial inhibition; -: no inhibition in KDH counter screening.

However, the partial inhibition of fluorescence signals catalyzed by KDH enzyme in the presence of FDi 2, 7, 8 and 10 could also be attributed to the direct destruction of NADH by the inhibitors. The strong negative slope suggests this possibility for the case of FDi7. Therefore, of the 15 predicted inhibitors, we found that five of them (FDi2, 7, 8, 10, and 11) strongly- and three of them (FDi12, 14 and 15) weakly inhibited FabD activity .

We also tested the effect of the 26 small molecules that are predicted to block the activity of various elongation cycle enzymes using an established, fractionated cytoplasmic protein

extract-based elongation cycle assay (Kodali, Galgoci et al. 2006) with slight modifications (see Materials and Methods). ECi8, 21, 23, 26 have shown strong- and ECi10, 11, 12, 13, 14, 16 weak inhibitions toward the FAS II elongation cycle enzymes (Table 5.4). Of the 41 potential inhibitors identified in the virtual screening protocol, eight were highly- and nine were weakly active. For each enzyme of the FAS II elongation cycle we could identify at least one validated inhibitor. This exceptionally high hit rate of >40% is remarkable considering the simple physical models used in the scoring functions and is significantly higher than what is typically obtained in virtual screening. This indicates that MM-PBSA rescoring is at least for the present systems able to describe side chain flexibility and hydrophobic effects and thus represents a significant improvement for virtual screening (Thompson, Humblet et al. 2008). It is expected that explicit treatment of entropy could further improve the hit rate, albeit at significant computational cost.

5.2.4 Testing the effect of predicted inhibitors by *in vivo* cell-based experiments

Next, we examined the effect of these compounds on bacterial cell viability by performing disc inhibition assays against two *E. coli* strains (MG1655 and a random patient isolate) and three *S. aureus* strains (*S. aureus* Mu50, USA300, and a strain isolated from a patient) using standard procedures (Woods and Washington 1995). As it was reported that serum is a rich source of unsaturated fatty acids excluding the use of FAS II inhibitors for the treatment of systemic infections (Brinster, Lamberet et al. 2009) we performed the experiments both on LB-agar and LB-agar plus 10% FCS plates at 10µg/µl (Fig. 5.3).

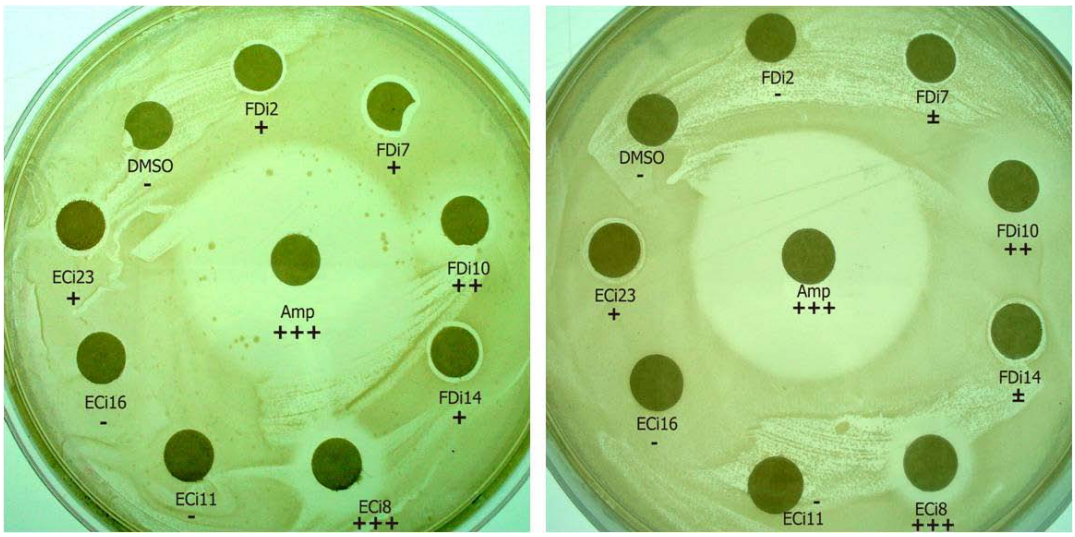


Figure 5.3 The inhibitory effects of small molecules against an *S. aureus* patient strain
Representative disc inhibition assay results at 10µg/µl (except ECi11) on **a.** LB agar plate, and **b.** LB agar plate plus 10% human serum (B) together with positive (Ampicillin (Amp), 1µg/µl) and negative, solvent (DMSO) controls are shown. Reproduced from (Shen and Liu et al., 2010).

When using standard LB-agar plates (Fig. 5.3a, and blue shaded columns in Table 5.4) among the identified FabD and elongation cycle enzyme inhibitors FDi8, 11, 12 and 15 did not show any inhibitory effect in this whole cell assay. In contrast, FDi2, 7, 10, 14 and ECi8, 23 were found to be active against all five strains tested, although their zone of inhibition was different both quantitatively and qualitatively for different strains (Fig. 5.3). ECi16 only displayed minor inhibitory effect on *E. coli* patient and MG1655 strains, while the remaining inhibitors, including ECi11, were without effect (Fig. 5.3). Of note, even though *S. aureus* Mu50 and USA300 strains are methicillin and vancomycin resistant opportunistic pathogens, respectively, in the Mu50 strain FDi10 and ECi8 and 23 displayed a clear zone of inhibition that was even stronger than the one caused by ampicillin (Fig. 5.3a). In the presence of unsaturated fatty acids (Fig. 5.3b, and pink- shaded columns in Table 5.4) FDi2 and ECi16 completely lost their inhibitory function, while FDi7, 14 and ECi23 displayed partial inhibitory activities.

	FabD (FD)								Elongation Cycle (EC)								Controls			
	i2		i7		i10		i14		i8		i11		i16		i23		Amp		DMSO	
***S. aureus* Mu50**	+	-	+	-	++	+	+	-	++	+	-	-	-	-	-	-	±	+	-	-
***S. aureus* patient**	+	-	+	±	++	++	+	±	+++	+++	-	-	-	-	+	+	+++	+++	-	-
***S. aureus* USA300**	+	-	++	-	++	++	+	±	+++	+++	-	-	-	-	+	-	+++	+++	-	-
***E. coli* MG1655**	+	-	+	-	+	-	+	±	+	+	-	-	+	ND	+	+	+++	+++	-	-
***E. coli* patient**	+	-	+	-	+	-	+	±	+	±	-	-	+	-	+	-	+++	+++	-	-
Mammalian cell toxicity																	**Triclosan**		**Cerulenin**	
MTT activity	-		±		±		±		+++		±		±		+++		±		+++	
Trypan blue excl.	-		+		+		+		+++		+		+		+++		+++		+++	

Table 5.4 Summary of bacterial disc inhibition and human fibroblast toxicity assay results

+: weak inhibition; ++: weak inhibition plus zone of partial inhibition; +++: strong inhibition; ±: partial inhibition; -: no inhibition. The blue shading indicates the inhibition effect on *S. aureus* and *E. coli* strains growing on the LB agar plate, while the pink shading displays the inhibition effects on the LB plus 10% human serum agar plate. Reproduced from (Shen and Liu et al., 2010).

In contrast, FDi10 and ECi8 retained most of their activity. These findings are consistent with the notion that the presence of unsaturated fatty acids can indeed bypass the need for FAS II enzymes (Brinster, Lamberet et al. 2009), and suggest that some of the identified inhibitors may also possess off-target growth inhibitory effects.

To further test if the identified inhibitors react with their targets' human orthologs and/ or cause toxicity in human cells through off-target effects, we employed MTT cell viability assay and trypan blue exclusion assay to examine the effect of potential inhibitors on mammalian cells. Normal human BJ foreskin fibroblast cells immortalized by hTERT (Hahn, Counter et al. 1999) were utilized for this purpose. After treating the cells with 400µg/ml of each inhibitor we found that FDi2 displayed comparable MTT activity and live cell numbers with the DMSO solvent control, indicating that it has no toxicity or inhibitory effect on cell viability (Table 5.4). In contrast, similar to the effect of cerulenin and triclosan, ECi8 and 23 completely blocked MTT

activity and killed the fibroblast cells. The other inhibitors, including FDi10, FDi14 and EC16, showed partial inhibition on mammalian cells and could potentially be further optimized for antibacterial selectivity. Of note, these experimental results are in agreement with the docking predictions for the bacterial selectivity (or lack of) of FabD inhibitors FDi2 and FDi14 (Fig. 5.1b).

5.3 DISCUSSION

Here, from a list of predicted essential enzymes shared among 13 *S. aureus* strains (Lee, Burd et al. 2009) with consistently lethal deletion phenotype in *E. coli*, we focused on the enzymes of the initiation step and elongation cycle of fatty acid biosynthesis (FAS II) as a case study. We discovered forty potential inhibitors of *E. coli* FabD, FabH, FabB/F, FabI, and FabA/Z through virtual screening of small-molecule chemical libraries against the enzymes' active sites, of which a small subset both inhibited enzyme activities *in-vitro* and bacterial cell viability *in vivo*. Thus, we provide a proof-of-principle demonstration that the predicted single essential enzymes or synthetic lethal pairs through metabolic network analysis can be potential "druggable" targets (Russ and Lampel 2005).

The promise of genome projects, functional genomics studies, and network biology in providing new molecular targets for pharmaceutical intervention in different disease states are now becoming a reality (Schadt, Lamb et al. 2005; Yang, Deignan et al. 2009). This is particularly important in the area of infectious diseases where the dual problems of multidrug resistance and the scarcity of effective therapy for a large number of diseases are well recognized. Although genome-scale gene deletions have been used for antimicrobial drug target

identification, this experimental strategy is time consuming and provides only growth condition-specific results. In some organisms that are hard to cultivate, such as *Plasmodium falciparum* or *Mycobacterium tuberculosis*, it is technically not feasible. With the large number of sequenced genomes of disease causing prokaryotes (bacteria) and eukaryotes (fungi, malaria, trypanosomes, etc.), the blueprint we demonstrate here starts from genomic information and combines the network biology-based molecular target identification with computational chemistry aided virtual screening as well as experimental validation (Figure 5.4), providing novel potential targets for antimicrobial drug discovery, thus representing an increasingly attractive alternative. This blueprint may be applicable for systematic drug target identification and subsequent lead candidate development in any cell type with high quality metabolic reconstruction.

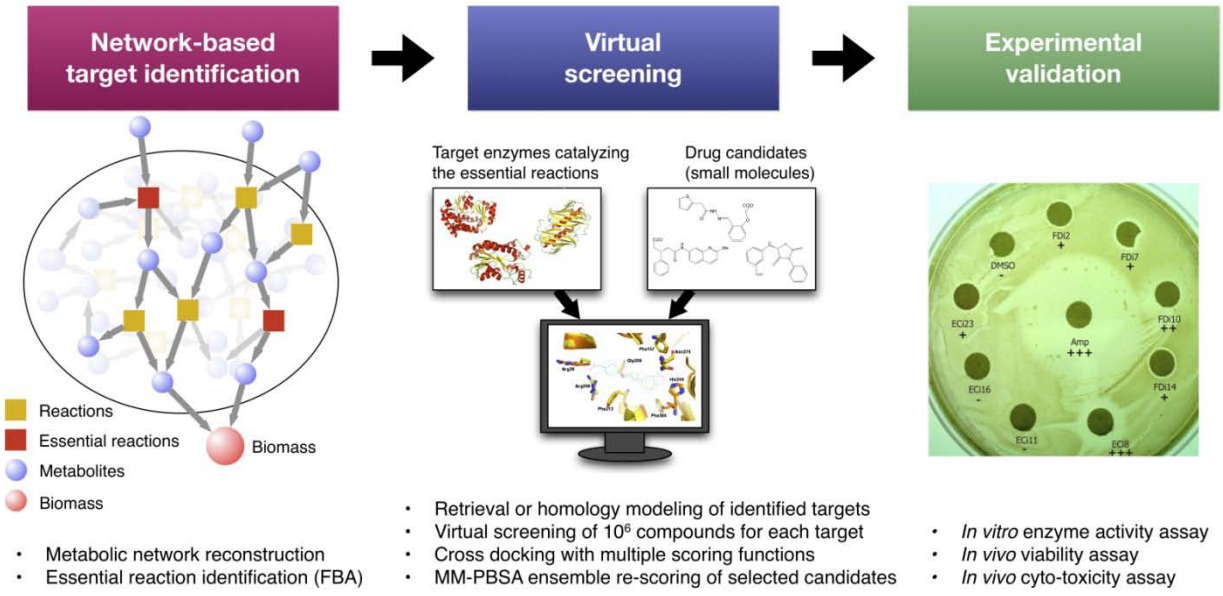


Figure 5.4 Blueprint for antimicrobial hit identification

Genome-scale metabolic network reconstruction and subsequent flux balance analysis identifies those reactions that are required for biomass formation under all growth conditions (left panel). The enzymes catalyzing the essential reactions are docked against a small molecule library to identify their potential small molecule inhibitors (middle panel). The potential inhibitors are then tested experimentally (right panel). Reproduced from (Shen and Liu et al., 2010).

5.4 MATERIALS AND METHODS

5.4.1 Identification of essential metabolic reactions in *E. coli*

Flux Balance Analysis (FBA) to identify the essential metabolic reactions of *E. coli* MG1655 using its recent metabolic network reconstruction (Feist, Henry et al. 2007) was performed as previously described (Lee, Burd et al. 2009). In brief, the reconstructed metabolic network and biomass composition of *E. coli* MG1655 was obtained from its recent reconstruction, called iAF1260 (Feist, Henry et al. 2007), that contains 1260 enzymes, 2077 reactions, and 1039 metabolites. This reconstruction introduces another compartment (periplasm) in addition to cytoplasm and extracellular space, and shows the exchange and transport reactions among these three compartments accordingly. The reconstructed metabolic network of *E. coli* MG1655 from (Feist, Henry et al. 2007) yields the stoichiometric matrix $\boldsymbol{S}$ with its element S_{ij} indicating the stoichiometric coefficient of metabolite *i* in reaction *j*. The FBA states that in the stationary state, the fluxes $\{v_j\}$ of the metabolic reactions are those which maximize $v_{biomass}$ subject to

$$\begin{aligned} &\boldsymbol{S} \cdot v = 0 \\ &v_j \geq 0 \qquad \text{for all irreversible reactions j} \\ &v_j \leq v_{\max,j} \qquad \text{for all uptake reactions j} \end{aligned} \tag{1}$$

where $v_{biomass}$ indicates the flux of the reaction that produces biomass, acting as a drain of those biomass compounds. The biomass generating reaction is given by

$$\sum_i c_i \, m_i \rightarrow biomass$$

in which m_i is each of the biomass compounds and c_i is the coefficient. In calculating the unconditionally essential enzymes, first we run the calculation using both biomass compositions (wild-type, core) obtained in *(1)*, and then consider only the enzymes that appear to be essential in both conditions.

For the FBA computations, we used ideally rich medium in which all the uptake reactions can occur without limitation on their fluxes, since we aim at identifying unconditionally essential enzymes, which we define below. The only exception was when the environment of the skin was considered. Here, due to the unique antimicrobial properties of the lipids on the skin surface (Drake, Brogden et al. 2008) the fatty acid transporters were turned off.

The FBA is basically a linear programming problem and we performed the analysis using the GNU Linear Programming Kit (GLPK). In the framework of the FBA, an enzyme is considered to be essential (Feist, Henry et al. 2007; Lee, Burd et al. 2009) in a given environment if the constraint

$$v_j = 0 \qquad (2)$$

for all reactions *j* catalyzed by an enzyme leads to zero flux of the biomass generating reaction,

$$v_{biomass} = 0$$

In the FBA scheme, the zero flux of the biomass generating reaction is interpreted as no growth of the organism. If such inactivation of an enzyme leads to no growth even in an ideally rich environment (medium) in which all the uptake reactions are allowed to occur without limitation, it would also lead to no growth in any other environment and therefore we identify

unconditionally essential - essential in all growth conditions - enzymes from the FBA (1) combined with (2) and setting all upper limits of uptake reactions to

$$v_{max,j} = \infty \qquad \text{for all uptake reactions j} \qquad (3)$$

5.4.2 Computational chemistry protocols for hit identification

Docking & virtual screening

The lead library of the ZINC database (Irwin and Shoichet 2005), containing approximately 10^6 commercially available small molecules pre-filtered for drug-like properties (Teague, Davis et al. 1999), was screened using Glide 4.5 (Friesner, Banks et al. 2004; Friesner, Murphy et al. 2006) The crystallographic structures for the enzymes involved in the different steps of the bacterial fatty acid pathway were downloaded from the Protein Database (See Table S2 for pdb codes of the crystal structures used for virtual screening, as well as the known inhibitors for each enzyme). Homology models for *E coli* FabZ were built where necessary using Prime (Zhu, Shirts et al. 2007) with the *P. falciparum* ortholog (pdb code1z6b) as template. Solvent molecules and ligands were removed and Macromodel (*Macromodel* 2007) atom types were assigned. Before the docking calculations, the protein and the ligands were preprocessed using the protein preparation and ligprep modules of the Schrodinger 2007 Suite. For the protein preparation, grid generation and docking procedures, default Glide settings were used.

For each protein, the 10^5 top-ranked structures of the Glide high throughput virtual screening were re-docked at the single precision level, saving the 10^4 top scored molecules. The resulting set was further screened using the Glide extra precision scoring algorithm (Glide XP), retaining a database of 2,000 molecules. 50 compounds of this database were selected based on

structural diversity, poses and scores, significance of the interactions in the active site, and further cross-docked using GOLD (Verdonk, Cole et al. 2003). 15 compounds available commercially were selected for each protein for a more detailed evaluation of the free energy of binding by means of ensemble MM-PBSA calculations.

Molecular dynamics (MD) simulations

To include protein flexibility into the rescoring of the putative ligands, MD simulations were performed for 15 FabD-ligand complexes and 26 FASII enzyme-ligands complexes. The SANDER.MPI module in the AMBER 9.0 (Case, Darden et al. 2006) program package was used throughout the production simulations as well as for the minimization and equilibration protocols. Long-range interactions were handled using the Particle-Mesh-Ewald (PME) summation (Essman, Perera et al. 1995). The SHAKE algorithm (Ryckaert, Ciccotti et al. 1977; *Macromodel* 2007) was employed on all atoms covalently bound to a hydrogen atom, allowing for an integration time step of 2 fs. Langevin dynamics (Pastor, Brooks et al. 1988) was used to control the temperature (300K) using a collision frequency of 1.0 ps^{-1}, with isotropic position scaling to maintain the pressure (1 atm). Initial atomic coordinates of each protein-ligand complex were taken from the pose derived from Glide XP docking. The complex was solvated in a periodic box of TIP3P (Jorgensen 1983) water molecules that extended 10Å from the protein, and was neutralized by Na^+ counterions using the LEaP module (Schafmeister, Ross et al. 1995; Case, Darden et al. 2006). The ff99 (Cieplak, Caldwell et al. 2001) and GAFF (Wang, Wolf et al. 2004) force fields of AMBER 9 (Case, Darden et al. 2006) were used to model the systems. Partial charges of the compounds were assigned using the antechamber program in AMBER (RESP) (Bayly, Cieplak et al. 1993) (Fox and Kollman 1998) (Pearlman, Case et al.

1995), after geometry optimizations at the B3LYP/6-31G* level (Becke and Yarkony 1995). For all MD simulations, 8ns trajectories were computed after minimization and equilibration protein-compound systems. The trajectories were analyzed using the PTRAJ module of AMBER.

Binding free energy calculations using Ensemble MM-PBSA

The MM-PBSA approach represents the postprocessing method to evaluate free energies of binding, combining molecular mechanical energies with continuum solvent approaches (Wang, Morin et al. 2001; Gohlke and Case 2003). It has been demonstrated to significantly improve correlation with experimental binding constants (Guimaraes and Cardozo 2008; Thompson, Humblet et al. 2008). The PBSA module of Amber 9 has been used in all cases with default parameters. Snapshots of the protein, small molecule ligand and complex were extracted every 10ps from the last 4ns of each MD trajectory. The electrostatic contribution to the solvation free energy was calculated by solving the PB equation in MM-PBSA (Onufriev, Bashford et al. 2000). The hydrophobic interaction term was derived from the solvent-accessible-surface-area-dependent term (Sitkoff, Sharp et al. 1994). The entropy term was neglected, assuming that it is similar when similar compounds bind to the same protein, and will not affect differences in binding energies.

5.4.3 In vitro enzyme assays and in vivo toxicity assays

Materials and reagents

Malonyl–CoA, CoA, BSA, NADP, NADPH, KDH (porcine heart), α-ketoglutaric acid, TPP, Cerulenin, Triclosan, MTT, and IPTG were purchased from Sigma. EDTA, TCEP, and

DMSO were obtained from Fluka. DMEM, glutamine, sodium pyruvate, phenol red, pET-151 vector and *E. coli* BL21 and Top10 competent cells were obtained from Invitrogen. The agarose gel-extraction kit and the plasmid DNA isolation kit were purchased from Qiagen. The FabD inhibitors and elongation cycle inhibitors were purchased from ChemBridge, ChemDiv, and Ryan Scientific. HisTrap FF columns were obtained from GE Healthcare. Bradford reagents and Coomassie blue were purchased from Bio-Rad.

FabD enzyme assay and the screening of potential FabD inhibitors

Purification of his-FabD and his-holo-ACP:

E. coli fabD was amplified from the genomic DNA of E. coli MG1655 strain using PCR and cloned into the pET151-TOPO vector from Invitrogen. Plasmids, pDSNdeIacpPec and pDSNdeIacpSec, containing the coding sequence for *E. coli* AcpP and AcpS, respectively, were generous gifts from Dr. Roland Lange (Actelion Pharmaceuticals Ltd., Switzerland). pET-151-fabD or pDSNdeacpPEc and pDSNdeIacpSEC were transfected or co-transfected into competent *E. coli* Top10 cells (Invitrogen). *E. coli* his-FabD and his-holo-ACP were overexpressed in *E. coli* BL21 cells induced with 0.5mM IPTG at 37 °C for 3 hours. Cells were then harvested by centrifugation and stored at -80°C. The bacterial cell pellet was resuspended in the lysis buffer and disrupted by sonication. After centrifugation, the supernatant was filtered and applied to AKTA system (GE Healthcare) using HisTrap FF column. The FabD-containing fractions were then pooled, concentrated and equilibrated in the enzyme assay buffer (50mM phosphate buffer, pH 6.8, 1mM EDTA, 1mM DTT). The protein concentration of his-FabD and his-holo-ACP was determined by Bradford assays.

Fluorometric coupled FabD enzyme assay:

The FabD enzyme assay was performed as described previously (Molnos, Gardiner et al. 2003) with minor modifications. Briefly, 46 µl 0.06nM his-FabD solution (50mM phosphate buffer, pH 6.8, 1mM EDTA, 1mM DTT and 0.1mg/ml BSA) was pre-incubated with 4 µl 10µg/ µl of the 15 different predicted FabD small molecule inhibitors or DMSO for 30min at 28 °C. 25 µl ACP/KDH mix (240µM his-ACP, 8mM α-ketoglutaric acid, 1mM NAD, 60 mU/100 µl KDH and 0.8mM TPP) and 25µl Malonyl-CoA solution (100 µM Malonyl-CoA in 50mM phosphate buffer, pH 6.8, 1mM EDTA, 1mM DTT) were added to start the reaction. NADH-dependent fluorescence was immediately measured for a minimum of 5 min using a microtiter plate reader (Beckman Coulter DTX880) equipped with an excitation filter at 340nm (bandwidth 20 nm) and an emission filter at 465nm (bandwidth 35 nm). The fluorescence signal was set to an integration time of 1s (no lag time) and three flashes per well. The experiments were run at 28°C in triplicates.

Fluorometric KDH counter screening assays:

The KDH assay was performed as follows: 3.75mU KDH was diluted in 75µl 50mM phosphate buffer, pH 6.8, 1mM EDTA, 1mM DTT, 2.67mM a-ketoglutaric acid, 0.33mM NAD, 0.267mM TPP and added to 4µl DMSO (controls) or 10µg/µl inhibitors dissolved in DMSO. A 60µM CoA solution was prepared in 50mM phosphate buffer, pH 6.8, 1mM EDTA, 1mM DTT. All solutions and the microplate reader were pre-warmed at 28°C; 21 µl CoA solution was added to start the reactions that were measured immediately over 5 min in triplicates.

Elongation cycle assay and the screening of potential inhibitors

We utilized an established elongation assay for type II fatty acid synthesis (Kodali, Galgoci et al. 2006) with slight modifications. Briefly, *E. coli* MG1655 was grown to a stationary phase in LB medium. The cultures were centrifuged and the pellets were washed and lysed in ice-cold buffer. The supernatant were then precipitated using ammonium sulfate. The 45%-80% ammonium sulfate saturated protein fraction which contained all the FASII enzymes was dialyzed. Subsequently, 2-16μg cell lysate was first incubated with different concentrations of inhibitors at 37°C for 20 min in 50μl buffer containing 5mM EDTA, 1mM NADPH, 1mM NADH, 5mM β-mercaptoethanol, 150μM DTT, 16μM ACP (pretreated with 3mM ice-cold DTT for 20 min), 40μM acetyl-CoA, 4% DMSO, 100mM sodium phosphate (pH 7.0). Then, 120μM 10μl malonyl-CoA was added into the 50μl solution to initiate the reaction. After 30 minutes reaction at 37°C, 60μl of each sample (plus 15μl 5X native sample buffer) were applied to a 16% polyacrylamide gel containing 0.5M Urea (to separate intermediate & final products) or 4M Urea (to separate different chain lengths of acyl-ACP products). After electrophoresis, the gels were stained with Coomassie Blue for visualization. Reactions with no inhibitor (only DMSO) were used as negative control. Known FASII elongation cycle inhibitors triclosan (167μg/ml) and thiolactomycin (167μg/ml) were used as positive controls.

Bacterial disc inhibition assay

The disc inhibition assay was performed as described previously (Woods and Washington 1995). Briefly, freshly grown *E. coli* (MG1655 strain and patient strain) and *S. aureus* (Mu50 strain, USA300 strain and patient strain) cultures (OD_{600}=0.3) were evenly inoculated onto the surface of LB or LB plus 10% human serum agar plates. Sterile filter discs soaked with 10μl of 10mg/ml of the different inhibitors were then applied to the surface of the

inoculated plates and incubated for 16-20h at 37 °C. 1mg/ml Ampicillin and 10µl DMSO were utilized as positive and negative controls, respectively.

MTT cell viability assay

Cell viability was determined by 3-[4,5-dimethyl-triazolyl-2]2,5-diphenyl tetrazolium bromide (MTT) assay. The human BJ fibroblast cells expressing hTERT (CL1 cells) (Hahn, Counter et al. 1999) were seeded in 24-well plates at a density designed to reach 90% confluency at the time of assay. The next day, the culture medium was replenished with fresh medium with 400µg/ml of the potential inhibitors. The same amount of DMSO was used as control. After 24 h of inhibitor treatment, MTT was added to a final concentration of 0.5 mg/ml, and the plate was incubated for 3 h at 37 °C. Cells having functional mitochondrial succinate dehydrogenase can convert MTT to formazan that generates a blue color when dissolved in dimethyl sulfoxide (DMSO). The absorbance was read at 540 nm using a plate reader (Beckman DTX880).

Trypan blue exclusion assay

CL1 cells were seeded at a density of 1×10^5 cells in 12-well plates and cultured in DMEM (25mM glucose, 4mM glutamine) containing 10% FBS at 37°C in an incubator with 5% CO_2 and 95% air. The next day, the culture medium was replenished with fresh medium with 400µg/ml of the inhibitors. The same amount of DMSO was used as control. After 24 h of inhibitor treatment, the cells were harvested by releasing them with 0.05% trypsin, 0.53mM EDTA, stained with trypan blue, and counted. Cell counting data were determined by blinded counting of both live (unstained) and dead (blue) cells on a standard grid-patterned hemocytometer. The mean number of cells and the standard deviation of the mean were

determined and data was then compared using a standard Student's *t* test to determine statistically significant differences between cell populations.

5.5 ACKNOWLEDGEMENT

We thank Drs. I. Bahar, D. Tobi, B. Chen and B. Shoichet for help with the structural analyses, and Dr. R. Lange (Actelion Pharmaceuticals Ltd., Switzerland) for providing plasmids, pDSNdeIacpPec and pDSNdeIacpSec. This research was supported by NIAID U01-0700499 to ZNO, ALB, VK and OW, and by a travel grant for YS from the Chemistry-Biochemistry-Biology Interface (CBBI) Program at the University of Notre Dame, supported by training grant NIGMS T32-075762.

6.0 IDENTIFICATION OF NOVEL BACTERIAL HISTIDINE BIOSYNTHESIS INHIBITORS USING DOCKING, ENSEMBLE RESCORING AND WHOLE CELL ASSAYS

6.1 INTRODUCTION

Although virtual screening has become an established tool for computer aided molecular design and frequently reproduces experimentally observed binding poses, there is usually no good correlation between docking scores and experimentally observed binding constants. Therefore, a significant number of compounds from virtual screens are usually selected for experimental confirmation by enzyme assays early in the hit discovery process. This requires significant effort in the acquisition and screening of the compounds and typically results in varying enrichment factors that depend on the scoring function and the enzyme studied. It would therefore be desirable to further refine the scoring to increase enrichment and possibly bypass the biochemical assay in favor of whole cell assays. To further refine the proposed blueprint in Chapter 5 and improve the accuracy of the virtual screening to the point where a whole cell assay is directly utilized to replace the often problematic enzyme assay, we select another one of the families of targets identified in these studies -- the histidine biosynthesis pathway, an unbranched pathway consisting of 10 enzymatic reactions with no routes to bypass any of the enzymes (Fig. 6.1) (Alifano, Fani et al. 1996).

Here, we report the results of a study of inhibitors of the histidine biosynthesis pathway, where ensemble rescoring was used to select compounds that were then directly tested in whole-cell assays. To demonstrate this novel strategy to identify potential inhibitors of the histidine biosynthesis, we chose three enzymes from the pathway as targets for antibiotic hit identification based on the availability of crystal structures and established biochemical assays: Phosphoribosyl-AMP Cyclohydrolase (HisI) (D'Ordine, Klem et al. 1999; Sivaraman, Myers et al. 2005), Imidazoleglycerol Phosphate Dehydratase (IGPD) (Glynn, Baker et al. 2005), and Histidinol Phosphate Aminotransferase (HisC) (Ames and Horecker 1956; Haruyama, Nakai et al. 2001; Sivaraman, Li et al. 2001; Fernandez, Vega et al. 2004; Marienhagen, Kennerknecht et al. 2005). The efficacy of the identified hits will then be tested in whole-cell assays. This suggested that the computational predictions can be sufficiently accurate to be tested directly in *in vivo* disk inhibition assays, which would accelerate the process.

6.2 RESULTS

6.2.1 The histidine biosynthesis pathway in *S. aureus*

The first step in the biosynthesis of histidine is catalyzed by the enzyme ATP phospho-ribosyltransferase (HisG) and results in the condensation of ATP and 5-phosphoribosyl 1-pyrophosphate to form *N*'-5'-phosphoribosyl-ATP. The triphosphate group of the product is then hydrolyzed by the enzyme phophoribosyl-ATP pyrophosphohydrolase (HisE) to form *N*'-5'-phosphoribosyl-AMP, and subsequently the

purine ring is hydrolyzed by phosphoribosyl-AMP cyclohydrolase (HisI) to form *N*'-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4 carboxamide-ribonucleotide. In the fourth step of the pathway, the product is converted to *N*'-[(5'-phosphoribulosyl)-formimino]-5-aminoimidazole-4 carboxamide-ribonucleotide *via* an Amadori rearrangement. This product is then cleaved into imidazole glycerol phosphate (IGP) and 5-aminoimidazole-4-carboxamide 1-beta-D ribofuranosyl 5'-monophosphate (AICAR). The first of these is used in the further synthesis of histidine, while the latter is used in the *de novo* synthesis of purine. In the sixth step of the pathway IGP is dehydrated by imidazole glycerole phosphate dehydratase (IGPD). This gives the product imidazole acetol phosphate, which in the seventh step of the biosynthesis is transaminated by the enzyme histidinol phosphate aminotransferase (HisC) to give L-histidinol phosphate. This product is then converted into L-histidinol, which is oxidized to L-histidine in two steps *via* the L-histidinal intermediate by a single enzyme (HisD).

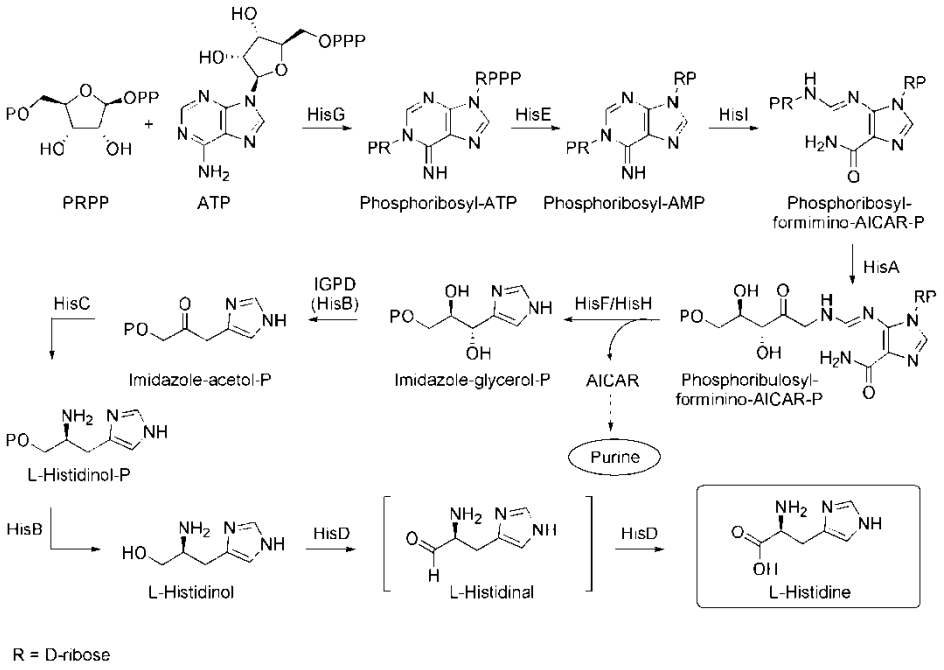


Figure 6.1 Histidine biosynthesis pathway. Reproduced with permission from (Henriksen, Liu et al., 2010).

Based on our analyses of the metabolic network of thirteen *S. aureus* strains by flux balance analysis,(Lee, Burd et al. 2009) inhibition of a single enzyme of the pathway is expected to stop the production of histidine, an essential biomass component. Evolutionarily, the pathway is highly conserved among different bacterial species, which makes the pathway a potential target for the treatment of bacterial infections other than caused by *S. aureus*. The histidine pathway has been extensively studied in the past, and these studies have contributed to the development of the operon theory of gene clustering and function.(Alifano, Fani et al. 1996) More recently, the histidine biosynthesis in plants has received a lot of attention due to its vital role for plant growth.(Stepansky and Leustek 2006) Particularly the sixth enzyme of the pathway, IGPD, has been used as a target for the design of herbicides (Mori, Iwasaki et al. 1995; Cox, Hawkes et al. 1997; Ohta, Mori et al. 1997; Gohda, Kimura et al. 1998; Schweitzer, Loida et al. 1999; Gohda, Mori et al. 2000; Schweitzer, Loida et al. 2002).

6.2.2 Identification of inhibitor scaffolds through computational chemistry

Phosphoribosyl-AMP Cyclohydrolase (HisI) (EC 3.5.4.19) catalyzes the third step of the histidine biosynthetic pathway, the hydrolysis of phosphoribosyl-AMP (Fig. 6.1). The crystal structure of HisI for *S. aureus* has not yet been determined, but the structure of HisI from *Methanobacterium thermoautotrophicum* (PDB ID: 1zps), which has a 51% homology to HisI from *S. aureus,* has been solved to a resolution of 1.70 Å. We have built a homology model of the *S. aureus* MRSA252 enzyme and compared it to the structure of the *M.*

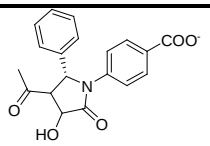
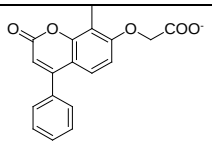
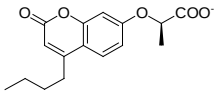
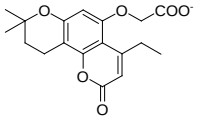
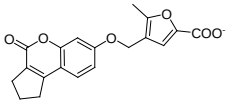
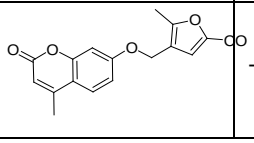
thermoautotrophicum enzyme to validate its use in the docking experiments. HisI is a homodimeric metalloenzyme dependent on both Zn^{2+} and Mg^{2+}. Two pairs of metal ions are bound on the interface between the two subunits, and make up two active sites, each containing one zinc and one magnesium ion, respectively. EDTA inhibits the enzyme by chelation of Mg^{2+}, and exogenous Zn^{2+} also reduces enzyme function, but no small molecule inhibitors of this enzyme have previously been reported.

Application of the virtual screening procedure to the lead library from ZINC yielded 12 small molecules that have good poses in both the docking and the subsequent MD simulations. Six of these molecules, shown in Table 6.1 together with their G-scores from Glide XP as well as predicted binding energies from MM-PBSA and MM-GBSA calculations, were selected based on their structural diversity. It should be noted that the absolute binding energies calculated by either method are not meaningful (Thompson, Humblet et al. 2008). Unfortunately, no known inhibitors are available for referencing the values and that the scores can only be used for a relative ranking of the compounds.

Figure 6.2 shows two representative examples of the binding modes observed after 8 ns MD simulation (for representations of other poses, see the Supporting Information). The HisI active site is solvent exposed and contains a zinc and a magnesium ion, which explains the high abundance of carboxylate derivatives among the docking hits. Inhibitor HisI2 is a typical case for the bidentate coordination of the zinc ion by the carboxylate group. This zinc coordination is always preferred over the coordination of the magnesium ion (only one of the selected docking poses, HisI10, showed the carboxylate group coordinating to both metal ions). The magnesium ion is surrounded by three aspartate residues from one side and a hydrophobic pocket consisting of the residues Ile25, Ile25', Val27, Val27', Leu37, Leu37',

Val39, Val39', Tyr41 and Tyr41' from both subunits (shown as four blue colored β-sheets in Figure 6.2). Several docking poses place a hydrophobic group pointing into this pocket (HisI2, 3, 4, 6, and 8), and most of these allow a cation-π interaction with the magnesium ion and an aromatic moiety.

Table 6.1: Selected Inhibitors for HisI

ID	Structure	G-score kcal/mol	MM-PBSA kcal/mol	MM-GBSA kcal/mol
HisI1		-15.23	-55.73	-22.50
HisI2		-14.44	-60.15	-29.58
HisI4		-14.08	-60.16	-25.46
HisI8		-13.80	-22.33	-22.28
HisI9		-13.69	-35.48	-30.28
HisI11		-13.42	-45.83	-19.40

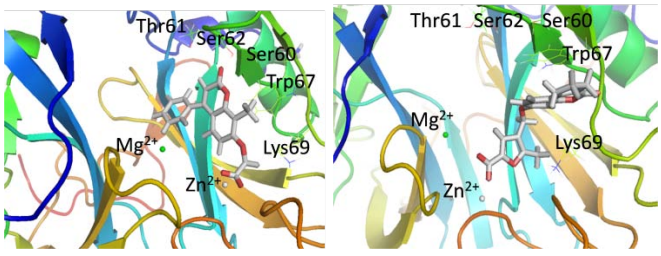


Figure 6.2 Active site of homology model of *S. aureus* with HisI2 (left) and HisI9 (right) in the active site after 8ns MD. Reproduced with permission from (Henriksen, Liu et al., 2010).

Imidazoleglycerol Phosphate Dehydratase IGPD (EC 4.2.1.19) catalyzes the dehydration of imidazoleglycerol phosphate to imidazoleacetol phosphate, which is the sixth step of histidine biosynthesis (Fig. 6.1). A crystal structure of the *S. aureus* N315 IGPD with a resolution of 2.01 Å (PDB ID: 2ae8) is available. However, many moieties close to the active site are not resolved, making this structure unsuitable for virtual screening approaches. We have therefore chosen the well characterized *Aridopsis thaliana* IGPD (PDB ID: 2f1d, 42% homology) to build a homology model of the *S. aureus* MRSA252 IGPD. IGPD is a manganese dependent metalloenzyme consisting of 24 identical subunits. Each active site consists of three subunits and two manganese ions. Each manganese ion is coordinated by a glutamate and three histidine residues from two different subunits ($Mn^{2+}{}_1$: His47A, His169A, Glu173A, and His74B; $Mn^{2+}{}_2$: His170A, His73B, Glu77B, and His145B). The IGPD enzyme has been used as target for the development of herbicides, and several inhibitors of this enzyme are known from these efforts (Mori, Iwasaki et al. 1995; Cox, Hawkes et al. 1997; Ohta, Mori et al. 1997; Gohda, Kimura et al. 1998; Schweitzer, Loida et al. 1999; Gohda, Mori et al. 2000; Schweitzer, Loida et al. 2002). The known inhibitors resemble the natural substrate imidazole-glycerol phosphate (IGP). They typically contain a phosphonate group instead of the phosphate group present in IGP, a triazole group instead of imidazole, and a number of hydroxy groups.

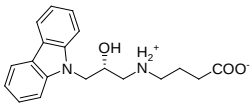
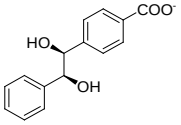
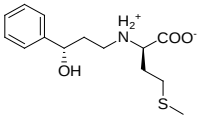
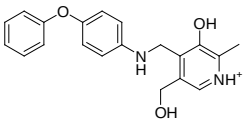
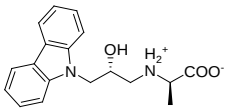
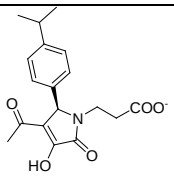
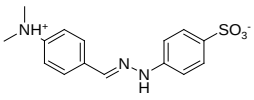
The selected potential IGPD inhibitors and their G-scores are listed in Table 6.2. It is evident that the MM-PBSA and MM-GBSA scores obtained for all inhibitors as well as for the known inhibitors (data not shown) are relatively unfavorable. This might be a

consequence of an underestimation of binding to the manganese ions, possibly due to the Mn^{2+} parameters employed for the MM-PBSA/GBSA calculations (see the Supporting Information for further details). Structurally, there is a high abundance of carboxylate derivatives among the best docking hits. In the majority of these, the carboxylate group is placed between the two manganese ions coordinating both of them. The key interactions are exemplified by the structures of IGPD13, 14, 17, and 18 in the IGPD active site after 4ns of MD simulation shown in Figure 6.3.

Most of the potential inhibitors interact with Asp69, either forming a salt bridge through an ammonium group like in the case of IGPD1, 2, and 5 or by hydrogen bond donating, either through a hydroxy group (IGPD6, 8, 10, 11, 12, 14, and 16) or a NH functionality (IGPD17 and 18), as shown in Figure 3 b-d. In accordance with the negatively charged substrate, the active site is highly positively charged with the presence of three arginines (Arg8 and Arg295, with an additional arginine residue Arg268 in the periphery of the active site) and Lys534. As a result, all selected potential inhibitors posses a negative charge, either as a carboxylate or a sulfonate. The sole exception, IGPD13, has an acidic phenol hydroxyl group that is likely deprotonated upon binding to the manganese, as shown in Figure 6.3a. Five of the selected poses interact with Lys534 (IGPD3, 5, 9, 10, and 18), and four interact with Glu12 (IGPD2, 6, 13, and 18). The rest of the active site pocket is filled out by aromatic groups stacking against the α-helix forming one wall of the active site (IGPD13, 17 and 18) or the β-sheet forming the other (IGPD14). As mentioned before, all known inhibitors of the IGPD enzyme are small, polar compounds. We have therefore chosen fewer hits containing large, hydrophobic groups and more relatively small, polar compounds for further evaluation than is actually representative of the abundance with

which they appear among the docking hits (e.g., IGPD2, 3, 5, and 6). We have also included two sulfonate derivatives (IGPD7 and 17) in the compound set to be evaluated experimentally, since a sulfonate group is a better bioisostere for the phosphonate group than the carboxylate group.

Table 6.2: Selected Inhibitors of IGPD

ID	Structure	G-score kcal/mol	MM-PBSA kcal/mol	MM-GBSA kcal/mol
IGPD1		-8.61	-17.22	-36.66
IGPD6		-8.05	12.28	-1.75
IGPD10		-7.78	-5.53	-24.67
IGPD13		-7.70	-1.30	-20.34
IGPD14		-7.68	-0.30	-20.53
IGPD16		-7.62	10.30	-13.25
IGPD17		-7.32	5.21	-12.98

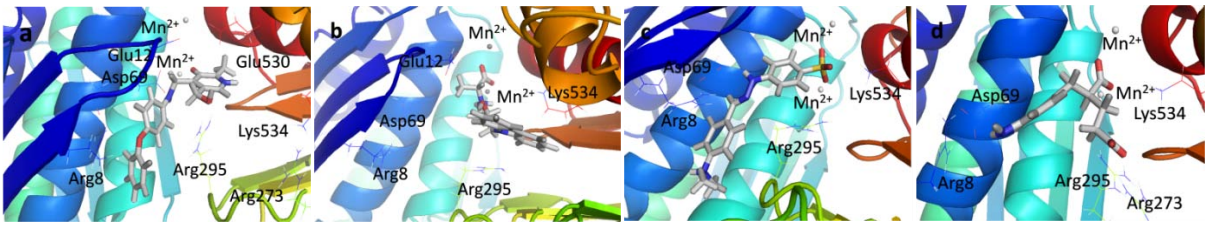
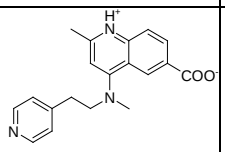
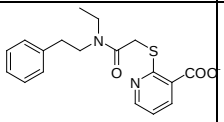
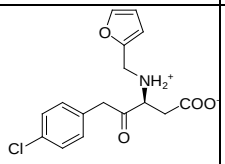
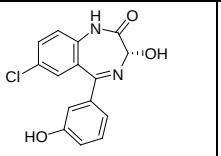
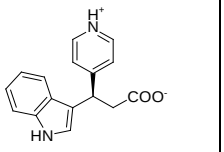
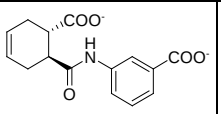


Figure 6.3 Structures of potential inhibitors a) IGPD13, b) IGPD14, c) IGPD17, and d) IGPD18 bound to IGPD active site from 4 ns MD simulations. Reproduced with permission from (Henriksen, Liu et al., 2010).

Histidinol Phosphate Aminotransferase (HisC). The seventh step of the histidine biosynthesis in eubacteria is the transamination between imidazoleacetol phosphate and glutamate to form L-histidinol phosphate (LHP) and α-ketoglutarate (Fig. 6.1), which is catalyzed by the enzyme HisC (EC 2.6.1.9). HisC uses the cofactor pyridoxal-5'-phosphate (PLP) to transfer the amino group from glutamate, forming pyridoxamine-5'-phosphate (PMP) that is subsequently transferred to the substrate. Except for the metal ions Cu^{2+} and Co^{2+} and hydroxyl amine, iodoacetate and semicarbazide, no HisC inhibitors are currently known. In our docking calculations, we have used the crystal structure of HisC complexed with PMP from *E. coli* (PDB ID: 1fg7, 29% homology), which has been solved to a resolution of 1.5 Å .(Sivaraman, Li et al. 2001) HisC from *S. aureus* MRSA252 has a high sequence similarity with its *E. coli* ortholog, making the latter a suitable model. One important difference is that the *E. coli* residue Tyr110 is substituted to a Phe in the *S. aureus* ortholog. Although the active enzyme functions as a homodimer, all relevant residues are contained in the monomer that was used in the docking calculations. In the crystal structure, the phosphate group of PMP forms a salt bridge with Arg222 and hydrogen bonds to Thr211, Ser213, and to the backbone amide groups of Ala84 and Asp85. The amino group forms a hydrogen bond to Lys214 and the hydroxy group hydrogen bonds to Tyr187 and

Asn157. Asp184 forms a salt bridge with the protonated pyridinium nitrogen and Tyr110 interacts with the pyridinium ring via π-stacking. The substrate, LHP, is known to form hydrogen bonds to Tyr20 and Asn157 and salt bridges to Arg322 and Arg335 via its phosphate group. Furthermore, the LHP imidazole ring forms hydrogen bonds to the residues Tyr20, Asp85, and Tyr110 in the *E. coli* HisC.[22] Based on the sequence alignment MRSA252 HisC cannot form the latter hydrogen bond.

Table 6.3: Selected Inhibitors of HisC

ID	Structure	G-score kcal/mol	MM-PBSA kcal/mol	MM-GBSA kcal/mol
HisC5		-15.02	-40.92	-49.55
HisC9		-14.50	-38.74	-43.95
HisC11		-14.25	-28.80	-35.94
HisC14		-13.84	-12.95	-35.13
HisC16		-13.70	-29.43	-37.33
HisC19		-13.53	-13.80	-26.48

The docking hits for the HisC enzyme are shown in Table 6.3 and are again negatively charged or polar compounds for the same reasons as discussed for IGPD. The

affinity of carboxylate derivatives to the HisC enzyme can be ascribed to the phosphate binding site of the PMP cofactor (Arg222, Thr209, and Ser211) and to the residues Asn157 and Arg335 which are part of the phosphate binding site of the substrate LHP. In all the docking hits that are carboxylate derivatives, the carboxylate functionality binds to either of the two phosphate binding sites.

The two main binding modes where the carboxylate group occupies the PMP phosphate binding site and its pyridine ring interacting with the LHP binding site and where the carboxylate group binds to the LHP phosphate binding site are exemplified by the calculated docking poses of HisC5 and HisC9 and are shown in Figure 6.4a and b, respectively. Other structures that do not contain a carboxylate functionality have other polar groups that follow the same motifs and bind to at least one of the phosphate binding sites, as for example the azepine moiety of HisC14 shown in Figure 6.4c. Other key interactions include π-stacking with Tyr110 and hydrogen bonding or salt bridges with Lys214 and Asp184, respectively. Overall, the calculated binding energies, shown in Table 12, are strongly favorable, but the lack of known inhibitors prevents the referencing to experimental data and allows only a relative ranking of the compounds.

Several hits with good G-scores contain two carboxylate groups, which would seem favorable, since both phosphate binding sites then can interact with a carboxylate group. This is demonstrated in Figure 6.4d, which shows the docking pose of HisC19 where both carboxylate groups interact with the two phosphate binding sites of PMP and LHP. However, since the active site is buried deep within the enzyme, the cost of desolvation of these bis-anions is likely to decrease the binding of very polar ligands, as is indicated by the much lower binding energies calculated by MM-PBSA and MM-GBSA. We therefore

selected only HisC19 as one example of this class of compounds for further experimental evaluation. Instead, we have chosen several structures where other polar groups are able to interact with the second phosphate binding site, e.g. the pyridine group of HisC5 and the amide group of HisC9 (Figure 6.4a, b, respectively).

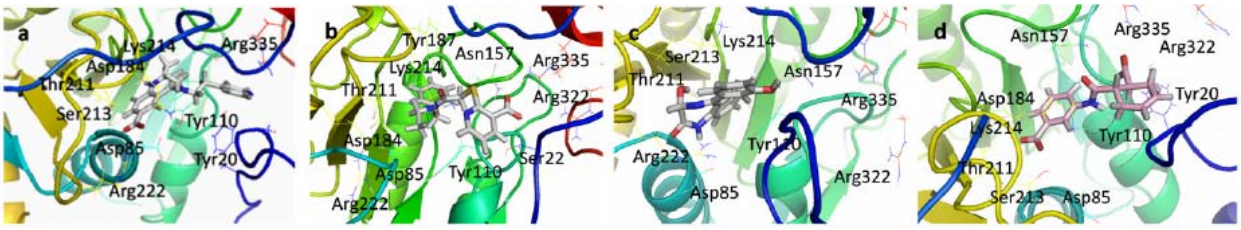


Figure 6.4 Structures of potential inhibitors
a) HicC5, b) HisC9, c) HisC14, and d) HisC19, bound to HisC active site from 8 ns MD simulations. Reproduced with permission from (Henriksen, Liu et al., 2010).

6.2.3 Biological validation of potential inhibitors of the histidine pathway

A recent study of inhibitors of the FASII biosynthetic pathway (Shen, Liu et al. 2010) demonstrated that 42% of the inhibitors predicted to be good binders by the MM-PBSA ensemble rescoring method were subsequently confirmed experimentally in enzyme- and whole cell assays. These results encouraged us to use a disk inhibition assay as a primary screen for the computationally predicted compounds. Although this approach can potentially yield both false positive results due to inhibition of targets other than the computational predicted ones and false negative results due to limited transport across the cell membrane, the simplicity of the assay and the fact that the results are more relevant for the ultimate goal of bacterial growth inhibition made this approach attractive.

We therefore selected several representative structures with good calculated MM-PBSA binding energies, structural features representative of the diversity of the larger dataset, and reasonable interactions to the active site of the respective enzymes to be tested

in disk inhibition assays. The goal of the selection was to identify the maximum number of active compounds with the minimum size of the total dataset. Of the potential HisI inhibitors, we selected HisI1, 2, 4, 8, 9, and 11 to be biologically tested. These compounds have been selected due to good calculated binding energies, and because they represent both of the two different binding modes discussed above. For IGPD, we have included the potential inhibitors IGPD1, 6, 10, 13, 14, 16, and 17. For HisC, we chose the potential inhibitors HisC9, 11, 14, 16, and 19 for biological evaluation. These 18 compounds were obtained commercially and tested in a standard bacterial growth assay(Woods and Washington 1995) in two MRSA strains (Mu50 and USA300) as well as a random patient isolate from the University of Pittsburgh Medical Center. As example of a gram-negative bacterium, we also tested the *E. coli* MG1655 strain and an *E. coli* random patient isolate. 1 mg/ml Ampicillin and 10 μl DMSO were used as positive and negative controls, respectively. Figure 6.5 shows some representative plates for t *E. coli* and *S. aureus* strains and the results for active compounds are summarized in Table 6.4.

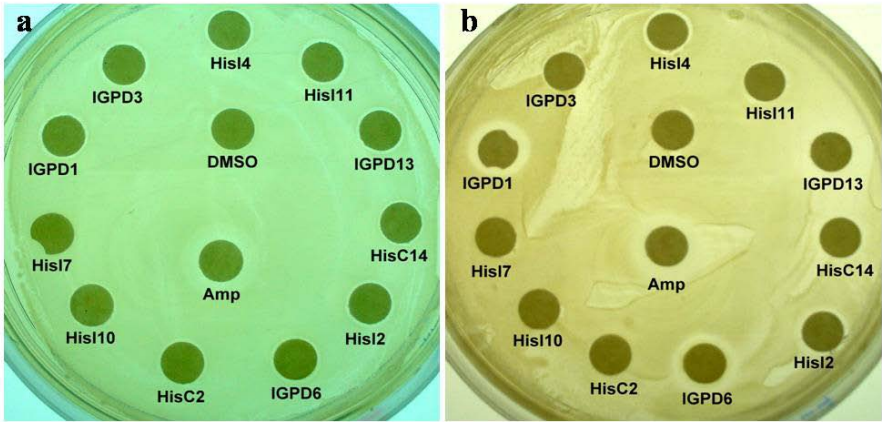


Figure 6.5 **Bacterial disc inhibition assays and human fibroblast toxicity assay.**

The inhibitory effect of selected small molecules against **a./** *E. coli* MG1655 strain, **b./** *S. aureus* Mu50 strain. Reproduced with permission from (Henriksen, Liu et al., 2010).

Table 6.4: Summary of results of growth inhibition and toxicity assays for 18 selected small molecules

ID	*S. aureus* patient strain	*S. aureus* Mu50	*S. aureus* USA300	*E. coli* patient strain	*E. coli* 1655	MTT activity	Trypan blue
HisI1	-	±/-	-	-	+	N.D.	N.D.
HisI2	+	+	+	+	+	-	-
HisI4	+	+	-	±	±	±	±
HisI8	-	+	±	-	+	N.D.	N.D.
HisI9	+	+/±	-	-	-	N.D.	N.D.
HisI11	±	±	-	±	+	±	-
IGPD1	+++	+	+++	±	±	±	-
IGPD6	+	+	+	±	+	±	+
IGPD10	-	-	-	-	-	N.D.	N.D.
IGPD13	++	++	++	+/-	+	±	+
IGPD14	-	-	-	-	-	N.D.	N.D.
IGPD16	-	+	-	+	+	N.D.	N.D.
IGPD17	-	-	-	-	±	N.D.	N.D.
HisC9	±	±	±	±	+	N.D.	N.D.
HisC11	-	-	-	-	-	N.D.	N.D.
HisC14	±	+	+/±	+/-	+	±	±
HisC16	+	+/±	-	+	±	N.D.	N.D.
HisC19	-	-	-	-	±	N.D.	N.D.
Amp	+++	+	+	+++	+	±	-
DMSO	-	-	-	-	-	-	-

Summary of bacterial disk inhibition and human fibroblast toxicity assay results. +: weak inhibition; ++: weak inhibition plus zone of partial inhibition; +++: strong inhibition; ±: partial inhibition; -: no inhibition; N.D. not determined. Reproduced with permission from (Henriksen, Liu et al., 2010).

Out of 18 tested compounds 10 compounds have different levels of activity against the *S. aureus* patient strain, 13 against *S. aureus* Mu50, and 7 against *S. aureus* USA300 as shown in Table 6.4. Especially encouraging is the fact that several of the compounds show significant activity towards the drug resistant strains of *S. aureus*, Mu50 and USA300 that show only weak susceptibility towards the well-established antibiotic ampicillin**.**

It is interesting to note that of the six tested HisI ligands, HisI1 and HisI9 were predicted to bind in a different orientation than the rest. These two ligands show the least activity of the potential HisI inhibitors in the disc inhibition assay, which suggests that their

mode of binding is not favorable. The compound HisI8 seems to be the third least active compound, which is in accordance with a lower calculated binding energy for this ligand than for the rest.

Out of the seven tested IGPD ligands, three show activity. This is the small polar compound IGPD6, the compound with a pyridinol group instead of a carboxylate group (IGPD13) and the carbazolyl derivative IGPD1. It is noteworthy that while IGPD1 shows strong inhibition of two *S. aureus* strains, the very similar compound IGPD14 shows no inhibitory effect at all. Glide does rank IGPD1 higher than IGPD14, but both poses are given high G-scores. However, MD and MM-PBSA/GBSA calculations are able to identify IGPD14 as a false positive. Two especially promising compounds are IGPD1 and IGPD13. The first shows strong inhibition against the *S. aureus* patient strain and the USA300 strain and it shows weak inhibition against the Mu50 strain. The second shows a clear zone of growth inhibition plus partial inhibition of the *S. aureus* strains Mu50 and USA300 respectively, and a weak inhibition of growth in the patient strain.

Three out of the five tested potential HisC ligands show an inhibitory effect. These three ligands all interact with one of the phosphate binding sites and with Tyr110, indicating the importance of these interactions for inhibitory activity on this enzyme. The positive results for HisC14 indicates that other groups than carboxylate can interact with the phosphate binding sites of this enzyme. The lack of activity of the dicarboxylate HisC19 might be a consequence of too high a desolvation cost for binding of this very polar compound to the deeply buried HisC active site pocket. It might also be due to HisC19 being too polar to effectively penetrate the cell wall.

Activity against *E. coli* is generally much lower, which is possibly a result of both the selection of the compounds towards the *S. aureus* orthologs of the three enzymes studied as well as the different properties of the cell wall in gram negative bacteria, which could make it difficult for the mostly negatively charged inhibitors to pass the outer membrane. Nevertheless, 14 compounds were found to be weakly active or give partial inhibition for the *E. coli* 1655 strain.

To check the computational prediction that the small molecules exert their effect through binding the the enzymes in the histidine biosynthesis pathway, we performed a complementation assay where the bacteria were grown on minimal medium agar plates, where the effect of the potential inhibitors is very similar to the LB plates. Upon addition to 0.2mM L-histidine, five of ten inhibitors selected showed complete complementation of growth of the *S. aureus*, while the other five showed partial complementation in at least one of the strains tested. These results are consistent with the hypothesis that the small molecules indeed inhibit whole cell growth through interference with the histidine biosynthesis.

Finally, we tested the toxicity of seven of the more active potential inhibitors in mammalian cells for off-target effects using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) cell viability and trypan blue exclusion assays. For this purpose, we used normal BJ foreskin fibroblast cells immortalized by hTERT, the catalytic subunit of human telomerase.(Hahn, Counter et al. 1999) The results of these studies are summarized in Table 4. Again, IGPD1 is the most promising compound for further development, since it has slightly lower human cell toxicity than the known antibiotic ampicillin. In comparison, the second best inhibitor, IGPD13, is significantly more cytotoxic (according to the trypan blue exclusion assay). Overall, the human cell toxicity of the selected compounds in these

assays is fairly limited; indicating that off-target activity of the selected compounds is relatively low.

6.3 CONCLUSIONS

The development of novel compounds with antibiotic activity is a pressing need in view of the rapidly expanding resistance of pathogens towards known antibiotics. In the ongoing race between drug development and new resistance, not only new compounds but more importantly, new methods for the discovery of new hits are needed. In the present study, the application of virtual screening and MM-PBSA ensemble rescoring to three enzymes in the histidine biosynthetic pathway, which was predicted to be essential for growth of *S. aureus* by systems-level approaches (Lee, Burd et al. 2009), yielded 49 potential inhibitors. Based on structural diversity and physicochemical properties, 18 compounds were selected for experimental tests in a whole cell growth inhibition assay. We found 13 and 14 compounds to be active in different strains of *S. aureus* and *E. coli*, respectively. Although the observed growth inhibitions are in many cases weak or only partial, two compounds (IGPD1 and IGPD13) show very promising activity comparable to the effects of the well-known antibiotic ampicillin.

Several conclusions can be drawn from these studies. First, the scalable and fast computer-based methods of flux balance analysis and virtual screening used for target and hit identification, respectively, is an effective strategy for the early stages of antibiotic drug discovery. Secondly, refinement of the docking results by ensemble rescoring together with judicious compound selection leads at least in the examples present here and in our earlier

studies (Shen, Liu et al. 2010) to excellent hit rates and justifies the immediate use of cell-based assays, which in many respects have a higher information content for drug discovery. Thirdly, at least two of the compounds disclosed here merit further development as inhibitors of the histidine biosynthesis pathway that is unconditionally essential for growth in *S. aureus* even though the possibility of transport of histidine or one of the relevant biosynthetic intermediates from external sources has to be, in analogy to studies of the FASII pathway,(Brinster, Lamberet et al. 2009) investigated in future studies.

6.4 MATERIALS AND METHODS

Computational methods Homology models of the *S. aureus* enzymes were built in Prime(2007) using comparative modeling using the template structures discussed in the text. The docking experiments were performed in Glide (Friesner, Banks et al. 2004; Friesner, Murphy et al. 2006), and using the "Lead" subset of the ZINC database (Irwin and Shoichet 2005) of commercially available compounds. This subset was obtained from the complete dataset by applying filters(Teague, Davis et al. 1999) to have good drug potential, resulting in ~10^6 small molecules docked to the enzyme of interest using Glide's high throughput mode. The highest scoring 100,000 hits were saved and docked to the enzyme again, this time using Glide's standard precision mode. The highest scoring 10,000 hits were then saved, and docked to the enzyme using the extra precision mode. The highest scoring 2,000 hits were saved, and by manual inspection we selected a small number of potential inhibitors representative of the chemical space covered by the best scored docking hits for ensemble rescoring. In this procedure, side chain flexibility is introduced through 8 ns MD

simulations, allowing the active site residues to move, and to test how these movements affect the binding mode of the ligand. Desolvation and hydrophobic effects are included through MM-PBSA/GBSA calculations of the binding energy on snapshots extracted from the last 4 ns of the MD run. These scores were then used to select compounds for experimental testing.

Experimental protocols For the experiments two *E. coli* (MG1655 and a random patient isolate) and three *S. aureus* (MU50, USA300 and a random patient isolate) and one *P. ariginosa (PA01)* strains were used. The propagation of bacteria and their experimental manipulation were performed as previously described (Shen, Liu et al. 2010). In brief, sterile filter discs soaked with 10µl of 10mg/ml of the different inhibitors were applied to the surface of the evenly-inoculated LB plates, minimal medium agar plates, or minimal medium agar plates complemented by 0.2mM L-histidine and incubated for 16-20h at 37 °C. 10µl 1mg/ml Ampicillin and DMSO were utilized as positive and negative controls, respectively. To test for potential human cell toxicity, human foreskin fibroblast cells were treated with 100 µg/ml of selected compounds or DMSO and their cell viabilities were determined by MTT (3-[4,5-dimethyl-triazolyl-2]2,5-diphenyl tetrazolium bromide) viability assay and trypan blue exclusion assay, as previously described (Shen, Liu et al. 2010).

6.5 ACKNOWLEDGEMENT

We gratefully acknowledge support of this research by the National Institutes of Health (NIAID U01-0700499 to Z.O and O. W.) and by travel grants from The American-Scandinavian Foundation, The Danish Chemical Society, The Otto Mønsted Foundation, The Knud Højgaard Foundation, The Augustinus Foundation, and The Oticon Foundation to S.T.H.

7.0 CONCLUSIONS

The main goal of our studies was to utilize a metabolic network analysis approach to understand and, ultimately pharmacologically target, cancer metabolism. I have shown that our approach, incorporating the additional biophysical constraint Molecular Crowding (MC) to metabolic modeling, was able to capture the main characteristics of the distinct cancer metabolism (Fig. 3.1), indicating that MC can potentially contribute to the Warburg effect.

Consistent with our theoretical work, I found experimentally that faster growing cells exhibit higher glycolytic activities than slower growing cells, even though they contain similar level of oncogenic Ras and tumorigenic potentials (Fig. 3.7-10). These results are consistent with what others have found using the same or similar model systems (Ramanathan A 2005; de Groof, Lindert et al. 2009). Not only in this fibroblast system, but also in primary lymphocytes (Wang, Marquardt et al. 1976) and thymocytes (Brand and Hermfisse 1997) it has been reported that glycolysis is tightly bound to proliferation. More interestingly, similar to *S. cerevisiae* (Tu, Kudlicki et al. 2005) the burst of glycolysis is coordinated in a cyclic fashion and coincides with the synthesis of cellular DNA in S phase in lymphocytes (Wang, Marquardt et al. 1976) possibly because the anaphase-promoting complex/cyclosome-Cdh1 (APC/C-Cdh1), a master regulator of G1-S transition inhibits glycolysis in proliferating lymphocytes by mediating the ubuquitinating two key metabolic enzymes for proteasomal degradation (Almeida et al. 2010, Colombo et al., 2010, Najafov and Alessi, 2010). Moreover, it has been shown that the accumulation of mitochondrial membranes, mitochondrial DNA and several mitochondrial proteins of the

respiratory chain also happens in S phase (Martinez-Diez, Santamaria et al. 2006). Thus, I speculate that MC may happen in S phase of fast proliferating cells to promote glycolysis and decrease mitochondrial respiration in order to achieve high ATP production rate while avoiding oxidative stress (Brand 1997) probably due to proteasome regulation. To validate this possibility, future experiments that simultaneously measure cell mass, density and volume during the cell cycle of single mammalian cells (Bryan et al. 2010; Godin et al. 2010) will shed lights on the direct involvement of MC during cell cycles and in cell metabolism regulation. It is also interesting to study the effect of MC on the cell proliferation related pathways and proteins, as well as proteasome activities to better understand how MC is executed and interacted with those signaling pathways.

All in all, this finding added a biophysical constraint, MC, to the previous proposed causes of the Warburg effect, that are permanent mitochondrial defects (Warburg 1956), hypoxic adaptation during carcinogenesis (Helmlinger, Yuan et al. 1997; Gatenby and Gillies 2004), and/or oncogenic alterations as well as metabolic enzyme mutations (Semenza 2009; Feng and Levine 2010; Locasale and Cantley 2010). Besides, this study helped explain why primary proliferating mammalian cells also utilize glycolysis pathway in the normoxic conditions.

It has been proposed that the Warburg effect may represent a unique target for cancer therapy (Kroemer and Pouyssegur 2008). However, our work suggests that the Warburg effect itself may be not an effective target for cancer therapy. Yet, it cannot be ruled out that individual tumor specific isozymes (Christofk, Vander Heiden et al. 2008) or different essential enzymes or lethal pairs can be identified by comparing the metabolic networks of cancer cells with normal cells. The reliable identification of enzyme targets within metabolic networks requires the

completely reconstructed high quality human metabolic networks to be cell type specific, which is not yet attainable. Therefore, to provide a proof-of-concept we instead investigated this potential in the single cell pathogen, *Staphylococcus aureus*, where the metabolic network of the strain N315 has been reliably reconstructed (Becker and Palsson 2005; Heinemann, Kummel et al. 2005).

Different *S. aureus* strains exhibit different pathogenesis, antibiotic susceptibility, virulence possibly due to their different genome sequences. In Chapter 4, we demonstrated that each of the examined 13 multidrug-resistant and sensitive *S. aureus* maintained a different set of unconditionally essential metabolic enzymes and synthetic lethal pairs. There are however 44 single enzymes and 10 lethal pairs common to all 13 *S. aureus* strains (Fig. 4.2). Therefore, to identify targets specific for *S. aureus*, metabolic reconstructions and *in silico* analyses of multiple strains of the same bacterial species are needed. In addition, we demonstrated that enzymes of the Fatty Acid Synthesis pathway (Chapter 5) and Histidine synthesis pathway (Chapter 6), which were among the predicted common single essential enzymes, were indeed essential and druggable, since we have been able to identify small inhibitors against them that kill *S. aureus* but not mammalian cells. Drug validation processes are always time-consuming and involve complicated in vitro enzyme assays. In Chapter 6, I have shown that whole-cell assays can be directly utilized for validation after identifying potential small inhibitors through virtual screening. Thus, the blueprint we presented in Chapter 5, together with the fast experimental improvement in Chapter 6, provided a proof-of-principle demonstration that the metabolic network analysis approach can be an attractive alternative to identify potential "druggable" targets (Russ and Lampel 2005).

This blueprint may be applicable for systematic drug target identification and subsequent lead candidate development in *any* cell type with high quality metabolic reconstruction, including cancer cells. More importantly, the notion of multiple metabolic networks comparison may be beneficial to identify generic targets against, e.g., different types of lung cancer or various species of Gram-positive or –negative bacteria, to simplify the treatment themes. By studying the metabolic networks with different drug resistant profiles, the predictions may help understand the mechanisms of drug resistance and further facilitate the route to target both drug sensitive and drug resistant metabolisms. When metabolic networks are available in a systematic fashion, the metabolic network analysis may be even capable to predict both the positive and negative effects of a given drug interruption at a specific metabolic state, thus defining the optimal metabolic state to maximize the positive effects while minimizing the side effects.

From a more general perspective, a key aim of network medicine is to reconstitute the physiological network states of cells (e.g., from cancer cell metabolism to normal cell metabolism) in a systematic and computable fashion (Pawson and Linding 2008). In the case of microbial infections, however, this entails the drug therapy-induced alteration of network states that are deleterious to the infectious organism, including its ability to synthesize biomass constituents or virulence factors, or to the function of its key cellular constituents (e.g., the ribosome) (Fig. 7.1a). The blueprint we demonstrate here combines genomic information and network biology-based strain-specific molecular target identification (Fig. 7.1b, left panel) with computational ranking of ~10^6 small molecules to identify potential antimicrobial scaffolds against the identified target enzymes (Fig. 7.1b, right panel). In turn, these define the chemical subspace for organism- or strain-specific lead development using traditional or structure-based methods of medicinal chemistry. This strategy could easily be extended to other pathways found

to be essential for bacteria, such as those regulating the generation of reactive oxygen intermediates that can potentiate the activity of individual antibiotic compounds (Kohanski, Dwyer et al. 2007; Kohanski, Dwyer et al. 2008; Dwyer, Kohanski et al. 2009).

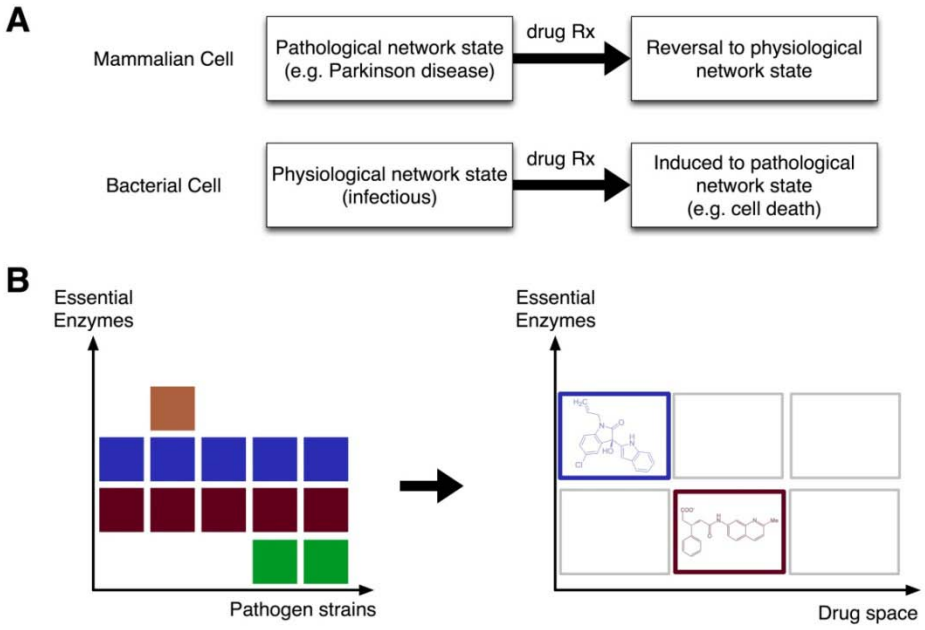


Figure 7. 1 Approach to network-based therapy, including strain-specific - antiinfective therapy.

a. For most diseases the aim of network medicine is to reestablish a physiological network state in a personalized fashion by affecting a ranked list of molecular targets, but for anti-microbial therapy the aim is just the opposite;
b. For each pathogenic bacteria strain the list of essential enzymes in a given strain can be calculated (left panel). In turn, for each target enzyme an experimentally validated small molecule hit that inhibits a given strain/essential enzyme combination defines the chemical subspace for subsequent compound optimization (right panel). In principle, a similar approach can be devised for all human diseases for which a molecular target list has been calculated and verified. Reproduced from (Shen and Liu et al. 2010).

While the work presented here provided encouraging results on the potential of systems-level analysis of metabolic networks in understanding cancer metabolism and facilitating antimicrobial drug discovery, there is still a significant amount of work to be done to specifically target cancer metabolism. First, the detailed mechanisms of how MC exerts its impact on cell

metabolism remain elusive. As MC can promote the stability of protein, DNA and RNA structures (Miyoshi and Sugimoto 2008; Zhou, Rivas et al. 2008; Dong, Qin et al. 2010), it is likely that the expression as well as the conformation of glycolytic enzymes are improved to convey higher reaction activities. Another possibility is through cell volume regulation, as MC is proposed to be a cell volume sensor (Lajvardi, Summers et al. 1996; Burg 2000; Al-Habori 2001), possibly due to its capability to increase the biochemical reaction activity of single macromolecules (Minton 2006; Zhou, Rivas et al. 2008). For example, cell swelling can increase glycogen synthesis and inhibit glycolysis (Peak, Alhabori et al. 1992; Haussinger and Schliess 1995; Lajvardi, Summers et al. 1996); In addition, cell swelling and MC had also been shown to interfere with the transfer of reducing equivalents through the mitochondrial malate/aspartate shuttle (Hussinger, Lang et al. 1990) and stimulate flux through the pentose phosphate pathway and NADPH generation (Saha, Stoll et al. 1992). The elucidation of the acting mechanisms of MC on cell metabolism in details will require the direct manipulation of MC levels, e.g., through over-expression of exogenous proteins, or injection of a crowding agent such as various concentrations of dextran or nanoparticles (Van Horn WD et al. 2010). Since the proteasome predominantly catalyzes protein degradation and clearance of damaged or regulatory proteins, thus being essential to maintain the protein concentration and turnover in the cell (Chondrogianni and Gonos 2010), the proteasome activities under the effect of MC are well worth investigating to better conceptualize the role of MC in regulating cell metabolism and cell growth.

Secondly, it has been reported that, in order to achieve high metabolic rates, cells maintain optimal MC (intracellular density) below or beyond which the overall metabolic rate would be diminished (Vazquez 2010). What the optimal density is for human normal and cancer

cells, respectively, and how MC is fine-tuned to achieve different metabolic states are interesting questions to be investigated.

In addition, beside the MC constraint, other biophysical or chemical constraints may exist that further confine the metabolism of tumor cells. One speculation is that reducing equivalents in the form of NADPH may also pose an essential impact on the cell metabolism as $NADPH/NADP^+$ ratio is critically important to protect from oxidative damage and for cell survival (Pollak N. et al. 2007). Therefore, the incorporation of additional physicochemical constraints to the existing cell metabolism model will be necessary to significantly improve the model in future.

Lastly, once the reconstruction of human cell metabolic networks in a tissue and cell type specific fashion are attainable, minimal essential enzyme targets or pathways may be calculated through metabolic network analysis to restore the metabolic state from the diseased state. As suggested in Chapter 4, most of the identified single essential enzymes are common among different strains, while essential synthetic lethal pairs are more strain-specific. To overcome the possible high mutation rate in those predicted essential single enzymes, a combination therapy which targets two single essential enzymes or both single essential enzymes together with the synthetic lethal pairs may improve the efficacy of the treatment. In this respect, poly-pharmacology, i.e., the inhibition of multiple targets with a single drug, may also be developed as tailored potent cancer therapy against cancer metabolism in different tissues or cell types.

BIBLIOGRAPHY

(2007). Prime 1.6. New York, NY, Schrödinger, LLC.

Abbanat, D., B. Morrow, et al. (2008). "New agents in development for the treatment of bacterial infections." Current Opinion in Pharmacology **8**(5): 582-592.

Acerenza, L. and M. Graña (2006). "On the Origins of a Crowded Cytoplasm." Journal of Molecular Evolution **63**(5): 583.

Adaleti, R., Y. Nakipoglu, et al. (2010). "Prevalence of phenotypic resistance of Staphylococcus aureus isolates to macrolide, lincosamide, streptogramin B, ketolid and linezolid antibiotics in Turkey." Brazilian Journal of Infectious Diseases **14**(1): 11-14.

Agnello, M., G. Morici, et al. (2008). "A method for measuring mitochondrial mass and activity." Cytotechnology **56**(3): 145-149.

Al-Habori, M. (2001). "Macromolecular crowding and its role as intracellular signalling of cell volume regulation." International Journal of Biochemistry & Cell Biology **33**(9): 844-864.

Alifano, P., R. Fani, et al. (1996). "Histidine biosynthetic pathway and genes: Structure, regulation, and evolution." Microbiological Reviews **60**(1): 44-&.

Almaas, E., Z. N. Oltvai, et al. (2005). "The activity reaction core and plasticity of metabolic networks." PLoS Comput. Biol. **1** (e68), 61-67

Almeida, A., J. P. Bolanos, et al. (2010). " E3 ubiquitin ligase APC/C-Cdh1 accounts for the Warburg effect by linking glycolysis to cell proliferation." PNAS, 107 (2) 738-741.

Ames, B. N. and B. N. Horecker (1956). "The Biosynthesis of Histidine: Imidazoleacetol Phosphate Transaminase." J. Biol. Chem. **220**: 113-128.

Ayala, F. R. R., R. M. Rocha, et al. (2010). "Glut1 and Glut3 as Potential Prognostic Markers for Oral Squamous Cell Carcinoma." Molecules **15**(4): 2374-2387.

Baba, T., T. Ara, et al. (2006). "Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection." Mol Systems Biol **2**: 2006.0008.

Baba, T., T. Bae, et al. (2008). "Genome sequence of Staphylococcus aureus strain newman and comparative analysis of staphylococcal genomes: Polymorphism and evolution of two major pathogenicity islands." Journal of Bacteriology **190**(1): 300-310.

Barabasi, A. L. and Z. N. Oltvai (2004). "Network biology: Understanding the cell's functional organization." Nature Reviews Genetics **5**(2): 101-U115.

Barbini, L., J. Rodriguez, et al. (2007). "Glyceraldehyde-3-phosphate dehydrogenase exerts different biologic activities in apoptotic and proliferating hepatocytes according to its subcellular localization." Molecular and Cellular Biochemistry **300**(1-2): 19-28.

Bayly, C. A., P. Cieplak, et al. (1993). "A well behaved elestrostatic potential based method using charge restraints for deriving atomic charges: The RESP model." J. Phys. Chem. **97**: 10269-10280.

Beard, D. A., S. C. Liang, et al. (2002). "Energy balance for analysis of complex metabolic networks." Biophysical Journal **83**(1): 79-86.
Becke, A. D. and D. R. Yarkony (1995). In Modern Electronic Structure Theory Part II. Singapore, World Scientific.
Becker, S. A. and B. O. Palsson (2005). "Genome-scale reconstruction of the metabolic network in Staphylococcus aureus N315: an initial draft to the two-dimensional annotation." Bmc Microbiology **5**.
Beg, Q. K., A. Vazquez, et al. (2007). "Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity." PNAS **104**(31): 12663-12668.
Bhattacharyya, A., S. Stilwagen, et al. (2002). "Draft sequencing and comparative genomics of Xylella fastidiosa strains reveal novel biological insights." Genome Research **12**(10): 1556-1563.
Biaglow, J. E., G. Cerniglia, et al. (1997). "Effect of oncogene transformation of rat embryo cells on cellular oxygen consumption and glycolysis." BBRC **235**(3): 739-742.
Bister, B., D. Bischoff, et al. (2004). "Abyssomicin C - A polycyclic antibiotic from a marine Verrucosispora strain as an inhibitor of the p-aminobenzoic acid/tetrahydrofolate biosynthesis pathway." Angewandte Chemie-International Edition **43**(19): 2574-2576.
Bonnet, S., S. L. Archer, et al. (2007). "A mitochondria-K+ channel axis is suppressed in cancer and its normalization promotes apoptosis and inhibits cancer growth." Cancer Cell **11**(1): 37-51.
Boran, A. D. W. and R. Iyengar (2010). "Systems approaches to polypharmacology and drug discovery." Current Opinion in Drug Discovery & Development **13**(3): 297-309.
Bosso, J. A., P. D. Mauldin, et al. (2010). "The association between antibiotic use and resistance: the role of secondary antibiotics." Eur J Clin Microbiol Infect Dis **29**(9): 1125-1129.
Brand, K. (1997). "Aerobic glycolysis by proliferating cells: Protection against oxidative stress at the expense of energy yield." Journal of Bioenergetics and Biomembranes **29**(4): 355-364.
Brand, K. A. and U. Hermfisse (1997). "Aerobic glycolysis by proliferating cells: A protective strategy against reactive oxygen species." FASEB Journal **11**(5): 388-395.
Bray, D. (1998). "SIGNALING COMPLEXES:Biophysical Constraints on Intracellular Communication." Annual Review of Biophysics and Biomolecular Structure **27**(1): 59-75.
Brezina, V. (2010). "Beyond the wiring diagram: signalling through complex neuromodulator networks." Philosophical Transactions of the Royal Society B-Biological Sciences **365**(1551): 2363-2374.
Brinster, S., G. Lamberet, et al. (2009). "Type II fatty acid synthesis is not a suitable antibiotic target for Gram-positive pathogens." Nature **458**: 83-86.
Brophy, S., K. M. Sheehan, et al. (2009). "GLUT-1 expression and response to chemoradiotherapy in rectal cancer." International Journal of Cancer **125**(12): 2778-2782.
Bryan A. K, Goranov A et al. (2010). “Measurement of mass, density, and volume during the cell cycle of yeast.” PNAS **107**(3):999-1004.
Bugrim, A., T. Nikolskaya, et al. (2004). "Early prediction of drug metabolism and toxicity: systems biology approach and modeling." Drug Discovery Today **9**(3): 127.

Buijs, M., J. A. Vossen, et al. (2009). "Specificity of the anti-glycolytic activity of 3-bromopyruvate confirmed by FDG uptake in a rat model of breast cancer." Investigational New Drugs **27**(2): 120-123.

Burg, M. B. (2000). "Macromolecular crowding as a cell volume sensor." Cellular Physiology and Biochemistry **10**(5-6): 251-256.

Butaye, P., L. A. Devriese, et al. (2003). "Antimicrobial growth promoters used in animal feed: Effects of less well known antibiotics on gram-positive bacteria." Clinical Microbiology Reviews **16**(2): 175-+.

Buzzai, M., D. E. Bauer, et al. (2005). "The glucose dependence of Akt-transformed cells can be reversed by pharmacologic activation of fatty acid beta-oxidation." Oncogene **24**(26): 4165-4173.

Campbell, J. W. and J. E. Cronan (2001). "*Escherichia coli* FadR positively regulates transcription of the fabB fatty acid biosynthetic gene." J. Bacteriol. **183**: 5982-5990.

Cao, W. G., S. Yacoub, et al. (2008). "Dichloroacetate (DCA) sensitizes both wild-type and over expressing Bcl-2 prostate cancer cells in vitro to radiation." Prostate **68**(11): 1223-1231.

Case, D. A., T. A. Darden, et al. (2006). AMBER 9. University of California, San Francisco.

Caspi, R., H. Foerster, et al. (2006). "MetaCyc: a multiorganism database of metabolic pathways and enzymes." Nucleic Acids Research **34**: D511-D516.

Chan, D. I. and H. J. Vogel (2010). "Current understanding of fatty acid biosynthesis and the acyl carrier protein." Biochemical Journal **430**: 1-19.

Chebotareva, N. A. (2007). "Effect of molecular crowding on the enzymes of glycogenolysis." Biochemistry-Moscow **72**(13): 1478-1490.

Chen, Z., W. Q. Lu, et al. (2007). "The Warburg effect and its cancer therapeutic implications." Journal of Bioenergetics and Biomembranes **39**: 267-274.

Chondrogianni N, Gonos ES. (2010). “Proteasome function determines cellular homeostasis and the rate of aging.” Adv Exp Med Biol, 694: 38-46

Christofk, H. R., M. G. Vander Heiden, et al. (2008). "The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth." Nature **452**(7184): 230-U274.

Cieplak, P., J. Caldwell, et al. (2001). "Molecular Mechanical Models for Organic and Biological Systems Going Beyond the Atom Centered Two Body Additive Approximation:Aqueous Solution Free Energies of Methanol and N-ethyl Acetamide, Nucleic Acid Base, and Amide Hydrogen Bonding and Chloroform/Water Partition Coefficients of the Nucleic Acid Bases." *J. Comput. Chem.* **22**: 1048-1057

Coe, E. L. and R. C. Strunk (1970). "effect of oxamate on glycolysis in intact ascites tumor cells .1. kinetic evidence for a dual glycolytic system." Biochimica Et Biophysica Acta **208**(2): 189-&.

Colombo, S. L., M. Palacios-Callender, et al. (2010). “ Anaphase-promoting complex/ cyclosome-Cdh1 coordinates glycolysis and glutaminolysis with transition to S phase in human T lymphocytes.” PNAS, 107 (44) 18868-18873.

Cossarizza, A., M. Baccaranicontri, et al. (1993). "A new method for the cytofluorometric analysis of mitochondrial-membrane potential using the j-aggregate forming lipophilic cation 5,5',6,6'-tetrachloro-1,1',3,3'-tetraethylbenzimidazolcarbocyanine iodide (JC-1)." Biochemical and Biophysical Research Communications **197**(1): 40-45.

Covert, M. W. and B. O. Palsson (2002). "Transcriptional regulation in constraints-based metabolic models of Escherichia coli." Journal of Biological Chemistry **277**(31): 28058-28064.

Cox, J. M., T. R. Hawkes, et al. (1997). "The Design and Synthesis of Inhibitors of Imidazole glycerol Phosphate Dehydratase as Potential Herbicides." Pestic. Sci. **50**: 297-311.
Cuezva, J. M., A. D. Ortega, et al. (2009). "The tumor suppressor function of mitochondria: Translation into the clinics." Biochimica Et Biophysica Acta-Molecular Basis of Disease **1792**(12): 1145-1158.
D'Costa, V. M., K. M. McGrann, et al. (2006). "Sampling the antibiotic resistome." Science **311**(5759): 374-377.
Dang, L., D. W. White, et al. (2010). "Cancer-associated IDH1 mutations produce 2-hydroxyglutarate (vol 462, pg 739, 2010)." Nature **465**(7300): 966-966.
David, M. Z. and R. S. Daum (2010). "Community-Associated Methicillin-Resistant Staphylococcus aureus: Epidemiology and Clinical Consequences of an Emerging Epidemic." Clinical Microbiology Reviews **23**(3): 616-+.
de Groof, A. J. C., M. M. T. Lindert, et al. (2009). "Increased OXPHOS activity precedes rise in glycolytic rate in H-RasV12/E1A transformed fibroblasts that develop a Warburg phenotype." Molecular Cancer **8**.
DeBerardinis, R. J., J. J. Lum, et al. (2008). "The Biology of Cancer: Metabolic Reprogramming Fuels Cell Growth and Proliferation." Cell Metabolism **7**(1): 11.
Deutscher, D., I. Meilijson, et al. (2006). "Multiple knockout analysis of genetic robustness in the yeast metabolic network." Nature Genetics **38**(9): 993-998.
Dong, H., S. B. Qin, et al. (2010). "Effects of Macromolecular Crowding on Protein Conformational Changes." PLoS Computational Biology **6**(7).
Drake, D. R., K. A. Brogden, et al. (2008). "Antimicrobial lipids at the skin surface." J Lipid Res **49**: 4-11.
Duarte, N. C., S. A. Becker, et al. (2007). "Global reconstruction of the human metabolic network based on genomic and bibliomic data." PNAS **104**(6): 1777-1782.
Durot, M., P. Y. Bourguignon, et al. (2009). "Genome-scale models of bacterial metabolism: reconstruction and applications." Fems Microbiology Reviews **33**(1): 164-190.
Dwarakanath, B. S. (2009). "Cytotoxicity, radiosensitization, and chemosensitization of tumor cells by 2-deoxy-D-glucose in vitro." J Cancer Res Ther **5 Suppl 1**: S27-31.
Dwyer, D. J., M. A. Kohanski, et al. (2009). "Role of reactive oxygen species in antibiotic action and resistance." Curr Opin Microbiol **12**: 482-489.
D'Ordine, R. L., T. J. Klem, et al. (1999). "N1-(5'-Phophoribosyl)adenosine-5'-Monophosphate Cyclohydrolase: Purification and Characterization of a Unique Metalloenzyme." Biochemistry **38**: 1537-1546.
Edwards, J. S., R. U. Ibarra, et al. (2001). "In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data." Nature Biotechnology **19**(2): 125-130.
Ellis, R. J. (2001). "Macromolecular crowding: an important but neglected aspect of the intracellular environment." Current Opinion in Structural Biology **11**(1): 114-119.
Ellis, R. J. (2001). "Macromolecular crowding: obvious but underappreciated." Trends in Biochemical Sciences **26**(10): 597-604.
Ellis, R. J. and A. P. Minton (2003). "Cell biology - Join the crowd." Nature **425**(6953): 27-28.
English, B. K. and A. H. Gaur (2010). "The Use and Abuse of Antibiotics and the Development of Antibiotic Resistance." Hot Topics in Infection and Immunity in Children Vi **659**: 73-82.
Essman, V., L. Perera, et al. (1995). "A smooth particle-mesh-Ewald method." J. Chem. Phys. **103**: 8577-8593.

Fajardo, A., N. Martinez-Martin, et al. (2008). "The Neglected Intrinsic Resistome of Bacterial Pathogens." PLoS One **3**(2).
Fantin, V. R., J. St-Pierre, et al. (2006). "Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology, and tumor maintenance." Cancer Cell **9**(6): 425-434.
Farooque, A., F. Afrin, et al. (2009). "Protection of normal cells and tissues during radio- and chemosensitization of tumors by 2-deoxy-D-glucose." J Cancer Res Ther **5 Suppl 1**: S32-35.
Favier, J., J. J. Briere, et al. (2009). "The Warburg Effect Is Genetically Determined in Inherited Pheochromocytomas." PLoS One **4**(9).
Feist, A. M., C. S. Henry, et al. (2007). "A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information." Mol Syst Biol. **3**: 121.
Feist, A. M., M. J. Herrgard, et al. (2009). "Reconstruction of biochemical networks in microorganisms." Nature Reviews Microbiology **7**(2): 129-143.
Feist, A. M. and B. O. Palsson (2008). "The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli." Nature Biotechnology **26**(6): 659-667.
Fell, D. A. S., J.A. (1986). Fat synthesis in adipose tissue. An examination of stoichiometric constraints. J Biochem. **238:** 781-786.
Feng, Y., C. J. Chen, et al. (2008). "Evolution and pathogenesis of Staphylococcus aureus: lessons learned from genotyping and comparative genomics." Fems Microbiology Reviews **32**(1): 23-37.
Feng, Z. and A. J. Levine (2010). "The regulation of energy metabolism and the IGF-1/mTOR pathways by the p53 protein." Trends in Cell Biology In Press, Corrected Proof.
Fernandez, F. J., M. C. Vega, et al. (2004). "Structural Studies of the Catalytic Reaction Pathway of a Hyperthermophilic Histidinol-phosphate Aminotransferase." J. Biol. Chem. **279**: 21478-21488.
Feron, O. (2010). "Tumor-Penetrating Peptides: A Shift from Magic Bullets to Magic Guns." Science Translational Medicine **2**(34): 34ps26-34ps26.
Fischbach, M. A. and C. T. Walsh (2009). "Antibiotics for Emerging Pathogens." Science **325**(5944): 1089-1093.
Fischer, K., P. Hoffmann, et al. (2007). "Inhibitory effect of tumor cell-derived lactic acid on human T cells." Blood **109**(9): 3812-3819.
Fleischmann, R. D., M. D. Adams, et al. (1995). "Hhole-genome random sequencing and assembly of haemophilus-influenzae RD." Science **269**(5223): 496-512.
Fong, S. S., J. Y. Marciniak, et al. (2003). "Description and interpretation of adaptive evolution of Escherichia coli K-12 MG1655 by using a genome-scale in silico metabolic model." Journal of Bacteriology **185**(21): 6400-6408.
Forsyth, R. A., R. J. Haselbeck, et al. (2002). "A genome-wide strategy for the identification of essential genes in Staphylococcus aureus." Molecular Microbiology **43**(6): 1387-1400.
Fox, T. and P. A. Kollman (1998). "Application of RESP Methodology in the Parametrization of Organic Solvents." J.Phys.Chem.B **102**: 8070-8079.
Frezza, C. and E. Gottlieb (2009). "Mitochondria in cancer: Not just innocent bystanders." Seminars in Cancer Biology **19**(1): 4.

Friesner, R. A., J. L. Banks, et al. (2004). "Glide: A New Approach for Rapid, Accurate Docking and Scoring.1.Method and Assessment of Docking Accuracy." J.Med. Chem. **47**: 1739-1749.

Friesner, R. A., R. B. Murphy, et al. (2006). "Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes." J.Med. Chem. **49**: 6177-6196.

Fulton, A. B. (1982). "How crowded is the cytoplasm?" Cell **30**(2): 345-348.

Gambhir, A., R. Korke, et al. (2003). "Analysis of cellular metabolism of hybridoma cells at distinct physiological states." Journal of Bioscience and Bioengineering **95**(4): 317-327.

Ganapathy-Kanniappan, S., M. Vali, et al. (2010). "3-Bromopyruvate: A New Targeted Antiglycolytic Agent and a Promise for Cancer Therapy." Current Pharmaceutical Biotechnology **11**(5): 510-517.

Garcia-Perez, A. I., E. A. Lopez-Beltran, et al. (1999). "Molecular crowding and viscosity as determinants of translational diffusion of metabolites in subcellular organelles." Archives of Biochemistry and Biophysics **362**(2): 329-338.

Gatenby, R. A. and R. J. Gillies (2004). "Why do cancers have high aerobic glycolysis?" Nature Reviews Cancer **4**(11): 891-899.

Gatenby, R. A. and R. J. Gillies (2007). "Glycolysis in cancer: A potential target for therapy." International Journal of Biochemistry & Cell Biology **39**(7-8): 1358-1366.

Gerdes, S. Y., M. D. Scholle, et al. (2003). "Experimental determination and system-level analysis of essential genes in *Escherichia coli* MG1655." J Bacteriol **185**: 5673-5684.

Glynn, S. E., P. J. Baker, et al. (2005). "Structure and Mechanism of Imidazoleglycerol-Phosphate Dehydratase." Structure **13**: 1809-1817.

Godin M., F.F. Delgado, et al. (2010). “Using buoyant mass to measure the growth of single cells.” Nat Methods **7**(5):387-90.

Gohda, K., Y. Kimura, et al. (1998). "Theoretical evidence of the existence of a diazafulvene intermediate in the reaction pathway of imidazoleglycerol phosphate dehydratase: design of a novel and potent heterocycle structure for the inhibitor on the basis of the electronic structure-activity relationship study." Biochim. Biophys. Acta **1385**: 107-114.

Gohda, K., I. Mori, et al. (2000). "A CoMFA analysis with conformational propensity: An attempt to analyze the SAR of a set of molecules with different conformational flexibility using a 3D-QSAR method." J. Comput.-Aided Mol. Design **14**: 265-275.

Gohlke, H. and D. A. Case (2003). "Converging Free Energy Estimates: MM-PB(GB)SAStudies on the Protein–Protein Complex Ras–Raf." J. Comput. Chem. **25**(2): 238-250.

Goins, A. B. (2007). "Relative size depennant effects of molecular crowding on translational mobility." Biophysical Journal: 329A-329A.

Gottschalk, S., N. Anderson, et al. (2004). "Imatinib (STI571)-mediated changes in glucose metabolism in human leukemia BCR-ABL-positive cells." Clinical Cancer Research **10**(19): 6661-6668.

Guimaraes, C. R. W. and M. Cardozo (2008). "MM-GB/SA rescoring of docking poses in structure-based lead optimization." J. Chem. Inf. Model **48**(5): 958-970.

Guppy, M., P. Leedman, et al. (2002). "Contribution by different fuels and metabolic pathways to the total ATP turnover of proliferating MCF-7 breast cancer cells." Biochemical Journal **364**: 309-315.

Hahn, W. C., C. M. Counter, et al. (1999). "Creation of human tumour cells with defined genetic elements." Nature **400**: 464-468.

Hanahan, D. and R. A. Weinberg (2000). "The Hallmarks of Cancer." Cell **100**(1): 57.
Hancock, R. (2004). "A role for macromolecular crowding effects in the assembly and function of compartments in the nucleus." Journal of Structural Biology **146**(3): 281-290.
Hann, H. W. L. and E. Bombardieri (2000). "Serum markers and prognosis in neuroblastoma: Ferritin, LDH, and NSE." Neuroblastoma: 371-381.
Harder, M. E., R. C. Ladenson, et al. (1974). "Mutants of *Escherichia coli* with temperature-sensitive malonyl coenzyme A-acyl carrier protein transacylase." J Biol Chem **249**: 7468-7475.
Harper, M. E., A. Antoniou, et al. (2002). "Characterization of a novel metabolic strategy used by drug-resistant tumor cells." Faseb Journal **16**(12): 1550-1557.
Haruyama, K., T. Nakai, et al. (2001). "Structures of Escherichia Coli Histidinol-Phosphate Aminotransferase and Its Complexes with Histidinol-Phosphate and N-(5'-Phosphopyridoxyl)-L-Glutamate: Double Substrate Recognition of the Enzyme." Biochemistry **40**: 4633-4644.
Haussinger, D. and F. Schliess (1995). "cell-volume and hepatocellular function." Journal of Hepatology **22**(1): 94-100.
Heath, R. J. and C. O. Rock (2004). "Fatty acid biosynthesis as a target for novel antibacterials." Curr Opin Investig Drugs **5**: 146-153.
Heath, R. J., S. W. White, et al. (2001). "Lipid biosynthesis as a target for antibacterial agents." Progress in Lipid Research **40**(6): 467-497.
Heinemann, M., A. Kummel, et al. (2005). "In silico genome-scale reconstruction and validation of the Staphylococcus aureus metabolic network." Biotechnology and Bioengineering **92**(7): 850-864.
Helmlinger, G., F. Yuan, et al. (1997). "Interstitial pH and pO(2) gradients in solid tumors in vivo: High-resolution measurements reveal a lack of correlation." Nature Medicine **3**(2): 177-182.
Herigon, J. C., A. L. Hersh, et al. (2010). "Antibiotic Management of Staphylococcus aureus Infections in US Children's Hospitals, 1999-2008." Pediatrics **125**(6): E1294-E1300.
Herrgard, M. J., B. S. Lee, et al. (2006). "Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in Saccharomyces cerevisiae." Genome Research **16**(5): 627-635.
Holm, A. K., L. M. Blank, et al. (2010). "Metabolic and Transcriptional Response to Cofactor Perturbations in Escherichia coli." Journal of Biological Chemistry **285**(23): 17498-17506.
Hoppert, M. and F. Mayer (1999). "Prokaryotes." American Scientist **87**(6): 518-525.
Hsu, P. P. and D. M. Sabatini (2008). "Cancer Cell Metabolism: Warburg and Beyond." Cell **134**(5): 703.
Hu, Z. Q., J. W. Jiang, et al. (2007). "Effects of macromolecular crowding on biochemical reaction equilibria: A molecular thermodynamic perspective." Biophysical Journal **93**(5): 1464-1473.
Hulten, K. G., S. L. Kaplan, et al. (2010). "Hospital-Acquired Staphylococcus aureus Infections at Texas Children's Hospital, 2001-2007." Infection Control and Hospital Epidemiology **31**(2): 183-190.
Hussinger, D., F. Lang, et al. (1990). "Control of hepatic nitrogen metabolism and glutathione release by cell volume regulatory mechanisms." Eur J Biochem **193**(3): 891-898.

Irwin, J. J. and B. K. Shoichet (2005). "ZINC--a free database of commercially available compounds for virtual screening." J Chem Inf Model **45**: 177-182.
Ishii, N., K. Nakahigashi, et al. (2007). "Multiple High-Throughput Analyses Monitor the Response of E. coli to Perturbations." Science **316**(5824): 593-597.
Jemal, A., R. Siegel, et al. (2007). "Cancer Statistics, 2007." CA Cancer J Clin **57**(1): 43-66.
Ji, Y. D., B. Zhang, et al. (2001). "Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA." Science **293**(5538): 2266-2269.
Jones, P. G., T. Sura, et al. (2003). "Mupirocin resistance in clinical isolates of *Staphylococcus aureus*." Infect Control Hosp Epidemiol **24**: 300-301.
Jorgensen, W. L. C., J.; Madura, J.; Impey, R. W.; Klein, M. L. (1983). "Comparison of simple potential functions for the simulation of liquid water." J. Chem. Phys. **79**: 926-935.
Jozefczuk, S., S. Klie, et al. (2010). "Metabolomic and transcriptomic stress response of Escherichia coli." Molecular Systems Biology **6**.
Kaelin, W. G. (2009). "SDH5 Mutations and Familial Paraganglioma: Somewhere Warburg is Smiling." Cancer Cell **16**(3): 180-182.
Kaelin, W. G. and C. B. Thompson (2010). "CANCER Clues from cell metabolism." Nature **465**(7298): 562-564.
Kanehisa, M. and S. Goto (2000). "KEGG: Kyoto Encyclopedia of Genes and Genomes." Nucleic Acids Research **28**(1): 27-30.
Kang, Y., T. Durfee, et al. (2004). "Systematic mutagenesis of the *Escherichia coli* genome." J Bacteriol **186**: 4921-4930.
Kapatral, V., I. Anderson, et al. (2002). "Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586." Journal of Bacteriology **184**(7): 2005-2018.
Kauffman, K. J., P. Prakash, et al. (2003). "Advances in flux balance analysis." Current Opinion in Biotechnology **14**(5): 491-496.
Kazakova, S. V., J. C. Hageman, et al. (2005). "A clone of methicillin-resistant *Staphylococcus aureus* among professional football players." N Engl J Med **352**: 468-475.
Keatinge-Clay, A. T., A. A. Shelat, et al. (2003). "Catalysis, specificity, and ACP docking site of *Streptomyces coelicolor* malonyl-CoA:ACP transacylase." Structure **11**: 147-154.
Kilburn, D., J. H. Roh, et al. (2010). "Molecular Crowding Stabilizes Folded RNA Structure by the Excluded Volume Effect." Journal of the American Chemical Society **132**(25): 8690-8696.
Kim, J.-w., L. B. Gardner, et al. (2005). "Oncogenic alterations of metabolism and the Warburg effect." Drug Discovery Today: Disease Mechanisms **2**(2): 233.
Kimura, M., M. Yamaguchi, et al. (2005). "CD19-negative diffuse large B-cell lymphoma shows high serum LDH level and poor prognosis." Blood **106**(11): 1924.
Kittlick, P. D. and W. Babin (1985). "redox status of cultured fibroblasts - possible relations with specific catabolic rates of proteoglycans." Experimental Pathology **27**(1): 41-48.
Klevens, R. M., M. A. Morrison, et al. (2007). "Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States." JAMA **298**: 1763-1771.
Kodali, S., A. Galgoci, et al. (2006). "Gel-elongation assay for type II fatty acid synthesis." Nature Protocols: DOI: 10.1038/nprot.2006.1134.
Kohanski, M. A., D. J. Dwyer, et al. (2007). "A common mechanism of cellular death induced by bactericidal antibiotics." Cell **130**: 797-810.

Kohanski, M. A., D. J. Dwyer, et al. (2008). "Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death." Cell **135**: 1153-1156.

Kolev, Y., H. Uetake, et al. (2008). "Lactate dehydrogenase-5 (LDH-5) expression in human gastric cancer: Association with hypoxia-inducible factor (HIF-1 alpha) pathway, angiogenic factors production and poor prognosis." Annals of Surgical Oncology **15**(8): 2336-2344.

Kondoh, H. (2008). "Cellular life span and the Warburg effect." Experimental Cell Research **314**(9): 1923.

Koukourakis, M. I., A. Giatromanolaki, et al. (2003). "Lactate dehydrogenase-5 (LDH-5) overexpression in non-small-cell lung cancer tissues is linked to tumour hypoxia, angiogenic factor production and poor prognosis." British Journal of Cancer **89**(5): 877-885.

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

Lajvardi, R., J. C. Summers, et al. (1996). "Role of cell swelling and of macromolecular crowding as volume sensors in barnacle muscle cells." FASEB Journal **10**(3): A388.

Lam, A. P. and R. G. Wunderink (2006). "Methicillin-resistant S. aureus ventilator-associated pneumonia: Strategies to prevent and treat." Seminars in Respiratory and Critical Care Medicine **27**(1): 92-103.

Lee, B. (1983). "Calculation of volume fluctuation for globular protein models." PNAS **80**(2): 622-626.

Lee, D.-S., H. Burd, et al. (2009). "Comparative genome-scale metabolic reconstruction and flux balance analysis of multiple *Staphylococcus aureus* genomes identify novel anti-microbial drug targets." J Bacteriol **191**: 4015-4024.

Legan, M., B. Luzar, et al. (2009). "Expression of cyclooxygenase-2, glucose transporter-1 and angiogenesis in gallbladder carcinomas and their impact on prognosis." Scandinavian Journal of Gastroenterology **44**(9): 1101-1108.

Lehár, J., B. R. Stockwell, et al. (2008). "Combination chemical genetics." Nat Chem Biol **4**: 674-681.

Li, Z., Y. Huang, et al. (2007). "The crystal structure of MCAT from Mycobacterium tuberculosis reveals three new catalytic models." J Mol Biol **371**: 1075-1083.

Locasale, J. W. and L. C. Cantley (2010). "Altered metabolism in cancer." BMC Biology **8**: Article No.: 88.

Lowy, F. D. (1998). "Medical progress - Staphylococcus aureus infections." New England Journal of Medicine **339**(8): 520-532.

Lowy, F. D. (2003). "Antimicrobial resistance: the example of Staphylococcus aureus." Journal of Clinical Investigation **111**(9): 1265-1273.

Ma, H. W., A. Sorokin, et al. (2007). "The Edinburgh human metabolic network reconstruction and its functional analysis." Molecular Systems Biology **3**.

Macromodel (2007). version 9.5; Schrödinger, LLC: New York, NY,.

Madhok, B. M., S. Yeluri, et al. (2010). "Dichloroacetate induces apoptosis and cell-cycle arrest in colorectal cancer cells." British Journal of Cancer **102**(12): 1746-1752.

Maertens, J. and P. A. Vanrolleghem (2010). "Modeling with a View to Target Identification in Metabolic Engineering: A Critical Evaluation of the Available Tools." Biotechnology Progress **26**(2): 313-331.

Mahadevan, R., J. S. Edwards, et al. (2002). "Dynamic flux balance analysis of diauxic growth in Escherichia coli." Biophysical Journal **83**(3): 1331-1340.
Maier, T., M. Leibundgut, et al. (2008). "The crystal structure of a mammalian fatty acid synthase." Science **321**: 1315-1322.
Marienhagen, J., N. Kennerknecht, et al. (2005). "Functional Analysis of All Aminotransferase Proteins Inferred from the Genome Sequence of Corynebacterium glutamicum." J. Bacteriol. **187**: 7639-7646.
Martinez, M. and P. Silley (2010). "Antimicrobial drug resistance." Handb Exp Pharmacol(199): 227-264.
Martinez-Diez, M., G. Santamaria, et al. (2006). "Biogenesis and Dynamics of Mitochondria during the Cell Cycle: Significance of 3 ' UTRs." Plos One **1**(1).
Martinezbeltran, J. and R. Canton (1994). "Mechanisms of antimicrobial resistance in gram-positive bacteria." Revista Clinica Espanola **194**: 803-813.
Mayevsky, A. (2009). "Mitochondrial function and energy metabolism in cancer cells: Past overview and future perspectives." Mitochondrion **9**(3): 165.
Mazure, N., F. Dayan, et al. (2006). "Hypoxia signaling - Impact on tumor metabolism, cell survival & cell death." Radiotherapy and Oncology **78**: 3.
Minton, A. P. (1993). "macromolecular crowding and molecular recognition." Journal of Molecular Recognition **6**(4): 211-214.
Minton, A. P. (2001). "The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media." Journal of Biological Chemistry **276**(14): 10577-10580.
Minton, A. P. (2006). "How can biochemical reactions within cells differ from those in test tubes?" J Cell Sci **119**(14): 2863-2869.
Minton, A. P. (2006). "Macromolecular crowding." Current Biology **16**(8): R269-R271.
Miyoshi, D. and N. Sugimoto (2008). "Molecular crowding effects on structure and stability of DNA." Biochimie **90**(7): 1040-1051.
Mo, M. L. and B. O. Palsson (2009). "Understanding human metabolic physiology: a genome-to-systems approach." Trends in Biotechnology **27**(1): 37-44.
Molnos, J., R. Gardiner, et al. (2003). "A continuous coupled enzyme assay for bacterial malonyl-CoA:acyl carrier protein transacylase (FabD)." Anal Biochem **319**: 171-176.
Mongkolrattanothai, K., J. C. Aldag, et al. (2009). "Epidemiology of community-onset Staphylococcus aureus infections in pediatric patients: an experience at a Children's Hospital in central Illinois." Bmc Infectious Diseases **9**.
Moran, G. J., A. Krishnadasan, et al. (2006). "Methicillin-resistant S-aureus infections among patients in the emergency department." New England Journal of Medicine **355**(7): 666-674.
Mori, I., G. Iwasaki, et al. (1995). "Synthesis of Inhibitors of Imidazole Glycerol Phosphate Dehydratase." J. Am. Chem. Soc. **117**: 4411-4412.
Motter, A. E., N. Gulbahce, et al. (2008). "Predicting synthetic rescues in metabolic networks." Molecular Systems Biology **4**: 10.
Moudgal, V. V. and G. W. Kaatz (2009). "Fluoroquinolone Resistance in Bacteria." Antimicrobial Drug Resistance, Vol 1: Mechanisms of Drug Resistance: 195-205.
Murabito, E., E. Simeonidis, et al. (2009). "Capturing the essence of a metabolic network: A flux balance analysis approach." Journal of Theoretical Biology **260**(3): 445-452.

Najafov A. and D. R. Alessi (2010) “Uncoupling the Warburg effect from cancer.” PNAS, 107 (45) 19135-19136.
Newman, D. J. and G. M. Cragg (2007). "Natural products as sources of new drugs over the last 25 years." Journal of Natural Products **70**(3): 461-477.
Newman, D. J. and G. M. Cragg (2009). "Natural product scaffolds as leads-to drugs." Future Medicinal Chemistry **1**(8): 1415-1427.
Nolop, K. B., C. G. Rhodes, et al. (1987). "Glucose-utilization invivo by human pulmonary neoplasms." Cancer **60**(11): 2682-2689.
Oefner, C., H. Schulz, et al. (2006). "Mapping the active site of *Escherichia coli* malonyl-CoA-acyl carrier protein transacylase (FabD) by protein crystallography." Acta Crystallogr D Biol Crystallogr **62**: 613-618.
Ohba, S., H. Fujii, et al. (2010). "Overexpression of GLUT-1 in the invasion front is associated with depth of oral squamous cell carcinoma and prognosis." Journal of Oral Pathology & Medicine **39**(1): 74-78.
Ohlsen, K., G. Dandekar, et al. (2008). "New trends in pharmacogenomic strategies against resistance development in microbial infections." Pharmacogenomics **9**(11): 1711-1723.
Ohta, D., I. Mori, et al. (1997). "Inhibitors of imidazoleglycerolphosphate dehydratase as herbicides." Weed Science **45**: 610-620.
Onufriev, A., D. Bashford, et al. (2000). J. Phys. Chem.B **104**: 3712-3715.
Ortega AD, S.-A. M., Giner-Sánchez D, Sánchez-Cenizo L, Willers I, Cuezva JM. (2009). "Glucose avidity of carcinomas." Cancer Letters **276**(2): 125.
Orth, A. P., S. Batalov, et al. (2004). "The promise of genomics to identify novel therapeutic targets." Expert Opin Ther Targets **8**: 587-596.
Overbeek, R., N. Larsen, et al. (2003). "The ERGO (TM) genome analysis and discovery system." Nucleic Acids Research **31**(1): 164-171.
Pollak N, Dölle C, Ziegler M. (2007). “The power to reduce: pyridine nucleotides--small molecules with a multitude of functions.” Biochem J. 402(2):205-18.
Pastor, R. W., B. R. Brooks, et al. (1988). "An analysis of the accuracy of Langevin and molecular dynamics algorithms. ." Mol. Phys. **65**: 1409-1419.
Pawson, T. and R. Linding (2008). "Network medicine." FEBS Lett **582**: 1266-1270.
Payne, D. J., M. N. Gwynn, et al. (2007). "Drugs for bad bugs: confronting the challenges of antibacterial discovery." Nature Reviews Drug Discovery **6**(1): 29-40.
Peak, M., M. Alhabori, et al. (1992). "regulation of glycogen-synthesis and glycolysis by insulin, ph and cell-volume - interactions between swelling and alkalinization in mediating the effects of insulin." Biochemical Journal **282**: 797-805.
Pearlman, D. A., D. A. Case, et al. (1995). Comput. Phys. Commun. **91**: 1.
Peters, K., G. Kamp, et al. (2009). "Changes in Human Endothelial Cell Energy Metabolic Capacities during in vitro Cultivation. The Role of "Aerobic Glycolysis" and Proliferation." Cellular Physiology and Biochemistry **24**(5-6): 483-492.
Pitkanen, E., J. Rousu, et al. (2010). "Computational methods for metabolic reconstruction." Current Opinion in Biotechnology **21**(1): 70-77.
Pouyssegur, J., F. Dayan, et al. (2006). "Hypoxia signalling in cancer and approaches to enforce tumour regression." Nature **441**(7092): 437-443.
Price, N. D., J. L. Reed, et al. (2004). "Genome-scale models of microbial cells: Evaluating the consequences of constraints." Nature Reviews Microbiology **2**(11): 886-897.

Qian, W. and B. Van Houten (2010). "Alterations in bioenergetics due to changes in mitochondrial DNA copy number." Methods **51**(4): 452-457.
Qin, J. Z., H. Xin, et al. (2010). "3-Bromopyruvate induces necrotic cell death in sensitive melanoma cell lines." Biochemical and Biophysical Research Communications **396**(2).
Raman, K. and N. Chandra (2009). "Flux balance analysis of biological systems: applications and challenges." Briefings in Bioinformatics **10**(4): 435-449.
Ramanathan A, W. C., Schreiber SL. (2005). "Perturbational profiling of a cell-line model of tumorigenesis by using metabolic measurements." PNAS **102**(17): 5992-5997.
Ravasz, E., A. L. Somera, et al. (2002). "Hierarchical organization of modularity in metabolic networks." Science **297**(5586): 1551-1555.
Reichert, A. S. and W. Neupert (2004). "Mitochondriomics or what makes us breathe." Trends in Genetics **20**(11): 555-562.
Resendis-Antonio, O., J. L. Reed, et al. (2007). "Metabolic reconstruction and Modeling of nitrogen fixation in rhizobium etli." PLoS Computational Biology **3**: 1887-1895.
Rodriguez-Enriquez, S., P. A. Vital-Gonzalez, et al. (2006). "Control of cellular proliferation by modulation of oxidative phosphorylation in human and rodent fast-growing tumor cells." Toxicology and Applied Pharmacology **215**(2): 208-217.
Rudin, L., J. E. Sjostrom, et al. (1974). "Factors affecting competence for transformation in staphylococcus-aureus." Journal of Bacteriology **118**(1): 155-164.
Russ, A. P. and S. Lampel (2005). "The druggable genome: an update." Drug Discovery Today **10**: 1577-1579.
Ryckaert, J. P., G. Ciccotti, et al. (1977). "Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes." J. Comput. Phys. **23**: 327-341.
Saha, N., B. Stoll, et al. (1992). "effect of anisotonic cell-volume modulation on glutathione-s-conjugate release, t-butylhydroperoxide metabolism and the pentose-phosphate shunt in perfused-rat-liver." European Journal of Biochemistry **209**(1): 437-444.
Schadt, E. E., J. Lamb, et al. (2005). "An integrative genomics approach to infer causal associations between gene expression and disease." Nat Genet **37**: 710-717.
Schafmeister, C., W. S. Ross, et al. (1995). LEaP, . University of California, San Francisco.
Schilling, C. H., J. S. Edwards, et al. (2000). "Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems." Biotechnology and Bioengineering **71**(4): 286-306.
Schilling, C. H. and B. O. Palsson (1998). "The underlying pathway structure of biochemical reaction networks." PNAS **95**(8): 4193-4198.
Schomburg, I., A. Chang, et al. (2004). "BRENDA, the enzyme database: updates and major new developments." Nucleic Acids Research **32**: D431-D433.
Schweitzer, B. A., P. J. Loida, et al. (2002). "Discovery of Imidazole Glycerol Phosphate Dehydratase Inhibitors through 3-D Database Searching." Bioorg. Med. Chem. Lett. **12**: 1743-1746.
Schweitzer, B. A., P. J. Loida, et al. (1999). "Design and synthesis of β-carboxamido phosphonates as potent inhibitors of imidazole glycerol phosphate dehydratase." Bioorg. Med. Chem. Lett. **9**: 2053-2058.
Segre, D., D. Vitkup, et al. (2002). "Analysis of optimality in natural and perturbed metabolic networks." PNAS **99**(23): 15112-15117.

Selvarasu, S., I. A. Karimi, et al. (2010). "Genome-scale modeling and in silico analysis of mouse cell metabolic network." Molecular Biosystems **6**(1): 152-161.
Selvarasu, S., V. V. T. Wong, et al. (2009). "Elucidation of Metabolism in Hybridoma Cells Grown in Fed-Batch Culture by Genome-Scale Modeling." Biotechnology and Bioengineering **102**(5): 1494-1504.
Semenza, G. L. (2009). "Regulation of cancer cell metabolism by hypoxia-inducible factor 1." Seminars in Cancer Biology **19**(1): 12.
Seo, S. and H. A. Lewin (2009). "Reconstruction of metabolic pathways for the cattle genome." BMC Systems Biology **3**.
Serre, L., E. C. Verbree, et al. (1995). "The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylase at 1.5-A resolution. Crystal structure of a fatty acid synthase component." J Biol Chem **270**: 12961-12964.
Sheikh, K., J. Forster, et al. (2005). "Modeling hybridoma cell metabolism using a generic genome-scale metabolic model of Mus musculus." Biotechnology Progress **21**(1): 112-121.
Shen, Y., J. Liu, et al. (2010). "A blueprint for antimicrobial hit discovery targeting metabolic networks." PNAS **107**: 1082-1087.
Sinclair, R. (1974). "Response of mammalian cells to controlled growth rates in steady-state continuous culture." In Vitro **10**(5-6): 295-305.
Sitkoff, D., K. A. Sharp, et al. (1994). J.Phys.Chem. **98**(1978-1988).
Sivaraman, J., Y. Li, et al. (2001). "Crystal Structure of Histidinol Phosphate Aminotransferase (HisC) from Escherichia Coli, and its Covalent Complex with Pyridoxal-5'-phosphate and L-Histidinol Phosphate." J. Mol. Biol. **311**: 761-776.
Sivaraman, J., R. S. Myers, et al. (2005). "Crystal Structure of Methanobacterium thermoautotrophicum Phosphoribosyl-AMP Cyclohydrolase His." Biochemistry **44**: 10071-10080.
Stacpoole, P. W. (1989). "the pharmacology of dichloroacetate." Metabolism-Clinical and Experimental **38**(11): 1124-1144.
Stacpoole, P. W., G. N. Henderson, et al. (1998). "Clinical pharmacology and toxicology of dichloroacetate." Environmental Health Perspectives **106**: 989-994.
Stausbol-Gron, B., H. Havsteen, et al. (1998). "Fibroblast growth in the soft agar clonogenic assay for cervix cancer radiosensitivity." British Journal of Cancer **78**(4): 550-551.
Stelling, J., S. Klamt, et al. (2002). "Metabolic network structure determines key aspects of functionality and regulation." Nature **420**(6912): 190-193.
Stepansky, A. and T. Leustek (2006). "Histidine biosynthesis in plants." Amino Acids **30**: 127-142.
Sudarshan, S., C. Sourbier, et al. (2009). "Fumarate Hydratase Deficiency in Renal Cancer Induces Glycolytic Addiction and Hypoxia-Inducible Transcription Factor 1 alpha Stabilization by Glucose-Dependent Generation of Reactive Oxygen Species." Molecular and Cellular Biology **29**(15): 4080-4090.
Sugahara, K. N., T. Teesalu, et al. (2010). "Coadministration of a Tumor-Penetrating Peptide Enhances the Efficacy of Cancer Drugs." Science **328**(5981): 1031-1035.
Sutphin, P. D., A. J. Giaccia, et al. (2007). "Energy Regulation: HIF MXIes It Up with the C-MYC Powerhouse." Developmental Cell **12**(6): 845.
Swietach, P., R. D. Vaughan-Jones, et al. (2007). "Regulation of tumor pH and the role of carbonic anhydrase 9." Cancer and Metastasis Reviews **26**(2): 299-310.

Teague, S. J., A. M. Davis, et al. (1999). "The design of leadlike combinatorial libraries." Angew. Chem. Int. Ed. Engl. **38**(24): 3743-3748.
Tenover, F. C. (2006). "Mechanisms of antimicrobial resistance in bacteria." American Journal of Infection Control **34**(5): S3-S10.
Thiele, I. and B. O. Palsson (2010). "A protocol for generating a high-quality genome-scale metabolic reconstruction." Nature Protocols **5**(1): 93-121.
Thiele, I., T. D. Vo, et al. (2005). "Expanded metabolic reconstruction of Helicobacter pylori (iIT341 GSM/GPR): an in silico genome-scale characterization of single- and double-deletion mutants." Journal of Bacteriology **187**(16): 5818-5830.
Thompson, D. C., C. Humblet, et al. (2008). "Investigation of MM-PBSA rescoring of docking poses." J. Chem. Inf. Model **48**(5): 1081-1091.
Torella, J. P., R. Chait, et al. (2010). "Optimal Drug Synergy in Antimicrobial Treatments." PLoS Computational Biology **6**(6).
Tu, B. P., A. Kudlicki, et al. (2005). "Logic of the yeast metabolic cycle: Temporal compartmentalization of cellular processes." Science **310**(5751): 1152-1158.
Tyson, J. J. and B. Novak (2010). "Functional Motifs in Biochemical Reaction Networks." Annual Review of Physical Chemistry, Vol 61 **61**: 219-240.
Van Horn WD, Ogilvie ME et al. (2009) “Reverse micelle encapsulation as a model for intracellular crowding” J Am Chem Soc. **131**(23):8030-9.
Vander Heiden, M. G. (2010). "Targeting Cell Metabolism in Cancer Patients." Science Translational Medicine **2**(31).
Vander Heiden, M. G., L. C. Cantley, et al. (2009). "Understanding the Warburg effect: the metabolic requirements of cell proliferation." Science **324**(5930): 1029-1033.
Vazquez, A. (2010). "Optimal cytoplasmatic density and flux balance model under macromolecular crowding effects." Journal of Theoretical Biology **264**(2): 356-359.
Vazquez, A., Q. K. Beg, et al. (2008). "Impact of the solvent capacity constraint on E. coli metabolism." BMC Systems Biology **2**.
Vazquez, A., M. A. de Menezes, et al. (2008). "Impact of Limited Solvent Capacity on Metabolic Rate, Enzyme Activities, and Metabolite Concentrations of S. cerevisiae Glycolysis." PLoS Computational Biology **4**(10).
Vazquez, A., J. X. Liu, et al. (2010). "Catabolic efficiency of aerobic glycolysis: The Warburg effect revisited." BMC Systems Biology **4**.
Verdonk, M. L., J. C. Cole, et al. (2003). "Improved Protein-Ligand Docking Using GOLD." Proteins **52**: 609-623.
Wang, J., S. Kodali, et al. (2007). "Discovery of platencin, a dual FabF and FabH inhibitor with in vivo antibiotic properties." PNAS **104**(18): 7612-7616.
Wang, J., P. Morin, et al. (2001). "Use of MM-PBSA in Reproducing the Binding Free Energies to HIV-1 RT of TIBO Derivatives and Predicting the Binding Mode to HIV-1 RT of Efavirenz by Docking and MM-PBSA." J. Am. Chem. Soc. **123**: 5221-5230.
Wang, J., S. M. Soisson, et al. (2006). "Platensimycin is a selective FabF inhibitor with potent antibiotic properties." Nature **441**(7091): 358-361.
Wang, J., R. M. Wolf, et al. (2004). "Development and testing of a general AMBER force field." J. Comput. Chem. **25**: 1157-1174.
Wang, T., C. Marquardt, et al. (1976). "Aerobic glycolysis during lymphocyte proliferation." Nature **261**(5562): 702.
Warburg, O. (1956). "origin of cancer cells." Science **123**(3191): 309-314.

Ward, P. S., J. Patel, et al. (2010). "The Common Feature of Leukemia-Associated IDH1 and IDH2 Mutations Is a Neomorphic Enzyme Activity Converting alpha-Ketoglutarate to 2-Hydroxyglutarate." Cancer Cell **17**(3): 225-234.
Werle, M., L. Jahn, et al. (2005). "Metabolic control analysis of the Warburg-effect in proliferating vascular smooth muscle cells." Journal of Biomedical Science **12**(5): 827-834.
Woods, G. L. and J. A. Washington (1995). Antibacterial susceptibility tests: dilution and disk diffusion methods. Manual of Clinical Microbiology. P. R. Murray, E. J. Baron, M. A. Pfaller, F. C. Tenover and R. H. Yolken. Washington, DC, ASM Press.
Wright, G. D. (2007). "The antibiotic resistome: the nexus of chemical and genetic diversity." Nature Reviews Microbiology **5**: 175-186.
Wright, G. D. (2010). "The antibiotic resistome." Expert Opinion on Drug Discovery **5**(8): 779-788.
Wu, D. J., Q. Wang, et al. (2010). "Epidemiology and molecular characteristics of community-associated methicillin-resistant and methicillin-susceptible Staphylococcus aureus from skin/soft tissue infections in a children's hospital in Beijing, China." Diagnostic Microbiology and Infectious Disease **67**(1): 1-8.
Wu, J., A. J. Daugulis, et al. (1992). "Correlation of ldh activity with loss of insect cell viability: an assessment of the LDH assay." Biotechnology Techniques **6**(4): 335-340.
Yang, D., M. T. Wang, et al. (2010). "Impairment of mitochondrial respiration in mouse fibroblasts by oncogenic H-RAS(Q61L)." Cancer Biology & Therapy **9**(2): 122-133.
Yang, X., J. L. Deignan, et al. (2009). "Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks." Nat Genet **41**: 415-423.
Zhang, Y.-M. and C. O. Rock (2008). "Membrane lipid homeostasis in bacteria." Nat Rev Microbiol **6**: 222-233.
Zhivotovsky, B. and S. Orrenius (2009). "The Warburg Effect returns to the cancer stage." Seminars in Cancer Biology **19**(1): 1.
Zhou, H. X., G. N. Rivas, et al. (2008). "Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences." Annual Review of Biophysics **37**: 375-397.
Zhu, K., M. R. Shirts, et al. (2007). "Improved Methods for Side Chain and Loop Predictions via the Protein Local Optimization Program: Variable Dielectric Model for Implicitly Improving the Treatment of Polarization Effects." J. Chem. Theory Comp. **3**: 2108-2119.
Zimmerman, S. B. and S. O. Trach (1991). "Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of escherichia-coli." Journal of Molecular Biology **222**(3): 599-620.
Zu, X. L. and M. Guppy (2004). "Cancer metabolism: facts, fantasy, and fiction." Biochemical and Biophysical Research Communications **313**(3): 459-465.