Seminars in Immunology 26 (2014) 610–622

Contents lists available at ScienceDirect

Seminars in Immunology

journal homepage: www.elsevier.com/locate/ysmim

Review

Systems-level modeling of mycobacterial metabolism for the identification of new (multi-)drug targets

Rienk A. Rienksma[a], Maria Suarez-Diez[a], Lucie Spina[b], Peter J. Schaap[a], Vitor A.P. Martins dos Santos[a,c,*]

[a] Laboratory of Systems and Synthetic Biology, Wageningen University and Research Centre, Dreijenplein 10, Wageningen 6703 HB, The Netherlands
[b] Centre National de la Rescherche Scientifique (CNRS), Institut de Pharmacologie et de Biologie Structurale (UMR 5089), Department of Tuberculosis and Infection Biology and Université de Toulouse (Université Paul Sabatier, Toulouse III), IPBS, 205 Route de Narbonne, BP 64182, F-31077 Toulouse, France
[c] Lifeglimmer GmbH, Markelstrasse 38, 12163 Berlin, Germany

ARTICLE INFO

Keywords:
Mycobacterium tuberculosis
Metabolic model
Constraint-based metabolic model
Gene essentiality
Metabolic state
Systems biology

ABSTRACT

Systems-level metabolic network reconstructions and the derived constraint-based (CB) mathematical models are efficient tools to explore bacterial metabolism. Approximately one-fourth of the *Mycobacterium tuberculosis* (Mtb) genome contains genes that encode proteins directly involved in its metabolism. These represent potential drug targets that can be systematically probed with CB models through the prediction of genes essential (or the combination thereof) for the pathogen to grow. However, gene essentiality depends on the growth conditions and, so far, no *in vitro* model precisely mimics the host at the different stages of mycobacterial infection, limiting model predictions. These limitations can be circumvented by combining expression data from *in vivo* samples with a validated CB model, creating an accurate description of pathogen metabolism in the host. To this end, we present here a thoroughly curated and extended genome-scale CB metabolic model of Mtb quantitatively validated using ^{13}C measurements. We describe some of the efforts made in integrating CB models and high-throughput data to generate condition specific models, and we will discuss challenges ahead. This knowledge and the framework herein presented will enable to identify potential new drug targets, and will foster the development of optimal therapeutic strategies.

1. The rise of multi-resistant *Mycobacterium tuberculosis* and the need for new intervention strategies

Mycobacterium tuberculosis (Mtb) is the etiological agent of tuberculosis (TB) and has re-emerged as a serious threat for human health. In 2012, TB claimed the lives of 1.3 million people [1]. The rapid appearance of multi, extensively and totally drug-resistant strains, emphasizes the adaptability of Mtb and has raised concerns of its impact to human health. Furthermore, due to the diverse genetic predisposition of the infected subjects, uncertainties on long-term adverse effects and other safety concerns regarding the rise of drug resistant strains, the development of new, effective and affordable TB drugs has been slow [2]. New (combined) therapeutic strategies are urgently required to combat these drug-resistant strains [3].

In vitro studies have revealed sets of genes that are essential for growth and survival under laboratory growth conditions [4,5]. Due to the differences between the *in vivo* and the *in vitro* environments this does not automatically imply that these sets of genes are suitable drug targets. Besides, given all cellular components from different types of networks, genes (and their products) that may be not essential on their own can be indispensable in combinations not immediately obvious. A vital improvement would be the expansion of these studies to *in vivo* or *ex vivo* models, such as animal models, which would as faithfully as possible mimic the onset and progression of the infection, as well as the strategies against it [6]. An alternative and complementary method to

Abbreviations: CB, constraint-based; Mtb, *Mycobacterium tuberculosis*; TB, tuberculosis; gdw, grams of cell dry weight; BCG, *Mycobacterium bovis* Bacillus Calmette-Guérin.

* Corresponding author at: Laboratory of Systems and Synthetic Biology, Wageningen University and Research Centre, Dreijenplein 10, Wageningen 6703 HB, The Netherlands. Tel.: +31 317 482865.
E-mail address: vitor.martinsdossantos@wur.nl (V.A.P. Martins dos Santos).

http://dx.doi.org/10.1016/j.smim.2014.09.013

identify suitable drug targets is to use mathematical descriptions of the metabolism of Mtb under *in vivo* conditions, circumventing experimental difficulties that arise with *in vivo* and *ex vivo* studies.

Approximately one-fourth of the annotated mycobacterial gene pool encodes structural proteins known to be involved in its metabolism presenting a wealth of enzymes and metabolites as potential drug targets. Stoichiometric genome-scale models of metabolism are essential to identify possible metabolic drug targets, as they provide a holistic view on metabolism. Drug targets in the form of enzymes encoded by their specific genes, have been identified by gene essentiality predictions based on modeling the *in vivo* environment [7]. Recent insights have clarified the picture of available metabolites to Mtb inside the host and shed new light on *in vivo* gene essentiality predictions [8–11].

Predictions on gene essentiality can be done using constraint-based (CB) metabolic models by simulating the effect of total loss of an enzyme function in a metabolic network. This black and white scenario where a drug is able to completely shut down an enzymatic reaction is not fully realistic. In most cases, drug effects are subtler, leading to only a partial loss of function [12]. Furthermore, and owing to the network structures in which they are embedded, genes may code for proteins that are not essential *per se*, but which do become so if equally non-essential proteins to which they are connected become dysfunctional or absent. A reliable metabolic network topology, knowledge of the available metabolites in the host, *in vivo* growth and survival requirements and strategies, and reliable and quantitative predictions of metabolic activity are important and thus far overlooked.

A stoichiometric genome-scale CB metabolic model that is experimentally validated, not only qualitatively for the correct network topology, but also quantitatively for predicting fluxes, provides many opportunities to further identify metabolic bottlenecks and weak spots. Instead of using only qualitative, topology based, methods such models can be explored for new drug targets and novel synergistic drug combinations using more realistic quantitative approaches. For example, in addition to simulating the effect of a knock out of given genes or combinations thereof, the effect of a partial loss of function induced by a drug can also be simulated. Simulating the effect of decreasing the function of enzymes that can be targeted with known drugs can highlight alternative metabolic escape routes that become more relevant under these conditions paving the way to the development of more efficient therapeutic strategies.

Here we present a new genome-scale CB model of Mtb metabolism, sMtb (*in silico Mycobacterium tuberculosis*), which builds upon three previously published models and which is experimentally validated in great detail. Our model also includes recently discovered or annotated reactions and pathways, has undergone extensive manual curation and outperforms its predecessors in terms of both qualitative and quantitative predictions. We discuss the applications of this model for the identification of possible drug targets, to the unraveling of potentially unknown interconnections and for the development of future intervention strategies.

2. Mathematical models of metabolism

There are different types of metabolic models, all of them based on networks of metabolites that are interconnected through enzymatic, spontaneous, or transport reactions. These metabolic networks are reconstructed from literature and annotated genome data.

CB metabolic models are stoichiometric, mass, charge and energy-balanced scaffolds that describe steady-state kinetics, whereas dynamic metabolic models are explicitly time-dependent and enable to determine the changes in the concentration of metabolites over time. Thus, dynamic metabolic models enable more accurate descriptions of metabolism, but require many detailed kinetic parameters, such as rate-constants of every enzyme. Such kinetic parameters are often unknown and obtaining them experimentally is often difficult or impossible. Therefore, for a genome-scale dynamic model, many of these parameters are unavailable and many of them would have to be fitted to the model, which would diminish its predictive power. In addition, simulations with these models are computationally costly, making dynamic models thus far unsuitable to describe metabolism on a genome-scale.

Genome-scale CB metabolic modeling provides a holistic view on metabolism and transport. A metabolic network forms the foundation of a CB metabolic model (Fig. 1). The stoichiometry of each reaction is written in a stoichiometric matrix where negative numbers represent the consumption of metabolites and positive numbers represent the formation of metabolites. This stoichiometric matrix ensures that the system is in steady-state, as for every reaction no metabolite can accumulate. Through the application of constraints, hence they are called 'constraint-based', the number of possible metabolic states can be lowered, to best predict the actual metabolic state of an organism under given genetic and environmental conditions [13]. Applying too many constraints can result in an infeasible model where no possible metabolic state can be found. CB metabolic models can be used to predict genes [14] and metabolites that are essential to synthesize precursors for growth [15]. A major advantage of genome-scale CB metabolic models as compared to dynamic models is that few parameters are required to describe the entire known metabolism of an organism. On the other hand, CB metabolic models are not easily adapted to describe the dynamics of the system, since they contain a stoichiometric matrix and are thus designed to operate in steady-state conditions where uptake and secretion fluxes are constant and there is no net accumulation of metabolic intermediates, which is only valid if the time scales under consideration are different enough. These metabolic models are based on optimization principles and need one or more optimization objectives to function. Optimization objectives in CB metabolic models can be multiple and describe what the organism

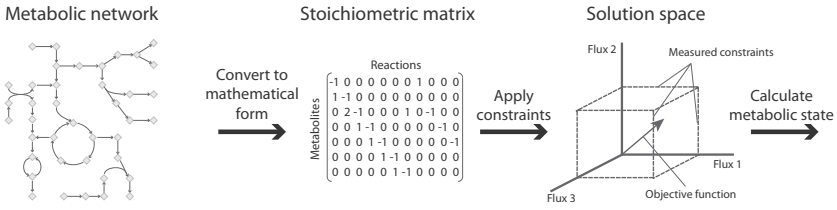

Fig. 1. Constraint-based model creation and functioning. A scaffold metabolic network is constructed from an annotated genome and completed after a rigorous survey of organism specific databases and literature. This metabolic network represents all the different possibilities for metabolites to travel through the network (metabolic states). After this network has been constructed, a stoichiometric matrix is created that encompasses the stoichiometry of all metabolic reactions under steady state conditions. Constraints on uptake and/or secretion rates are subsequently set, and the optimization of one or multiple objectives leads to the prediction of a metabolic state.