Confronting Metabolic-Related Conditions by Constraint-Based Simulations: A Review

Sayed-Amir Marashi [a]*

[a] Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran

Abstract. Finding drug targets is one of the main goals of systems biology. In this mini-review, we focus on the main advances which are recently achieved in drug target identification by analysing metabolic networks and pathways. We first introduce the mathematical definitions which are used in the analysis of metabolic networks. Then, we show how such models can be exploited to discover new targets to stop metabolic dysfunction, or alternatively, to prevent developing drug resistance by the new idea of metabolic transformation.

Keywords: Drug target prediction; Constraint-based models; Knockout simulation; Metabolic transformation algorithm (MTA).

1. 1. A formal introduction to metabolic networks

Traditionally, biochemists have analyzed biochemical reactions *in vitro* to find the mechanisms and parameters of metabolic reactions [1]. Throughout the last century, this type of study has been done for several enzymatic reactions in a variety of different organisms. If such information is available, one can model a biochemical system with ordinary differential equations (ODEs) to study its dynamical behavior. ODE models are deterministic models, that is, if the state of the system is known at a certain time point, it is possible to predict the system behavior at any other time point [2].

For the analysis of metabolic networks at the systems level, even in the absence of regulation, ODE modeling is not very appropriate for at least two reasons [1]. Firstly, for genome-scale metabolic networks, mechanisms and parameters are unknown for a considerable number of reactions. Secondly, the reaction mechanism and parameters are usually determined by analyzing enzymes in isolation. Even if these data are available for all the reactions, there is no guarantee that the enzymes and reactions behave similarly *in vivo*. In the analysis of biochemical reaction networks, some data are often available about each reaction without knowing the exact details. Based on this fact, a different strategy to model reaction networks is constraint-based modeling (CBM). By imposing the physicochemical constraints in a model, one can find out what is impossible, but the precise prediction of the system behavior impossible [3]. Informally speaking, there is a trade-off between the information content of the results and the knowledge required for the modeling.

In CBM, no prior knowledge (or assumptions) regarding the mechanisms and parameters of the reaction system is required. However, when such information is available, e.g. a certain intracellular flux is experimentally measured, this knowledge can be introduced into the model as an additional constraint [1, 4].

1.1. The basics of constraint-based modeling

Genetic networks and metabolic networks are the two systems controlling the fundamental mechanisms that govern biological systems [5]. While we are aware of the important interplay between these two networks [6], metabolic networks are usually studied without considering regulation. The reason lies in the fact that genome-scale metabolic networks are much easier to construct and analyze compared to genome-scale gene regulatory networks. Moreover, even in the absence of the regulation assumption, genome-scale metabolic networks have proven to be useful in predicting the intracellular fluxes and growth phenotypes of different organisms[7, 8].

As mentioned in Section 1.1, a metabolic network has two sets of elements: the metabolites, and the reactions. A reaction determines what metabolites (and at what ratio) react with each other to produce other metabolites. A system boundary (which is often the cell membrane) separates the internal metabolites, (i.e., those metabolites which are inside the system) from the external metabolites (those which are outside

* Corresponding author.
E-mail address: marashi@ut.ac.ir

the system). Boundary reactions are those reactions which convert internal metabolites to external ones (or vice versa), while internal reactions are the ones which convert internal metabolites to each other.

For a metabolic network *N* with m internal metabolites and n reactions, the stoichiometric matrix **S** is an $m \times n$ matrix, where element S_{ij} is the stoichiometric coefficient of metabolite i in reaction j.

1.2. The constraints, flux cone and flux space

In a metabolic network, flux through the i-th reaction, $v_i(t)$, is equivalent to the rate of this reaction, which is a function of time. If $\dot{c}$ is the vector denoting changes of internal metabolite concentrations, then for every flux distribution v(t), the equation $Sv(t)=\dot{c}$ holds [9].

In the analysis of metabolic networks, it is often assumed that the system is in steady state, which means that there is no net change in the concentration of the internal metabolites ($\dot{c} = 0$). Therefore, the flux balance equation can be written as $Sv = 0$, where $v \in R^n$ is the vector of flux distribution values [10].

Assuming that the metabolic network is in (quasi-)steady state is a good approximation, since the metabolic reactions are generally much faster than other biological functions, e.g., protein expression [11]. Therefore, it is reasonable to assume the system to be in steady-state, unless the dynamics of the system are studied over long time intervals.

There are two types of reactions in metabolic networks: *Irr* is the set of irreversible reactions, for which the corresponding flux values are always nonnegative; and *Rev* is the set of other reactions which are allowed to have both negative and non-negative values.

The flux cone *C* of a metabolic network N = (S; *Irr*), which is the set of all possible flux vectors in steady-state, is defined as:

$$C = \{v \in R^n \mid Sv = 0;\ v_i \geq 0 \text{ for all } i \in Irr\} \quad (1)$$

Sometimes, lower bounds and upper bounds of the flux values are also given (capacity constraints). In this case, the flux space, which is a polyhedron in n-dimensional space, is defined as [12]:

$$F = \{v \in R^n \mid Sv = 0;\ v_i \geq 0 \text{ for all } i \in Irr,\ \alpha \geq v \geq \beta\} \quad (2)$$

where β and α are the lower and upper bound vectors, respectively.

Equations 1 and 2 are the fundamental equations of constraint-based analysis of metabolic networks.

It should be emphasized here that the lower bound (resp. upper bound) of a reaction flux does not necessarily determine the minimum (resp. maximum) possible flux value through this reaction. Because of the flux balance assumption, constraints on one flux can result in constraining other fluxes.

Finding the minimum and maximum possible flux values is done by a technique called linear programming. A linear program (LP) is defined by a linear objective function and a set of linear constraints. Here, the objective function can be the flux through a certain reaction, or a linear combination of the fluxes through the reactions. For finding the optimal value of the objective function $k_1v_1+\ldots+k_nv_n$ in the flux space defined by Equation 1.3, the following LP should be solved:

$$\begin{aligned} \text{maximize (or minimize)} \quad & k_1 v_1 + \ldots + k_n v_n && k_1, \ldots, k_n \in R \\ \text{Subject to:} \quad & \mathbf{Sv} = \mathbf{0} \\ & v_i \geq 0 && \text{for all } i \in Irr \\ & \boldsymbol{\alpha} \geq v \geq \boldsymbol{\beta} \end{aligned}$$

As described above, if in the analysis of a metabolic system the focus is not on the dynamics, a low level of detail is required for building a genome-scale metabolic model [13]: for every reaction, the only required information is the reversibility type of reaction and the precise stoichiometric coefficients of the substrate(s) and product(s). The main challenge of reconstruction of a genome-scale metabolic network for a certain organism is to provide a comprehensive set of reactions in that organism.

There are at least two main sources of information which can be used in metabolic network reconstruction: bibliomic (literature) data and genomic data.

For several decades, metabolic reactions in different organisms have been identified by biochemists. This is an invaluable and indispensable source of information, which can be retrieved mainly from certain databases, like BRENDA. Further information about the reactions and enzymes may be directly obtained from literature mining.

The second important source of information for metabolic network reconstruction is functional genomics [14]. Sequence similarity search techniques are the basic tools to annotate sequenced genomes. When the sequence of a new genome is determined, the search techniques are used to find sequences with 'high' similarity to known enzymes. High similarity in the sequences implies the same enzymatic function.

Although the reactions can be identified automatically with a variety of methods, the resulting metabolic networks are very error-prone. In practice, most of the published metabolic networks are hand-curated. This is a very difficult and time consuming task, which may take from several months to a few years of work for a research team [15]. The manual curation and refinement of the model often relies on experimental, organism-specific information [15, 16].

It should be noted here that sometimes additional "auxiliary" reactions may be added to the metabolic network. An example is the case of biomass-producing reaction [13]. If the stoichiometric ratio of each component in the biomass is known, the biomass-producing reaction can be defined as a boundary reaction which produces an external metabolite (called 'Biomass') and consumes each of its precursors with the corresponding ratios. The biomass objective function, which is the flux through the biomass-producing reaction, is often used in flux balance analysis to model cell growth [17].

2. Drug target identification via CBM

Disease, in general, is caused by abnormal biological activities. Such abnormal activities might be caused by external agents, e.g., viruses and bacteria. On the other hand, there

are diseases which are caused by internally originated factors, like the case of cancer. In many cases, metabolism of cells is directly influenced under the pathological conditions [18]. In the following, we will discuss different strategies which can be used to confront pathological conditions, and hopefully, to cure diseases.

2.1. Targeting essential reactions and metabolites

Most of the known antibiotics have enzymatic targets [19]. Therefore, from the systems biology point of view, it is reasonable to search for potential drug targets within metabolic networks.

The simplest strategy to attack a pathogen is to block its "essential" reactions or metabolites, that is, reactions or metabolites which are essential to the growth of the pathogen [20, 21]. Several studies have ever tried to find drug targets from this viewpoint. As an example, Perumal et al. [22] used FBA to find potential drug targets in the multidrug resistant pathogen *Pseudomonas aeruginosa*. A similar strategy has been used to suggest drug targets in other organisms, including *Yersinia pestis* [23], *Leishmania major* [24] and even in human diseases like renal dysfunction [25] or cancer [26, 27]. However, the predictions of flux balance analysis are not always accurate, as in case of the prostate adenocarcinoma [28].

In case of multidrug resistance, sometimes it is useful to find non-trivial double knockouts that can prevent growth. Such double knock-outs, which are sometimes referred to as synthetic lethal pairs, have been reported in case of the drug resistant pathogen *Mycobacterium tuberculosis* [29]. Moreover, a comparable strategy has been successfully tested on human cancer cells, as it is known that fumarate hydratase is downregulated in cancer cells, and haem oxygenase is correctly predicted to be synthetically lethal with fumarate hydratase [30]. Inhibition of haem oxygenase does not harm normal cells because of the higher rates of fumarate hydratase in normal tissues.

Sometime, a good drug target lies in the metabolic network of the normal host cell, and not the target cell. Understanding host-pathogen interactions can play a crucial role in the modeling of metabolic links between human cells and the pathogenic factor. Developing systems-level models of host and pathogen metabolic networks can successfully simulate metabolite exchanges between the microbe and host, which in turn, can be exploited for determining the bottlenecks of pathogen metabolism.

A good example of this approach is the work of Huthmacher et al. [31], in which the authors tried to simulate the metabolism of a combined model of red blood cell and the protozoan parasite *Plasmodium falciparum*, the main cause of malaria.

2.1.1. Metabolic transformation algorithm

Metabolic transformation algorithm (MTA) [32] is a new strategy for facing pathological conditions. The idea behind MTA is to target metabolism in such a way that its active pathways become similar to the active pathways of normal healthy cells, or at least become as similar as possible.

MTA is originally suggested to revert disrupted metabolism of the senescence state [32]. Nevertheless, its applications are presumably much wider. For example, in the treatment of cancer, cancerous cells usually become resistant to the treatments. The reason is that, some mutant cancerous cells might become resistant to the administered drugs by chance, but these cells then will proliferate and replace the sensitive cancerous cells. MTA has the advantage that its application does not result in any selection against certain cells, and therefore, it would not result in drug resistance [33]. The same strategy would be useful in the treatment of other complex diseases like FSGS [34].

3. Conclusion

In the present paper, we briefly reviewed the main CBM strategies for encountering diseases via perturbation of metabolism. Genome-scale metabolic models have recently found a wide range of applications in Medical Biotechnology and Systems Medicine [35, 36]. In this manuscript, we discussed how these models can be useful in finding new drugs, either from the classical strategy (to fight the pathogenic cells) or the novel strategy of converting diseased cells to normal cells. Such strategies may find their application in other fields of biomedicine, like male infertility, by reverting the dysfunctional states of the cells to functional states.

References

1. Beard, Daniel A., Shou-dan Liang, and Hong Qian. "Energy balance for analysis of complex metabolic networks." Biophysical journal 83.1 (2002): 79-86.
2. Wolkenhauer, Olaf. "Systems biology: The reincarnation of systems theory applied in biology?." Briefings in bioinformatics 2.3 (2001): 258-270.
3. Palsson, Bernhard. "The challenges of in silico biology." Nature biotechnology 18.11 (2000): 1147.
4. Price, Nathan D., Jennifer L. Reed, and Bernhard Ø. Palsson. "Genome-scale models of microbial cells: evaluating the consequences of constraints." Nature Reviews Microbiology 2.11 (2004): 886-897.
5. Kitano, Hiroaki. "Perspectives on systems biology." New Generation Computing 18.3 (2000): 199-216.
6. Shlomi, Tomer, et al. "A genome-scale computational study of the interplay between transcriptional regulation and metabolism." Molecular systems biology 3.1 (2007): 101.
7. Edwards, Jeremy S., Rafael U. Ibarra, and Bernhard O. Palsson. "In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data." Nature biotechnology 19.2 (2001): 125-130.
8. Lee, Kyung Yun, et al. "The genome-scale metabolic network analysis of Zymomonas mobilis ZM4 explains physiological features and suggests ethanol and succinic acid production strategies." Microbial cell factories 9.1 (2010): 94.
9. Schuster, Stefan, and Ronny Schuster. "Detecting strictly detailed balanced subnetworks in open chemical reaction networks." Journal of Mathematical Chemistry 6.1 (1991): 17-40.
10. Fell, David A., and J. Rankin Small. "Fat synthesis in adipose tissue. An examination of stoichiometric constraints." Biochemical Journal 238.3 (1986): 781-786.

11. Edwards, Jeremy S., and Bernhard O. Palsson. "How will bioinformatics influence metabolic engineering?." Biotechnology and Bioengineering 58.2-3 (1998): 162-169.
12. Price, Nathan D., Jan Schellenberger, and Bernhard O. Palsson. "Uniform sampling of steady-state flux spaces: means to design experiments and to interpret enzymopathies." Biophysical journal 87.4 (2004): 2172-2186.
13. Durot, Maxime, Pierre-Yves Bourguignon, and Vincent Schachter. "Genome-scale models of bacterial metabolism: reconstruction and applications." FEMS microbiology reviews 33.1 (2008): 164-190.
14. Fell, David A., Mark G. Poolman, and Albert Gevorgyan. "Building and analysing genome-scale metabolic models." (2010): 1197-1201.
15. Thiele, Ines, and Bernhard Ø. Palsson. "A protocol for generating a high-quality genome-scale metabolic reconstruction." Nature protocols 5.1 (2010): 93-121.
16. Feist, Adam M., et al. "Reconstruction of biochemical networks in microorganisms." Nature Reviews Microbiology 7.2 (2009): 129-143.
17. Feist, Adam M., and Bernhard O. Palsson. "The biomass objective function." Current opinion in microbiology 13.3 (2010): 344-349.
18. Kaelin, William G., and Steven L. McKnight. "Influence of metabolism on epigenetics and disease." Cell 153.1 (2013): 56-69.
19. Bakheet, Tala M., and Andrew J. Doig. "Properties and identification of antibiotic drug targets." BMC bioinformatics 11.1 (2010): 195.
20. Lee, Sang Yup, et al. "Method for screening essential metabolites in growth of microorganisms." U.S. Patent No. 8,494,779. 23 Jul. 2013.
21. Kim, Hyun Uk, Seung Bum Sohn, and Sang Yup Lee. "Metabolic network modeling and simulation for drug targeting and discovery." Biotechnology journal 7.3 (2012): 330-342.
22. Perumal, Deepak, et al. "Targeting multiple targets in Pseudomonas aeruginosa PAO1 using flux balance analysis of a reconstructed genome-scale metabolic network." Journal of drug targeting 19.1 (2011): 1-13.
23. Navid, Ali, and Eivind Almaas. "Genome-scale reconstruction of the metabolic network in Yersinia pestis, strain 91001." Molecular bioSystems 5.4 (2009): 368-375.
24. Chavali, Arvind K., et al. "Systems analysis of metabolism in the pathogenic trypanosomatid Leishmania major." Molecular systems biology 4.1 (2008): 177.
25. Chang, Roger L., et al. "Drug off-target effects predicted using structural analysis in the context of a metabolic network model." PLoS Comput Biol 6.9 (2010): e1000938.
26. Folger, Ori, et al. "Predicting selective drug targets in cancer through metabolic networks." Molecular systems biology 7.1 (2011): 501.
27. Hadi, Mahdieh, and Sayed-Amir Marashi. "Reconstruction of a generic metabolic network model of cancer cells." Molecular BioSystems 10.11 (2014): 3014-3021.
28. Gatto, Francesco, et al. "Flux balance analysis predicts essential genes in clear cell renal cell carcinoma metabolism." Scientific reports 5 (2015): 10738.
29. Fang, Xin, Anders Wallqvist, and Jaques Reifman. "Development and analysis of an in vivo-compatible metabolic network of Mycobacterium tuberculosis." BMC systems biology 4.1 (2010): 160.
30. Frezza, Christian, et al. "Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase." Nature 477.7363 (2011): 225-228.
31. Huthmacher, Carola, et al. "Antimalarial drug targets in Plasmodium falciparum predicted by stage-specific metabolic network analysis." BMC systems biology 4.1 (2010): 120.
32. Yizhak, Keren, et al. "Model-based identification of drug targets that revert disrupted metabolism and its application to ageing." Nature communications 4 (2013).
33. Yizhak, Keren, et al. "Modeling cancer metabolism on a genome scale." Molecular systems biology 11.6 (2015): 817.
34. Sohrabi-Jahromi, Salma, Sayed-Amir Marashi, and Shiva Kalantari. "A kidney-specific genome-scale metabolic network model for analyzing focal segmental glomerulosclerosis." Mammalian Genome 27.3-4 (2016): 158-167.
35. Zhang, Cheng, and Qiang Hua. "Applications of genome-scale metabolic models in biotechnology and systems medicine." Frontiers in physiology 6 (2015).
36. Robinson, Jonathan L., and Jens Nielsen. "Integrative analysis of human omics data using biomolecular networks." Molecular BioSystems 12.10 (2016): 2953-2964.