BRIEFINGS IN BIOINFORMATICS. VOL 10. NO 4. 435–449
Advance Access publication March 15, 2009
doi:10.1093/bib/bbp011

Flux balance analysis of biological systems: applications and challenges

Karthik Raman and Nagasuma Chandra

Submitted: 28th November 2008; Received (in revised form): 6th February 2009

Abstract

Systems level modelling and simulations of biological processes are proving to be invaluable in obtaining a quantitative and dynamic perspective of various aspects of cellular function. In particular, constraint-based analyses of metabolic networks have gained considerable popularity for simulating cellular metabolism, of which flux balance analysis (FBA), is most widely used. Unlike mechanistic simulations that depend on accurate kinetic data, which are scarcely available, FBA is based on the principle of conservation of mass in a network, which utilizes the stoichiometric matrix and a biologically relevant objective function to identify optimal reaction flux distributions. FBA has been used to analyse genome-scale reconstructions of several organisms; it has also been used to analyse the effect of perturbations, such as gene deletions or drug inhibitions *in silico*. This article reviews the usefulness of FBA as a tool for gaining biological insights, advances in methodology enabling integration of regulatory information and thermodynamic constraints, and finally addresses the challenges that lie ahead. Various use scenarios and biological insights obtained from FBA, and applications in fields such metabolic engineering and drug target identification, are also discussed. Genome-scale constraint-based models have an immense potential for building and testing hypotheses, as well as to guide experimentation.

Keywords: *network reconstruction; metabolic network analysis; objective functions; genome scale modelling; reactome modelling*

INTRODUCTION

Systems biology, being a more holistic approach to study biological systems than the traditional reductionist approach, involves modelling and analysis of metabolic pathways, regulatory and signal transduction networks for understanding cellular behaviour. Cellular metabolism is often altered in disease, leading to an increased recognition of the importance of metabolic analysis in drug discovery. Metabolic engineering and related biotechnological applications benefit immensely from a systems view of the metabolism. The extreme complexity of cellular systems poses several challenges for a systematic analysis of various biochemical reactions taking place in a cell. A wide spectrum of techniques has been applied for the simulation and analysis of biochemical systems. These include stoichiometric techniques that rely on reaction stoichiometry and other constraints, kinetic pathway modelling using comprehensive mechanistic models, interaction-based analyses, Petri nets and qualitative modelling formalisms [1, 2]. A pre-requisite for utilizing these techniques is the availability of a detailed description of the possible components of a pathway and their inter-connections. While complete genome sequences provide pointers to such information, databases containing well-annotated pathways are still largely under development and not always easily available. At the present time, therefore, the first step of pathway modelling is a

Corresponding author. Nagasuma Chandra, Bioinformatics Centre and Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore 560012, India. Tel: +91-80-2293-2892; E-mail: nchandra@serc.iisc.ernet.in

Karthik Raman is currently a postdoctoral research associate in the Department of Biochemistry, at the University of Zurich. His research interests include the modelling of complex biological networks, and the analysis of their robustness and evolvability.

Nagasuma Chandra is a principal research scientist at the Bioinformatics Centre and Supercomputer Education and Research Centre at the Indian Institute of Science, Bangalore. Her group has ongoing projects on systems biology of *M. tuberculosis*, structural bioinformatics, computational genomics and in applying these to address fundamental issues in drug discovery.

careful curation of models from existing pathway databases and biochemical data buried in literature. Once an *in silico* model of the cell has been constructed, various genotypes, in terms of the genes that are present in the model can be analysed and the corresponding phenotypes, which pertain to the behaviour of the cell can be used to infer various aspects of cellular metabolism and function. The reconstruction and simulation of biological networks has applications in industrial biotechnology, drug discovery as well as for improving the functional annotation of micro-organisms and understanding cellular metabolism.

Several techniques are available for the modelling, simulation and analyses of pathways and networks involved in metabolism, regulation or signalling. Models vary in the level of detail that can be incorporated into them, which influence the accuracy of simulation results and the nature of insights that can be obtained from them. In this context, a simple classification of the modelling techniques would be: (a) interaction-based modelling involving graph-based representations of networks (static models), (b) constraint-based modelling involving stoichiometry (static models) and (c) mechanistic modelling, involving kinetic parameters as well as stoichiometry (dynamic models) [2].

Kinetic data available for the simulation of networks are quite scarce, limiting the number and size of systems in different species that can be studied through this approach. A useful alternative to mechanistic modelling techniques is constraint-based modelling. Constraint-based methods enforce cellular limitations on biological networks such as physico-chemical constraints, spatial [3] or topological constraints, environmental constraints or gene regulatory constraints [4]. This review focuses on flux balance analysis (FBA), a constraint-based approach for modelling metabolic networks.

CONSTRAINT-BASED MODELLING

Reconstruction of metabolic networks

Metabolic reconstruction is a process through which the various components of the metabolic network of a biological system, viz. the genes, proteins, reactions and metabolites that participate in metabolic activity are identified, categorized and inter-connected to form a network. Most often, the system is a single cell of interest and building on the genomic sequence as a scaffold, reconstructions can incorporate hundreds of reactions that approximate the entire metabolic activity of a cell. A comprehensive review of metabolic reconstruction has been published in [3]. Table 1 lists some resources useful for metabolic reconstruction and analysis.

Genome-wide annotations of protein sequences obtained through bioinformatics analyses are available for hundreds of organisms, providing excellent starting points for reconstruction of metabolic networks. The draft models obtained from these are refined and enriched by incorporating genetic, biochemical data and known metabolic functional data [5], the reconstructed models containing stoichiometric information for proteins in the metabolic networks. Flux-based analyses, which primarily require only stoichiometric information, can be more readily applied for the study of metabolic networks. The study of fluxes through such networks is informative and can give interesting insights even in the absence of detailed kinetic information. Metabolic fluxes can be seen as a fundamental determinant of cell physiology as they show quantitatively, the contributions of various pathways to overall cellular functions. A common way of relating the cell genotype to phenotype is by analysing the metabolic fluxes [6]. Constraint-based analyses of reconstructed metabolic networks have proved to be quite effective in various applications such as metabolic engineering [7–9], prediction of outcomes of gene deletions [10], drug-target identification, as reported in [11] and in the elucidation of cellular regulatory networks [12].

Flux balance analysis (FBA)

One specific example of metabolic modelling using a constraint-based approach is FBA, which uses linear optimization to determine the steady-state reaction flux distribution in a metabolic network by maximizing an objective function, such as ATP production or growth rate [13]. FBA involves carrying out a steady state analysis, using the stoichiometric matrix for the system in question. An important assumption is that the cell performs optimally with respect to a metabolic function, such as maximization of biomass production or minimization of nutrient utilization, on the premise that selection pressures during evolution guide systems towards optimality. Once an objective function is fixed, the system of equations can be solved to obtain a steady state flux distribution. This flux distribution is then used to interpret the metabolic capabilities of the system.

Table 1: Resources for metabolic reconstruction and software tools for FBA

Resource	URL	Description
Metabolic reconstruction		
Kyoto Encyclopaedia of Genes and Genomes (KEGG)	http://www.genome.jp/kegg/	Pathway databases for several organisms
BioCyc	http://www.biocyc.org	Pathway databases for several organisms
PEDANT	http://pedant.gsf.de/	Genome annotations
Reactome	http://www.reactome.org/	Curated database of biological processes in humans
Biomodels.net	http://www.biomodels.net/	Kinetic models of pathways, many published models from literature
BRENDA	http://www.brenda-enzymes.info/	Biochemical and molecular information on enzymes
SABIO-RK Database	http://sabio.villa-bosch.de/	System for the analysis of biochemical pathways – reaction kinetics
Software tools		
Constraint-based reconstruction and analysis (COBRA) toolbox	http://www.bioeng.ucsd.edu/research/research_groups/gcrg/downloads/COBRAToolbox/	Interfaces with MATLAB for extensive analysis of networks using FBA; performs gene deletions – single and multiple (can interface with LINDO, GLPK, CPLEX)
MetaFluxNet	http://mbel.kaist.ac.kr/lab/mfn/	Metabolic flux analysis
CellNetAnalyzer	http://www.mpi-magdeburg.mpg.de/projects/cna/cna.html	Structural and functional analysis of cellular networks
SNA: Stoichiometric network analysis	http://www.bioinformatics.org/project/?group_id=546	Mathematica toolbox for stoichiometric network analysis
Yana	http://yana.bioapps.biozentrum.uni-wuerzburg.de/	Network reconstruction, visualization and analysis
PathwayAnalyser	http://sourceforge.net/projects/pathwayanalyser	FBA and MoMA of metabolic networks; gene deletion studies
Systems Biology Research Tool	http://www.bioc.uzh.ch/wagner/software/SBRT/	Multiple methods for analysing stoichiometric networks
SBML Software Guide	http://sbml.org/SBML_Software_Guide	Resource list for software tools, model databases
Solvers for FBA/MoMA		
LINDO	http://www.lindo.com/	Commercial solver for optimization problems
CPLEX	http://www.ilog.com/products/cplex/	Commercial optimization software package
GNU linear programming toolkit (GLPK)	http://www.gnu.org/software/glpk/	Solver for LP problems
Object oriented quadratic programming (OOQP)	http://pages.cs.wisc.edu/~swright/ooqp/	Solver for QP problems

Figure 1 gives an overview of the steps involved in performing an FBA. Essentially, it involves four steps: (i) system definition, (ii) obtaining reaction stoichiometries, (iii) defining biologically relevant objective function and addition of other biochemical constraints and (iv) optimization. To define the system, the individual reactions in the model are listed in detail in terms of the metabolites involved, the genes and the corresponding enzymes involved in catalysing the reactions, as well as compartmentalization and reversibility. Transport reactions involved and external metabolites, which account for those metabolites which will be in exchange with the rest of the system, such as carbon sources for growth, co-factors that are ubiquitous, or end-products from a pathway or components of the biomass, are also identified at this stage. External metabolites are essentially Next, the dynamic mass balance of the metabolic system is described using the stoichiometric matrix $\mathbf{S}_{m\times n}$, relating the flux rates of enzymatic reactions, $v_{n\times 1}$ to time derivatives of metabolite concentrations, $x_{m\times 1}$ as $\mathrm{d}x/dt = \mathbf{Sv}$ where $\mathbf{v} = [v_1\, v_2 \cdots v_{n_i} \quad b_1\, b_2 \cdots b_{n_{ext}}]^T$; v_i signifies the internal fluxes, b_i represents the exchange fluxes in the system, n_i is the number of internal metabolites and n_{ext} is the number of external metabolites in the system. At steady state, $\mathrm{d}x/\mathrm{d}t = \mathbf{Sv} = 0$. Therefore, the required flux distribution belongs to the null space of $\mathbf{S}$. Since there are many more reactions than metabolites ($n > m$), the system is under-determined (with n–m degrees of freedom), necessitating the imposition of additional constraints to obtain meaningful solutions of steady state flux distributions [13].

In general, constraints may be of four types [4]: (i) physico-chemical constraints, (ii) spatial or topological constraints, (iii) condition dependent environmental constraints and (iv) regulatory

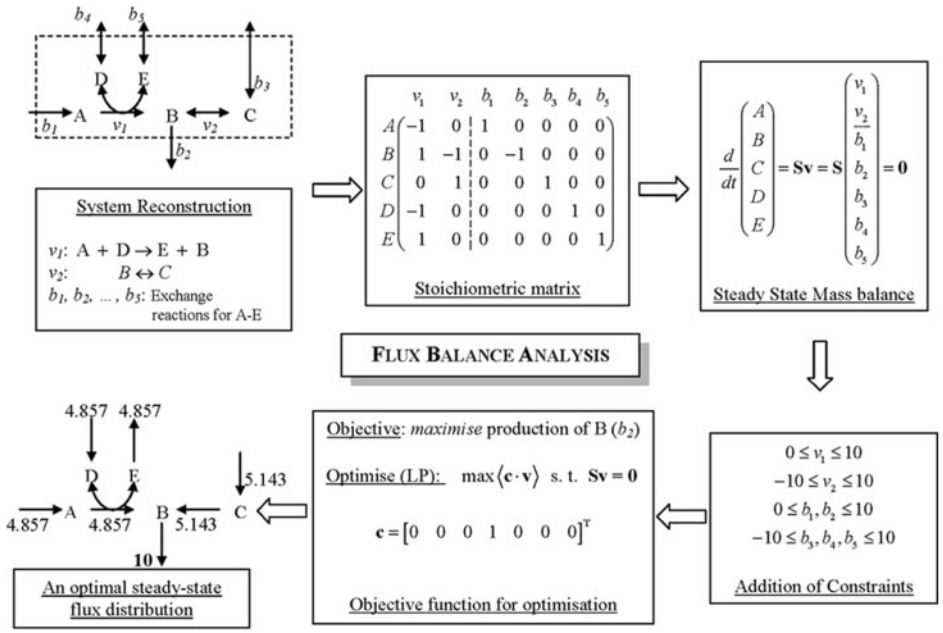

Figure 1: An outline of the major steps involved in FBA.

constraints. A detailed discussion of the various possible constraints and their mathematical representation has been reported elsewhere [4]. To solve the under-determined system, additional flux constraints can be imposed through the measurement of certain fluxes. More commonly, additional constraints are imposed by defining lower and upper bound for the fluxes. For example, the lower and upper bounds of the fluxes can be constrained as follows:

$$\begin{matrix} 0 < v_i < \infty \\ -\infty < b_i < \infty \end{matrix}.$$

which necessitates all internal irreversible reactions to have a flux in the positive direction and allows exchange fluxes to be in either direction. Practically, a finite upper bound can be imposed, which may also be decided based on the knowledge of cellular thermodynamics or actual measurements [13].

Constraints on particular exchange fluxes, nutrient uptake rates, consumption of ATP, phosphate-oxygen ratio may also be defined, depending on the *in silico* strain being analysed. Other constraints such as regulatory and thermodynamic constraints that can improve the predictive capability of FBA have been discussed in a later section.

Given that the measurement of fluxes is a tedious task, it is also possible to solve for a flux distribution by assuming that the under-determined metabolic network is optimized with respect to a certain objective [13–15]. This enables the formulation of the under-determined system as an optimization problem. The next critical step is to define an objective function that captures the biochemical goal of the system itself. A linear objective function results in a linear programming (LP) problem:

$$\max_{\mathbf{v}} \mathbf{c}^T\mathbf{v} \quad \text{s.t. } \mathbf{S}\cdot\mathbf{v} = 0$$

where **c** represents the objective function composition, in terms of the fluxes.

A variety of objective functions have been used in flux balance models [16]. The most common objective function involves the maximization of growth, or biomass, which allows for a wide range of predictions consistent with experimental observations [17, 18]. It has even been demonstrated that *Escherichia coli* tends to evolve towards maximization of biomass [19], although under some conditions, the behaviour of cellular systems is incompatible with biomass maximization [15, 20]. The biomass is usually represented as a stoichiometrically balanced reaction, describing the formation of biomass from various cellular components, as well as various co-factors, which are required to drive the process. Details of how a biomass objective function is determined are described in [5]. Other objective functions include: (i) minimization of ATP production, used to determine conditions of optimal metabolic energy efficiency [21, 22], (ii) minimization of nutrient uptake, (iii) maximization of

metabolite production, particularly to determine production capabilities of a particular cell, (iv) maximization of biomass *and* metabolite production [7] and (v) optimal metabolite channelling, by minimization of the absolute norm of the flux vector, or the Euclidean norm (quadratic objective function).

The choice of objective function is also influenced by the end goal of the study. Objective functions can be used to explore the capabilities and limitations of biochemical networks [13] and even for the analysis of metabolic network robustness [18, 23]. Sauer and co-workers have examined the predictive capacity of 11 linear and non-linear objectives functions, by evaluating the accuracy of FBA-based flux predictions through rigorous comparison to ^{13}C-based flux data from *E. coli* grown under six environmental conditions [24]. The different objective functions for FBA have also been mathematically defined and illustrated well in [24]. In a study of the mycolic acid pathway in *Mycobacterium tuberculosis*, we have evaluated two different objective functions describing the production of mycolic acids, illustrating how the biology is better captured by an objective function that accounts for the production of one or more mycolates rather than one necessitating the production of all mycolates [11].

Next, in the optimization phase, the set of mass balance constraints (along with other additional constraints), for a given objective function, are solved to obtain the steady state flux distribution. A variety of solvers can be used to solve LP problems. The COBRA Toolbox for MATLAB, developed by Palsson and co-workers [25], interfaces with a variety of these solvers, such as LINDO, CPLEX and GLPK. Table 1 lists some of the software tools and solvers useful for performing FBA. It is possible that sometimes, multiple optimal solutions may be obtained for a system. These alternate optimal solutions can be analysed to identify redundancies in the metabolic network [26]. In *Haemophilus influenzae*, a high degree of redundancy was found, especially for pathways involved in the production of non-essential amino acids, where 49 externally indistinguishable states were observed (on average), for a particular exchange flux [26].

Minimization of metabolic adjustment (MoMA)

A variant of FBA called MoMA that refers to the Minimization of Metabolic Adjustment adopts a quadratic optimization function, resulting in a quadratic programming (QP) problem [15, 27]. MoMA, a flux-based technique has the same stoichiometric constraints as FBA, but relaxes the optimal growth flux for mutants and seeks an approximate solution for a sub-optimal growth flux state, which is nearest in flux distribution to the unperturbed ('wild') state [15]. An important feature of MoMA is that the wild-type flux distribution used need not be obtained by performing an FBA; an experimentally determined flux distribution could serve better. Thus, objective functions for optimization, which at times may not reflect the physiological situation very accurately, can be circumvented using MoMA. MoMA also does not assume optimality of growth or any other metabolic function. Church and co-workers show that MoMA correctly identifies some lethal gene deletions in *E. coli*, which were not identified by FBA [15].

Another approach, in a similar vein, is known as regulatory on-off minimization (ROOM) [28], which attempts to minimize the *number* of significant flux changes from the wild-type flux distribution. The ROOM formulation requires the solution of a Mixed Integer Linear Programming (MILP) problem ROOM affords Furthermore improvements over FBA/MoMA; for instance, in an *E. coli* pyruvate kinase (pyk) knockout, ROOM predictions were much closer to experimentally observed values than those from FBA/MoMA [28]. ROOM also outperformed MoMA in lethality predictions on *Saccharomyces cerevisiae*; ROOM predictions concurred with FBA for all genes identified as viable by FBA but falsely classified as lethal by MoMA.

Analysis of perturbations

The most common perturbation studied using FBA is the deletion of one or more genes from the system. Gene deletion studies can be performed by constraining the reaction flux(es) corresponding to the gene(s) [and therefore, of their corresponding proteins(s)], to zero. Effects of inhibitors of particular proteins can also be studied in a similar way, by constraining the upper bounds of their fluxes to any defined fraction of the normal flux, corresponding to the extents of inhibition.

FBA gives a general idea of the metabolic capabilities of an organism; gene deletion studies using FBA yield information on the criticality of genes for the growth/survival of an organism. The analysis of perturbations using flux balance models

of metabolic networks provides a handle to analyse the lethality of individual gene deletions, as well as double knock-outs, to identify pairs of genes that are indispensable [29], as well as to determine and analyse synthetic genetic interactions [30].

FBA can also be employed to analyse the growth of an organism in various media (various carbon sources) and also the dependence on nutrient uptake rates. Such perturbations can give further insights into the metabolic capabilities of a system. For example, the growth of *E. coli* on glucose, under both aerobic and anaerobic conditions, as well as on a medium containing both glucose and lactose has been reported in [31], highlighting the differences in the utilization of different carbon sources and secretion of metabolites. The model GSMN-TB [32] analyses the fluxes of *M. tuberculosis* grown on a Middlebrook 7H10 medium, predicting lethality. The online web service for the model permits various perturbations to the medium and can predict gene essentiality under different media conditions.

ENHANCEMENTS TO FBA MODELS

Incorporation of regulatory information

The first attempts to integrate regulatory information into metabolic networks were made by Palsson and co-workers, who integrated regulatory constraints into FBA models (regulatory FBA; rFBA), using Boolean logic operators [12, 31, 33, 34]. The regulatory constraints essentially represent temporary flux constraints that arise due to a specific environment rather than physicochemical constraints that represent more fundamental restrictions. Integrated metabolic and transcriptional regulatory models consist of two interconnected components that represent metabolism and regulation. While the functional state of the metabolic component is represented by steady-state reaction fluxes, the functional state of the regulatory network at steady state is represented by a Boolean function, indicating the expression of each gene. The combined functional state of the entire system in a given environment, referred to as the metabolic-regulatory steady state (MRS), is described by a pair of consistent metabolic and regulatory steady states, which satisfy both the metabolic and regulatory constraints [35].

rFBA essentially involves the prediction of a regulatory and a metabolic steady state for short successive time intervals. For each interval, a regulatory state consistent with the metabolic steady state of the previous interval is computed, followed by an FBA to find a steady state flux distribution consistent with the current regulatory state. A new metabolic state could potentially lead to a new regulatory state, and the process is further iterated. An important limitation of this approach is that only a single metabolic state is chosen (arbitrarily) at each time interval from a space of possible solutions provided by FBA. This arbitrary choice leads to only a fraction of the space of dynamic flux profiles being explored. Nonetheless, rFBA provides a first glimpse, albeit qualitative, of the transcriptional events in the cell and their integration with metabolism. An excellent example of the advantage of rFBA over FBA has been illustrated in [31], where the growth of *E. coli* on glucose and lactose has been studied. FBA predicts a concurrent uptake of lactose and glucose, resulting in a rapid depletion of substrates and a higher growth rate, as well as a secretion of acetate and formate. Whereas, rFBA predicts a shift in gene expression, with the up-regulation of the lactose uptake and degradation machinery, alongside key galactose metabolism enzymes, enabling the system to use lactose as a carbon source following the depletion of glucose.

Ruppin and co-workers have proposed an alternative to rFBA, namely the steady state regulatory FBA (SR-FBA) [36], which attempts to comprehensively characterize the steady-state behaviours in a genome-scale integrated metabolic-regulatory network. To identify an MRS, SR-FBA involves the solution of an MILP, formulated by translating the Boolean regulatory constraints and the mapping between genes and reactions to linear equations. Applying SR-FBA to study the metabolism of *E. coli*, Ruppin and co-workers have analysed the effects of metabolic and regulatory constraints on metabolic behaviour. While metabolic constraints were found to determine the flux activity state of a majority of the genes, the role of transcriptional regulation was also found to play a role in determining the state of about 13–20% of the genes. While a large majority of these genes were direct targets of transcription factors, the rest were not directly regulated, indicating that transcriptional regulation can indirectly determine the activity of reactions that are not subjected to transcriptional factor regulation, by regulating associated pathway reactions. Thus, this study further emphasizes the need to obtain an

integrated view of the metabolic-regulatory network in a cell, also illustrating the versatility of FBA, for addressing such problems.

Incorporation of thermodynamic constraints

In traditional FBA, thermodynamic constraints are somewhat naively accounted for, in the specification of reaction reversibility. At this level of detail, dependence of reversibility on intracellular conditions, which may change in response to environmental changes are not accounted for. Qian and co-workers have proposed an alternative method, involving an energy balance analysis, through the imposition of non-linear constraints, which arise from the introduction of free energy changes into the constraints [37]. The energy balance analysis was also able to explain some of the incorrect essentiality predictions by FBA, earlier reported for *E. coli* [38]. Holzhütter and co-workers have proposed a method to include metabolite concentrations into FBA, to ensure thermodynamic realisability [39], overcoming the conventional reliance of FBA on the intuitive assumptions of reversibility of biochemical reactions. The incorporation of the additional constraints results in an MILP problem, with a quadratic scoring function. Applying their method to the metabolic network of *E. coli iJR*904 [40], they show that increasing network complexity entails increasing sensitivity of predicted flux distributions to variation of standard Gibb's free energy changes and metabolite concentration ranges.

Bi-level optimization

Prediction of gene deletion strategies for enhancing the production of specific metabolites has been reported using OptKnock [7, 41], which extends FBA to solve for the optimal flux distribution that simultaneously optimizes two objective functions, biomass growth as well as the secretion of a target metabolite, using a bi-level optimization technique.

In more complex systems such as mammalian systems, which perform multiple functions, multi-objective optimizations may be required to identify optimal flux distributions. Yarmush and co-workers have proposed a multi-objective optimization approach to perform both FBA and energy balance analysis to obtain optimal solutions, also demonstrating its application for analysing metabolic control in hepatocytes. Hepatocytes perform several metabolic functions at various levels depending on the environmental conditions. The authors have proposed that the multi-objective FBA of the system can be used in the design of an optimal bio-artificial liver support device [42].

APPLICATIONS OF FBA

Flux balance models of metabolic pathways enable the simulation of systems under varying experimental conditions. Such models have value in a variety of applications, such as optimization of bio-processes in industries, identification of drug targets and an improved annotation of genomes. Figure 2 depicts the various extensions to FBA, as well as the multiple applications of FBA.

Analysis of genome-scale metabolic networks (GSMNs)

GSMNs have been constructed and analysed using FBA, for various organisms in the past, including bacteria, archaea and eukaryotes. Reconstructions of human metabolism have also been reported [43, 44]. Table 2 summarizes some examples of the available studies on GSMNs, indicating the organism studied as well as the major insights obtained. GSMNs can also be readily subjected to a wide array of analyses. Large-scale gene deletion studies for organisms such as *S. cerevisiae* [10, 45] and *E. coli* [38] have been reported in literature. GSMNs thus have a wide range of applications, from improving the understanding of microbial metabolism and the capabilities of a metabolic network (for metabolic engineering applications), to insights obtained from gene deletion analyses, which can be applied for the identification of potential drug targets, in case of pathogenic organisms [32, 46]. With the availability of high-throughput transcriptomics data, which can be integrated with GSMNs and analysed using techniques such as rFBA, better predictions of metabolic capabilities and phenotypes will be possible.

Flux coupling analyses

It is possible to derive further insights from reconstructed networks, by an examination of their structure and topology based on convex analysis. Methods such as elementary flux mode analysis [47] and extreme pathway analysis [48] have been in vogue for analysing large metabolic networks and have been reviewed elsewhere [49]. Flux Coupling Finder (FCF), a method to analyse different types of

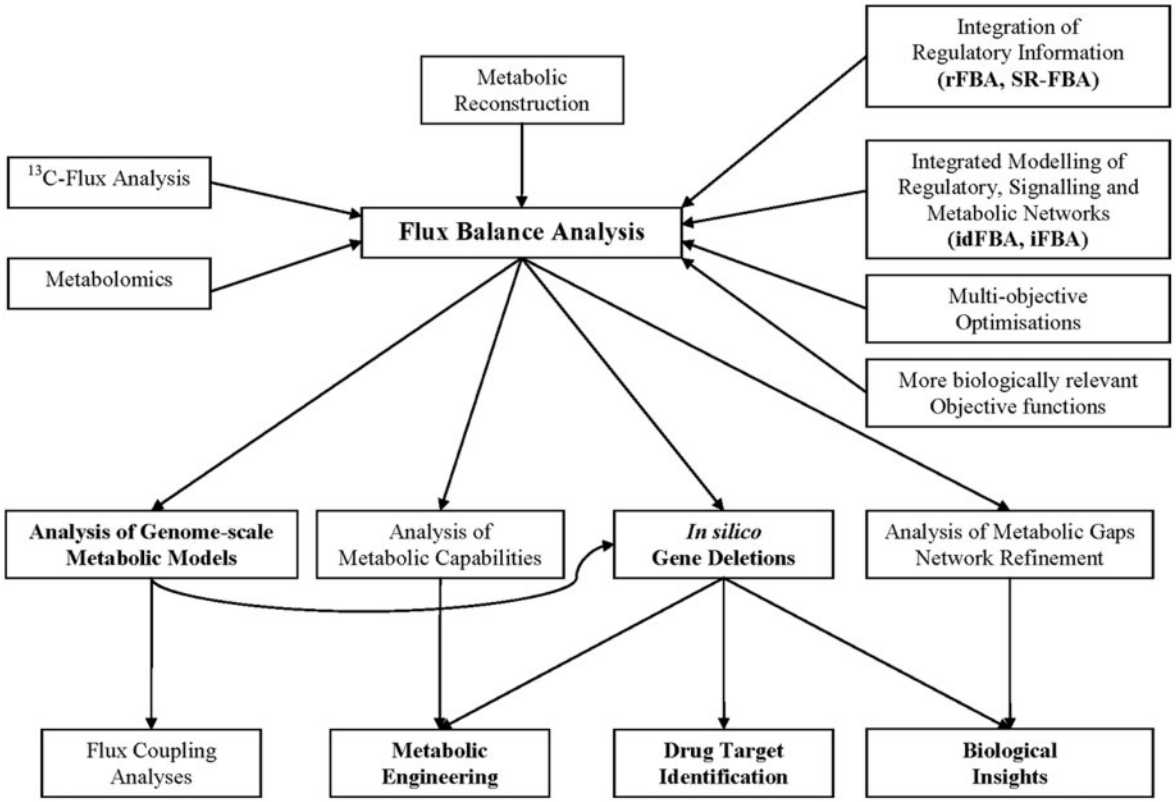

Figure 2: Extensions and applications of FBA.

coupling that can exist between fluxes in GSMNs has been proposed Maranas and co-workers [50]. The analysis enables the global identification of 'blocked' reactions, which are incapable of carrying flux under a certain condition, equivalent knock-outs—the set of all possible reactions whose deletion forces the flux through a particular reaction to zero, and sets of affected reactions—reactions whose fluxes are forced to zero if a particular reaction is deleted. It is easy to envisage applications for flux coupling analyses in metabolic reconstruction (refinement), as well as for detailed analyses of reconstructed networks. The FCF procedure applied to stoichiometric models of *Helicobacter pylori*, *E. coli* and *S. cerevisiae* indicated that 10%, 14% and 29% of their respective reactions are blocked unconditionally. This provides a ready list of the reactions that must be 'refined' in the network.

Drug target identification

In silico gene deletion studies help in identifying those enzymes in a metabolic network, which when deleted, adversely affects the fluxes across the entire network. Joyce and Palsson have given a good overview of using genome-scale *in silico* models to evaluate gene essentiality [51]. Gene deletions that are lethal can serve as a first list of putative drug targets, which can be further characterized by sequence analyses and structural studies. We have earlier constructed a model of mycolic acid biosynthesis in *M. tuberculosis*, analysing genes essential to the pathway, which led to the identification of seven new possible targets, apart from InhA, an already well known target [11]. Desaturases illustrate examples of such potential drug targets, whose prediction is strengthened by different types of computational target identification approaches [52, 53] and by an experimental study that revealed potent anti-mycobacterial activity by two lead compounds NAS-91 and NAS-21, owing in part to their ability to inhibit mycolic acid biosynthesis [54]. In a separate study, hard-coupled reaction sets in the genome-scale reconstruction of *M. tuberculosis* have been analysed for predicting potential drug targets [46]. The hard-coupled reaction sets mapped to several known drug targets, as well as potential targets from processes such as mycolic acid biosynthesis, mycothiol synthesis and menaquinone synthesis. Beste *et al.* [32] have used a GSMN of *M. tuberculosis* to identify essential genes, again providing important

Table 2: Details of some examples of GSMNs analysed using FBA, along with network statistics and insights obtained

Organism	Genes	Meta-bolites	Reactions	Analysis and insights	References
H. influenzae	296	343	488	• Definition of six different optimal metabolic phenotypes • Variation of metabolic 'genotype'	[78]
H. influenzae	400	451	461	• Genome-scale extreme pathway analysis • Network refinement • Prediction of co-regulated/co-expressed gene products • *In silico* gene deletions—minimal substrate, normal host conditions	[79]
Escherichia coli	660	436	720	• *In silico* gene deletions	[38]
Escherichia coli	904	625	931	• Genome-scale proton balancing • Analysis of metabolic gaps, dead-ends • Improved phenotype predictions	[40]
E. coli	1260	1039	2077	• Thermodynamic consistency analysis • Growth rate predictions • *In silico* gene deletions	[72]
Helicobacter pylori	291	340	388	• Minimal media identification • *In silico* gene deletions • Extreme pathway analysis	[80]
Saccharomyces cerevisiae	708	584	1175	• Analysis of metabolic capabilities • Prediction of metabolite productions	[77]
S. cerevisiae	750	646	1149	• Compartmentalized GSMN • *In silico* gene deletion and analysis	[10]
Staphylococcus aureus	619	571	640	• Minimal media identification • *In silico* gene deletions • Sensitivity of growth rate to oxygen uptake on different carbon sources • Potential drug targets	[75]
Methanosarcina barkeri	692	558	509	• Prediction of cellular phenotypes • Characterization of methanogenic growth • Improved genome annotation	[74]
Bacillus subtilis	844	988	1020	• Gap analysis • Growth rate predictions • *In silico* knockouts • Functional assignment to genes (improved genome annotation)	[73]
Homo sapiens	1496	2712	3311	• Importance of comparison highlighted • Metabolic gap analysis • Coupled reaction sets, potential drug targets predicted	[43]
Mycobacterium tuberculosis	726	739	849	• Model integrated with experimentally determined data • *In silico* knockouts • Growth rate predictions • Online resource available for simulations	[32]
M. tuberculosis	661	828	939	• Growth rate predictions • *In silico* knockouts • Hard-coupled reaction sets identified; mapped to drug targets	[46]
B. subtilis	534	456	563	• Includes enzymatic and genetic regulation • Functional organization of genetic and metabolic regulatory networks	[81]
Aspergillus niger	871	1047	1190	• Model validated with data on yields, fluxes and transcription • Prediction of yields, intracellular distribution of carbon fluxes and physiological responses	[82]
P. aeruginosa	1056	760	883	• Genome annotation refinement • Growth predictions (on various carbon sources) • *In silico* gene essentiality	[58]
P. putida	815	888	877	• Gap analysis • Growth rate predictions, flux variability analysis • Model validated with data from continuous cell cultures, high-throughput phenotyping data, ^{13}C-measurement of internal fluxes, specifically generated knock-out mutants • Auxotrophy predictions • Potential metabolic engineering strategies from *in silico* gene deletions	[55]

clues to the identification of drug targets. Systems-level analyses such as FBA can be a very useful starting point in the 'pipeline' of drug target identification, as has been illustrated by us recently [52].

Metabolic engineering

Papin and co-workers have reported a reconstruction of *Pseudomonas putida* [55], an organism that has a proven potential in environmental and industrial biotechnological applications due to its metabolic versatility, stress resistance and amenability to genetic modifications. They have used FBA and flux variability analysis to analyse the potential of the metabolic network of the organism, as well as identifying key parameters such as growth yield, network robustness and gene essentiality. By validating the model with data from experimental cell cultures, the authors have provided a valuable framework for biotechnological applications using *P. putida*. An excellent application of systems biology in metabolic engineering, with commercial potential, has been illustrated by Stephanopoulos and co-workers, for improving lysine production in a strain of *Corynebacterium glutamicum*, by the coordinated over expression of two genes, encoding pyruvate carboxylase and aspartate kinase [9, 56]. Stephanopoulos and co-workers have also reported a genome-wide FBA of *E. coli* to discover putative genes impacting network properties and cellular phenotype, for re-engineering lycopene synthesis [8]. Targets identified using this model improve product synthesis on the basis of increased availability of metabolic precursors and cofactor balancing. For lycopene biosynthesis, a triple knock-out construct was identified that exhibited almost a 40% increase over an engineered high-producing parental strain [8].

Refinement of metabolic networks

The study of metabolic networks through FBA also finds application in the refinement of the knowledge on metabolism of an organism, as well as the reconciliation of conflicting knowledge in the literature. The 'metabolic gaps' as well as inconsistencies with experimental data that may be observed during simulations of metabolic networks may help to refine the networks and consequently, to improve the knowledge on the metabolism of an organism. Palsson and co-workers have proposed an optimization-based algorithm to predict the missing reactions required to reconcile disagreements between reconstructed metabolic networks and experiment [57]. Papin and co-workers have constructed a GSMN of *Pseudomonas aeruginosa*, also illustrating the application of FBA for 'gap analysis', to identify and resolve 'knowledge gaps' in the metabolic network [58]. A systematic network analysis has a potential to identify and potentially resolve gaps in the knowledge of metabolic networks.

Predicting novel regulatory mechanisms

Palsson and co-workers have described the analysis of an integrated model of metabolism and transcriptional regulation in *S. cerevisiae* [34]. By identifying the discrepancies between predicted growth phenotypes and experimentally observed phenotypes, which arise from missing regulatory effects in the model, they have shown that it is possible to investigate novel regulatory mechanisms. This study also highlights how modelling can direct experimentation.

CHALLENGES AND FUTURE PERSPECTIVES

Limitations of FBA

In general, the solution obtained by FBA is only as good as the constraints used to build the model [13]. Therefore, it is very important to invest a lot of time and effort in a quality reconstruction of metabolic networks, including the selection of constraints. It has been shown, for the genome-scale metabolic reconstruction of *P. putida* that the structure of the metabolic network is the critical factor in determining the accuracy of FBA predictions, and the objective function detailing biomass composition has a lesser influence [55]. FBA suffers from incomplete annotation of the proteins in a genome, although it can provide clues to enhance the current knowledge. Furthermore, FBA focuses only on part of the entire genome of an organism, involving mostly enzymes, which catalyse the various metabolic reactions in the cell. Due to the incomplete nature of annotation, several reactions may appear to have zero fluxes from FBA, since the reactions involving metabolites, downstream or upstream from these reactions may not have been characterized (metabolic gaps).

Challenges in reconstruction and analysis of GSMNs

Most reconstructions rely fundamentally on the availability of genome sequences and annotations. For organisms with low sequence homology to other organisms, such as *Plasmodium falciparum*, automated reconstructions generally result in highly incomplete GSMNs [5]. Furthermore, many organisms have unique pathways, for example, the mycolic acid pathway in *M. tuberculosis*, which need to be manually curated, through extensive literature analysis [11]. A detailed account of the challenges in the reconstruction of parasitic metabolic networks has been discussed in [59].

Biologically relevant objective functions

One of the major challenges for FBA is the definition of a biologically relevant objective function. While the maximization of biomass production has been commonly used as an objective function in the genome-scale reconstructions of several prokaryotes, phenotypes may be more accurately predicted with more biologically relevant functions, particularly in case of higher organisms. Techniques such as MoMA [15, 27] and ROOM [28], which alter the objective for optimization, resulting in QP and MILP formulations respectively, have already been discussed earlier. A recent approach towards resolving the problem of selecting a suitable objective function is a framework proposed by Papin and co-workers, known as the Biological Objective Solution Search (BOSS) [60]. In this framework, the biological objective is a new stoichiometric reaction added to the stoichiometric matrix, which is not confined to be a subset of the existing reactions. This reaction is added to the existing constraints and optimized, also minimizing the difference between the resulting flux distribution and available experimental data.

Srivastava and co-workers have described a simple Bayesian-based method to quantitatively select a single 'most probable' objective function out of a choice of plausible biologically relevant alternative objective functions, by comparison against experimental data [61]. While the method provides a useful approach to discriminate between objective functions, it also emphasizes the need for experimental flux data, which is required to compare predictions using different objective functions. It must be emphasized here, that a synergy between *in silico* simulations and biochemical experiments can indeed help in multiple ways to synthesize better models of metabolic networks.

The choice of objective function cannot be made independent of the conditions of simulation; for instance, it would not be reasonable to use a biomass maximization function to accurately predict the fluxes for an organism that is grown under starvation of nutrients. Sauer and co-workers have illustrated for *E. coli* that no single objective function was capable of predicting experimentally observed fluxes under different conditions; it was important to identify the most relevant objective for each condition [24]. Thus, the choice of objective function is quite important in the context of FBA, and it is important to choose a biologically relevant objective function. Improvement in methodology is required in two ways: first, the identification of an appropriate objective function and second, the description the chosen function at high resolution, which may require detailed large-scale quantitative experimentation under various conditions.

Impact of high-throughput experiments

The importance of integration with experimental data has already been emphasized in previous sections. With the advance in high-throughput techniques for estimating metabolomic data, it is possible to generate large amounts of data for use in FBA models; FBA can benefit from metabolomic measurements, which could aid in identifying more constraints. Furthermore, FBA can also cope with the uncertainty and incompleteness in metabolomic data, since it allows for the incorporation of partial metabolic information [62]. The metabolic 'gaps' identified through FBA can also be useful in guiding metabolomic experiments. With the advances in high-throughput ^{13}C flux analysis [63, 64], which can be applied at a genome-scale to estimate intracellular fluxes [65, 66], it is possible to generate more data for hypothesis validation, improving constraints in FBA models, as well as to aid in the choice of objective functions. Lee and co-workers have illustrated the use of even partial information obtained from the ^{13}C-labelling experiments to generate mass balance equations for 'artificial' metabolites, which are used as additional constraints during FBA [67]. The availability of genome-scale transcriptomics data can be advantageous in the integrated reconstruction of metabolic and regulatory networks.

Integration of metabolic, regulatory and signal transduction networks

The complexity of biological function arises from the concerted interplay between metabolism, regulation and signal transduction. However, till recently, most models of biological networks have focussed only on one of these networks, rather than analysing the complexity in its entirety.

Papin and co-workers have proposed an FBA–based strategy, referred to as integrated dynamic FBA (idFBA), that dynamically simulates cellular phenotypes arising from integrated networks [68]. The idFBA framework requires an integrated stoichiometric reconstruction of signalling, metabolic, and regulatory processes. A major challenge for such an integration is the fact that the various processes operate on vastly different time-scales. idFBA attempts to address this issue by including slow reactions in a time-delayed fashion. Time is discretized into small steps; at each step, an FBA is performed. An incidence matrix $\mathbf{I}_{R_s \times t_N}$ is computed, which keeps track of which of the R_s reactions are to be included at a particular time-step t of t_N, thereby accounting for the difference in time scales for the reactions. Based on the computed flux, the constraints and the incidence matrix (for the next time-step) are updated. The choice of objective function for such an integrated system is also important; idFBA utilizes BOSS [60] (described earlier) to identify objectives for the integrated system. The authors have shown the utility of idFBA for analysing a portion of the high-osmolarity glycerol response pathway in *S. cerevisiae*, generating time-course predictions comparable to an equivalent kinetic model. It appears that idFBA might serve to improve the accuracy and versatility of constraint-based analyses, at the same time avoiding the stringent requirements of kinetic parameters and detailed mechanisms, imposed by kinetic models.

Covert and co-workers have proposed another method to simultaneously model the metabolic, regulatory and signal transduction networks, by integrating FBA with regulatory Boolean logic and ordinary differential equations for describing signal transduction [69]. They have used this approach, called integrated FBA (iFBA), to create an integrated model of *E. coli*, which combines a flux-balance-based central carbon metabolic and transcriptional regulatory model [31] with an ODE-based detailed model [70] of carbohydrate uptake control. They have shown that the iFBA framework better captures the dynamics of the system, compared to either rFBA or ODE modelling paradigms. This approach is an improvement over rFBA in that the kinetic description is much more detailed; a dynamic picture of the system is obtained, rather than just the final steady state that would be obtained using rFBA. Furthermore, certain enzymes, which would never be part of a strictly optimal growth scenario, are expressed and active since they are utilized for important functions such as signal transduction, though not for their metabolic contribution to growth. Phenotypes for certain knock-outs, such as *galP* and *glk* on glucose/glucose-6-phosphate, were also much better predicted by iFBA, due to its ability to account for the subtle effects of the dynamics of internal metabolites such as glucose-6-phosphate.

Such methods provide interesting extensions to the well-established FBA paradigm, giving a greater impetus towards accurate prediction of phenotypes from models of biological networks. More accurate predictions from biological networks can be obtained only by an integration of models of metabolism, regulation and signal transduction.

CONCLUSIONS

Although it may be very desirable to have genome-scale mechanistic models of microbial systems, the lack of available metabolomic data and thermodynamic quantities has rendered the probability of achieving cell-scale kinetic models quite low [71]. Constraint-based models, particularly those using FBA, have filled in the void admirably, enabling analysis of several large systems, including entire GSMNs for prokaryotes [32, 55, 72–75], eukaryotes [10, 76, 77] and even the human [43, 44], with wide-ranging applications from metabolic engineering [8, 55, 56] to drug discovery [11]. The potential of FBA for addressing several biological problems is now well-established, as evident from the number of reports in literature (Table 2). The stage seems all set to realize the promise in obtaining biological insights of chosen sets of proteins, in a systematic manner.

However, several challenges remain in the construction and analysis of constraint-based models, particularly in terms of the accurate definition of the metabolic network, the various constraints, as well as the definition of biologically relevant objective functions. It is very likely that constraint-based models will continue to grow in popularity and a wide spectrum of objective

functions for analysis, with increasing biological relevance, will be used to enable various types of predictions on the capabilities of metabolic networks. There have already been interesting advances in the area of FBA, with the integration of regulatory information as well as signalling networks into the metabolic models. The integration of various types of models—kinetic, constraint-based and topological—to draw conclusions at various levels is another exciting challenge ahead of modelling in systems biology, which holds the key to many of the varied applications of systems biology.

Key Points

- Flux balance analysis (FBA) is a powerful tool for the constraint-based analyses of (genome-scale) metabolic networks, to identify steady state flux distributions and metabolic capabilities of biochemical networks.
- The critical steps in FBA are the reconstruction of a metabolic network, followed by mass balance, imposition of constraints, choice of a suitable (biologically relevant) objective function and (linear) optimization.
- FBA is highly versatile and various recent extensions to FBA such as the rFBA, iFBA and idFBA, to account for the interdependence of metabolic networks on transcriptional regulatory networks and signal transduction networks, have empowered FBA to make better *in silico* predictions on the phenotypes of biological systems.
- FBA has a wide variety of applications in metabolic engineering, drug discovery, construction and analysis of genome-scale metabolic models, as well as in refining the existing knowledge on biochemical/metabolic networks.

References

1. Raman K, Rajagopalan P, Chandra N. Principles and practices of pathway modelling. *Curr Bioinformatics* 2006;**1**: 147–60.
2. Stelling J. Mathematical models in microbial systems biology. *Curr Opin Microbiol* 2004;**7**:513–8.
3. Reed JL, Famili I, Thiele I, *et al*. Towards multidimensional genome annotation. *Nat Rev Genet* 2006;**7**:130–41.
4. Price ND, Reed JL, Palsson BØ. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat Rev Microbiol* 2004;**2**:886–97.
5. Feist AM, Herrgard MJ, Thiele I, *et al*. Reconstruction of biochemical networks in microorganisms. *Nat Rev Microbiol* 2009;**7**:129–43.
6. Covert MW, Schilling CH, Famili I, *et al*. Metabolic modeling of microbial strains *in silico*. *Trends Biochemical Sci* 2001;**26**:179–86.
7. Burgard AP, Pharkya P, Maranas CD. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng* 2003;**84**:647–57.
8. Alper H, Jin Y-S, Moxley JF, *et al*. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab Eng* 2005;**7**:155–64.
9. Koffas MA, Jung GY, Stephanopoulos G. Engineering metabolism and product formation in Corynebacterium glutamicum by coordinated gene overexpression. *Metab Eng* 2003;**5**:32–41.
10. Duarte NC, Herrgard MJ, Palsson BØ. Reconstruction and validation of *Saccharomyces cerevisiae iND750*, a fully compartmentalized genome-scale metabolic model. *Genome Res* 2004;**14**:1298–309.
11. Raman K, Rajagopalan P, Chandra N. Flux balance analysis of mycolic acid pathway: targets for anti-tubercular drugs. *PLoS Comput Biol* 2005;**1**:e46.
12. Covert MW, Reed JL, Knight EM, *et al*. Integrating high-throughput and computational data elucidates bacterial networks. *Nature* 2004;**429**:92–6.
13. Kauffman KJ, Prakash P, Edwards JS. Advances in flux balance analysis. *Curr Opin Biotechnol* 2003;**14**:491–6.
14. Bonarius HPJ, Schmid G, Tramper J. Flux analysis of underdetermined metabolic networks: the quest for the missing constraints. *Trends Biotechnol* 1997;**15**:308–14.
15. Segre D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci USA* 2002;**99**:15112–7.
16. Palsson BØ. *Systems Biology-Properties of Reconstructed Networks*. New York: Cambridge University Press, 2006.
17. Edwards JS, Ibarra RU, Palsson BØ. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat Biotechnol* 2001;**19**:125–30.
18. Deutscher D, Meilijson I, Kupiec M, *et al*. Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nat Genet* 2006;**38**:993–8.
19. Ibarra RU, Edwards JS, Palsson BO. Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* 2002;**420**:186–9.
20. Burgard AP, Maranas CD. Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnol Bioeng* 2003;**82**:670–7.
21. Ramakrishna R, Edwards JS, McCulloch A, *et al*. Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints. *Am J Physiol Regul Integr Comp Physiol* 2001;**280**: R695–704.
22. Vo TD, Greenberg HJ, Palsson BO. Reconstruction and functional characterization of the human mitochondrial metabolic network based on proteomic and biochemical data. *J Biol Chem* 2004;**279**:39532–40.
23. Edwards JS, Palsson BO. Robustness analysis of the Escherichia coli metabolic network. *Biotechnol Prog* 2000; **16**:927–39.
24. Schuetz R, Kuepfer L, Sauer U. Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Mol Syst Biol* 2007;**3**:119.
25. Becker SA, Feist AM, Mo ML, *et al*. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat Protocols* 2007;**2**:727–38.
26. Papin JA, Price ND, Edwards JS, *et al*. The genome-scale metabolic extreme Pathway structure in *Haemophilus influenzae* shows significant network redundancy. *J Theor Biol* 2002;**215**:67–82.
27. Segre D, Zucker J, Katz J, *et al*. From annotated genomes to metabolic flux models and kinetic parameter fitting. *OMICS* 2003;**7**:301–16.

28. Shlomi T, Berkman O, Ruppin E. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc Natl Acad Sci USA* 2005;**102**:7695–700.
29. Papp B, Pal C, Hurst LD. Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature* 2004;**429**:661–4.
30. Harrison R, Papp B, Pal C, *et al*. Plasticity of genetic interactions in metabolic networks of yeast. *Proc Natl Acad Sci USA* 2007;**104**:2307–12.
31. Covert MW, Palsson BO. Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *J Biol Chem* 2002;**277**:28058–64.
32. Beste DJV, Hooper T, Stewart G, *et al*. GSMN-TB: a web-based genome-scale network model of Mycobacterium tuberculosis metabolism. *Genome Biol* 2007;**8**:R89.
33. Covert MW, Schilling CH, Palsson BØ. Regulation of gene expression in flux balance models of metabolism. *J Theor Biol* 2001;**213**:73–88.
34. Herrgard MJ, Lee BS, Portnoy V, *et al*. Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in Saccharomyces cerevisiae. *Genome Res* 2006; **16**:627–35.
35. Covert MW, Palsson BØ. Constraints-based models: regulation of gene expression reduces the steady-state solution space. *J Theor Biol* 2003;**221**:309–25.
36. Shlomi T, Eisenberg Y, Sharan R, *et al*. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Mol Syst Biol* 2007;**3**:101.
37. Beard DA, Liang SD, Qian H. Energy balance for analysis of complex metabolic networks. *Biophys J* 2002;**83**: 79–86.
38. Edwards JS, Palsson BØ. The *Escherichia coli* MG1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities. *Proc Natl Acad Sci USA* 2000;**97**:5528–33.
39. Hoppe A, Hoffmann S, Holzhutter HG. Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks. *BMC Syst Biol* 2007;**1**:23.
40. Reed JL, Vo TD, Schilling CH, *et al*. An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR). *Genome Biol* 2003;**4**:R54.
41. Pharkya P, Burgard AP, Maranas CD. Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock. *Biotechnol Bioeng* 2003;**84**:887–99.
42. Nagrath D, Avila-Elchiver M, Berthiaume F, *et al*. Integrated energy and flux balance based multiobjective framework for large-scale metabolic networks. *Ann Biomed Eng* 2007;**35**:863–85.
43. Duarte NC, Becker SA, Jamshidi N, *et al*. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc Natl Acad Sci USA* 2007; **104**:1777–82.
44. Ma H-W, Sorokin A, Mazein A, *et al*. The Edinburgh human metabolic network reconstruction and its functional analysis. *Mol Syst Biol* 2007;**3**:135.
45. Förster J, Famili I, Palsson BO, *et al*. Large-scale evaluation of in silico gene deletions in *Saccharomyces cerevisiae*. *OMICS* 2003;**7**:193–202.
46. Jamshidi N, Palsson BØ. Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the *in silico* strain *iNJ*661 and proposing alternative drug targets. *BMC Syst Biol* 2007;**1**:26.
47. Schuster S, Dandekar T, Fell DA. Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol* 1999;**17**:53–60.
48. Schilling CH, Letscher D, Palsson BO. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *J Theor Biol* 2000;**203**:229–48.
49. Papin JA, Stelling J, Price ND, *et al*. Comparison of network-based pathway analysis methods. *Trends Biotechnol* 2004;**22**:400–5.
50. Burgard AP, Nikolaev EV, Schilling CH, *et al*. Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Res* 2004;**14**:301–12.
51. Joyce AR, Palsson BO. Predicting gene essentiality using genome-scale in silico models. *Methods Mol Biol* 2008;**416**: 433–57.
52. Raman K, Kalidas Y, Chandra N. targetTB: a target identification pipeline for Mycobacterium tuberculosis through an interactome, reactome and genome-scale structural analysis. *BMC Syst Biol* 2008;**2**:109.
53. Raman K, Rajagopalan P, Chandra N. Hallmarks of mycolic acid biosynthesis: a comparative genomics study. *Proteins* 2007;**69**:358–68.
54. Gratraud P, Surolia N, Besra GS, *et al*. Antimycobacterial activity and mechanism of action of NAS-91. *Antimicrob Agents Chemother* 2008;**52**:1162–6.
55. Puchalka J, Oberhardt MA, Godinho M, *et al*. Genome-scale reconstruction and analysis of the Pseudomonas putida KT2440 metabolic network facilitates applications in biotechnology. *PLoS Comput Biol* 2008;**4**:e1000210.
56. Koffas M, Stephanopoulos G. Strain improvement by metabolic engineering: lysine production as a case study for systems biology. *Curr Opin Biotechnol* 2005;**16**:361–6.
57. Reed JL, Patel TR, Chen KH, *et al*. Systems approach to refining genome annotation. *Proc Natl Acad Sci USA* 2006; **103**:17480–4.
58. Oberhardt MA, Puchalka J, Fryer KE, *et al*. Genome-scale metabolic network analysis of the opportunistic pathogen Pseudomonas aeruginosa PAO1. *J Bacteriol* 2008; **190**:2790–803.
59. Pinney JW, Papp B, Hyland C, *et al*. Metabolic reconstruction and analysis for parasite genomes. *Trends Parasitol* 2007; **23**:548–54.
60. Gianchandani EP, Oberhardt MA, Burgard AP, *et al*. Predicting biological system objectives de novo from internal state measurements. *BMC Bioinformatics* 2008;**9**:43.
61. Knorr AL, Jain R, Srivastava R. Bayesian-based selection of metabolic objective functions. *Bioinformatics* 2007;**23**: 351–7.
62. Lee JM, Gianchandani EP, Papin JA. Flux balance analysis in the era of metabolomics. *Brief Bioinform* 2006;**7**: 140–50.
63. Sauer U. High-throughput phenomics: experimental methods for mapping fluxomes. *Curr Opin Biotechnol* 2004; **15**:58–63.

64. Sauer U. Metabolic networks in motion: 13C-based flux analysis. *Mol Syst Biol* 2006;**2**:62.

65. Fischer E, Sauer U. Large-scale in vivo flux analysis shows rigidity and suboptimal performance of Bacillus subtilis metabolism. *Nat Genet* 2005;**37**:636–40.

66. Blank LM, Kuepfer L, Sauer U. Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol* 2005;**6**: R49.

67. Choi HS, Kim TY, Lee DY, *et al*. Incorporating metabolic flux ratios into constraint-based flux analysis by using artificial metabolites and converging ratio determinants. *J Biotechnol* 2007;**129**:696–705.

68. Lee JM, Gianchandani EP, Eddy JA, *et al*. Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Comput Biol* 2008;**4**:e1000086.

69. Covert MW, Xiao N, Chen TJ, *et al*. Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli. *Bioinformatics* 2008;**24**:2044–50.

70. Kremling A, Bettenbrock K, Gilles ED. Analysis of global control of Escherichia coli carbohydrate uptake. *BMC Syst Biol* 2007;**1**:42.

71. Jamshidi N, Palsson BO. Formulating genome-scale kinetic models in the post-genome era. *Mol Syst Biol* 2008;**4**:171.

72. Feist AM, Henry CS, Reed JL, *et al*. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 2007;**3**:121.

73. Oh Y-K, Palsson BØ, Park SM, *et al*. Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data. *J Biol Chem* 2007;**282**:28791–9.

74. Feist AM, Scholten JCM, Palsson BØ, *et al*. Modeling methanogenesis with a genome-scale metabolic reconstruction of *Methanosarcina barkeri*. *Mol Syst Biol* 2006;**2**, doi:10.1038/msb4100046 [Epub ahead of print 31 January 2006].

75. Becker SA, Palsson BØ. Genome-scale reconstruction of the metabolic network in *Staphylococcus aureus* N315: an initial draft to the two-dimensional annotation. *BMC Microbiol* 2005;**5**:8.

76. Famili I, Förster J, Nielsen J, *et al*. *Saccharomyces cerevisiae* phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proc Natl Acad Sci USA* 2003;**100**:13134–9.

77. Förster J, Famili I, Fu P, *et al*. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res* 2003;**13**:244–53.

78. Edwards JS, Palsson BØ. Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *J Biol Chem* 1999;**274**:17410–6.

79. Schilling CH, Palsson BØ. Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis. *J Theor Biol* 2000;**203**:249–83.

80. Schilling CH, Covert MW, Famili I, *et al*. Genome-scale metabolic model of Helicobacter pylori 26695. *J Bacteriol* 2002;**184**:4582–93.

81. Goelzer A, Brikci FB, Martin-Verstraete I, *et al*. Reconstruction and analysis of the genetic and metabolic regulatory networks of the central metabolism of *Bacillus subtilis*. *BMC Syst Biol* 2008;**2**:20.

82. Andersen MR, Nielsen ML, Nielsen J. Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. *Mol Syst Biol* 2008;**4**:178.