This article has been accepted for the publication in Computers in Biology and Medicine, Volume 88, 1 September 2017, Pages 150-160, https://doi.org/10.1016/j.compbiomed.2017.07.005

Computational modelling of genome-scale metabolic networks and its application to CHO cell cultures

Živa Rejc [☆a], Lidija Magdevska [☆b], Tilen Tršelič [☆a], Timotej Osolin[b], Rok Vodopivec[b], Jakob Mraz[c], Eva Pavliha[c], Nikolaj Zimic[b], Tanja Cvitanović[d], Damjana Rozman[d], Miha Moškon[b,*], Miha Mraz[b]

[a] *Faculty of Chemistry and Chemical Technology, University of Ljubljana, Ljubljana, Slovenia*
[b] *Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, Slovenia*
[c] *Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia*
[d] *Centre for Functional Genomics and Bio-Chips, Institute of Biochemistry, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia*

Abstract

Genome-scale metabolic models (GEMs) have become increasingly important in recent years. Currently, GEMs are the most accurate *in silico* representation of the genotype-phenotype link. They allow us to study complex networks from the systems perspective. Their application may drastically reduce the amount of experimental and clinical work, improve diagnostic tools and increase our understanding of complex biological phenomena. GEMs have also demonstrated high potential for the optimisation of bio-based production of recombinant proteins.

Herein, we review the basic concepts, methods, resources and software tools used for the reconstruction and application of GEMs. We overview the evolution of the modelling efforts devoted to the metabolism of Chinese Hamster Ovary (CHO) cells. We present a case study on CHO cell metabolism under different amino acid depletions. This leads us to the identification of the most influential as well as essential amino acids in selected CHO cell lines.

Keywords: metabolic networks, genome-scale metabolic models, Chinese

[☆] Authors contributed equally to this work.
[*] Corresponding author
Email address: miha.moskon@fri.uni-lj.si (Miha Moškon)

hamster ovary cells, flux balance analysis, modelling and analysis

1. Introduction

Ongoing development of systems biology in recent years derives mainly from the successful integration of many computational approaches into the experimental work. One of the most successful applications of computer science in biology is in the annotation of genomes from the vast data generated by DNA sequencing experiments. Here, the computational approaches have been essential for the analysis of large amounts of sequenced data as well as for their presentation and applications. The success stories of the annotation of genomes of different simple organisms [1] as well as of the human genome [2] were followed by the establishment of the first *genome-scale metabolic models* (GEM) [3].

GEMs are the most accurate *in silico* representation of the genotype-phenotype link [4]. These models are continously improved with the accuracy of their descriptions as well as the strength of the predictions they make. For example, Recon, the GEM describing human metabolism, was first published in 2007 but it has now gone through six iterations of improvements [5, 6]. These improvements are a consequence of the evolution of experimental and computational approaches used in systems biology. They result from the publicly available large scale data through the literature and through different general as well as specific web databases, such as KEGG (Kyoto Encyclopaedia of Genes and Genomes)[7], BRENDA (BRaunschweig ENzyme DAtabase)[8] and BioCyc [9]. Publicly available computational models through databases such as BioModels [10] and BiGG Models [11] present an additional driving force for exchangeability of knowledge and computational tools. The main motivation towards the continuous development of GEMs is a vast scope of their applications. These range from (1) the design and optimisation of environmentally friendly bio-based production of fine chemicals with simple organisms [12], and so called genome-scale synthetic biology [13] to (2) the optimisation of biopharmaceutical manufacturing cell lines and processes in non-mammalian cells, such as *Pichia pastoris* [14], as well as in mammalian cells [15], such as Chinese Hamster Ovary (CHO) cells [16], and finally to (3) the identification and analysis of possible biomarkers of complex diseases, such as non-alcoholic fatty liver disease (NAFLD) [17, 18, 19] and cancer [20, 21].

Computational models can be used to identify the segments that need to be explained more accurately to obtain a valid representation of the system's response. This allows us to systematically increase the knowledge describing biological mechanisms governing complex networks. Integration of data obtained from experiments, literature and databases into metabolic computational models can be described with a circular iteration scheme of knowledge acquisition and model improvements as shown in Fig. 1. It consists of (1) data acquisition and refinement through experimental work, literature and publicly available databases; (2) establishment and optimisation of computational models using the acquired data; and (3) analysis and validation of computational models and their potential refinement through another iteration of the cycle. The computational approaches are essential in all three steps described in the scheme. Novel computational approaches, which can be used in the reconstruction, analysis, refinement and visualisation of metabolic models, are therefore vital for the continuous progress of systems biology.

Numerous computational methods are available in the field of metabolic modelling and analysis. Majority of these are derived from the *constraint-based analysis*. The development and application of these methods is driven by the publicly available toolboxes, such as Pathway Tools [22], RAVEN (Reconstruction, Analysis, and Visualisation of mEtabolic Networks)[23] and probably the most popular COBRA (COnstraint-Based Reconstruction and Analysis) toolbox [24, 25]. These toolboxes implement the majority of the available computational methods. They follow open source concepts and are easy to update with novel methods. Computational methods applied to the analysis of metabolic networks include basic analyses, which can predict the reaction fluxes that bring the network to its optimal state (for example flux balance analysis - FBA)[26]. Implemented methods can be used to tailor the metabolic model with a specific context (see for example [27, 28]), and automatic identification of reactions, which need to be blocked in order to achieve the optimal state of the metabolic network, for example the state in which the production of selected metabolite is optimal [29, 30]. Large attention has also been devoted to the development of different visualisation approaches (see for example Escher [31]). Visualisation is, however, still mostly performed manually (see for example ReconMap [32] for the visualisation of human metabolism model).

In the following chapters the review of the state-of-the-art methods for the analysis, reconstruction and visualisation of metabolic networks is described. We begin with the description of some general approaches for the modelling

Figure 1: Circular iteration scheme describing the integration of experimental work with computational models. Data acquisition through the execution of new experiments, application of publicly available databases and literature is followed by the establishment and optimisation of computational models. These have to be sufficiently validated before their practical application. Validation is usually successful only after several iterations of model refinements and extensions. These usually require additional experiments and acquisition of supplementary data.

and analysis of biological systems with the emphasis on the GEMs (see Section 2). Furthermore, we describe the most comprehensive publicly available databases containing large experimental datasets and computational models (see Section 3). We comment on the approaches that can be used in the process of the reconstruction and visualisation of GEMs (see Sections 4 and 5). We overview the progress in the development of CHO GEMs in recent years (see Section 6). We demonstrate the application of selected computational methods on the analysis of the most recent and most complete CHO GEM, i.e. iCHO1766 [33] (see Section 7).

2. Constraint-based methods for the analysis of metabolic networks

Numerous computational methods have been developed for the computational reconstruction and analysis of metabolic networks in recent years. Most of these approaches have been integrated within different publicly available computational toolboxes, such as COBRA [24, 25] and RAVEN [23].

Computational analysis of molecular networks is usually performed on the basis of their stoichiometric description [34]. Here, each reaction is described with its *stoichiometric coefficients* [35]. Stoichiometric description can be used to establish a set of ordinary differential equations (ODEs), which are solved numerically to compute the changes in metabolite concentration over time. *Dynamical approaches*, such as numerical integration of ODEs, can be applied only when reaction kinetics and parameters are known for the whole system. This is usually not the case for the large-scale metabolic networks. Steady-state assumption can be applied to the analysis of the organisms living in a constant environment that are in exponential or log phase of growth [26, 36]. This assumption transforms the systems of ODEs to a system of linear equations with infinite solutions. The solution space can be reduced to a single solution with the flux balance analysis (FBA) [26].

FBA defines the optimal flux through the observed reactions with additional constraints. These constraints define the lower and the upper bound of the reaction fluxes. They are derived from different properties of the system, such as the availability of enzymes catalysing observed metabolic reactions and the reversibility of reactions. Additional information can be encoded within the flux boundaries. These include cell line specific parameters and properties of its environment, compartmentalisation of the metabolic reactions and their regulation by signalling or gene regulatory networks [37, 38].

The optimal flux is evaluated on the basis of the optimisation criterion. The optimization criterion defines a biological objective that is relevant to the problem being studied [26]. It is modelled as a linear function, which describes the reactions' contributions to the specific phenotype of the observed system. Examples of optimisation criteria are maximal cell growth or maximal production of a recombinant protein [26]. The optimal flux distribution within the reconstructed metabolic network can be acquired with the maximization (sometimes also minimization) of this criterion. The constraints and the objective criterion constitute the general form of a linear program. This can be solved as linear programming problem [37]. An example of an FBA application to the analysis of a metabolic network can be found in section 7.

Several extensions of FBA have been developed in recent years. These include dynamic flux balance analysis (dFBA) [39] and miniFBA [40], which incorporate rate of change of flux constraints. Regulatory flux balance analysis (rFBA) [41] incorporates regulatory constraint in the analysis. Another example is parsimonious or parsimonious enzyme usage FBA (pFBA), which classifies genes according to their contribution to the optimal solution [42]. The repertoire of constraint-based methods currently includes more than 100 methods. Further information regarding other extensions of the FBA method can be found in references [30], [43] and [44].

3. Biological databases

Large experimental datasets together with different computational models have been made available in the form of publicly available databases in recent years. Some of these cover metabolic pathways for different organisms (e.g. KEGG [7] and MetaCyc [9]), while others focus on experimental data or computational implementations of different metabolic models (e.g. BiGG Models [45] and BioModels [10]). Table 1 lists some of the commonly used databases in the field of GEMs reconstruction and analysis, and describes their intended purpose. Databases have a significant impact on the analysis of biological systems. They provide data for the reconstruction, improvement and validation of metabolic networks, and serve as a repository for the storage of knowledge that has been obtained with different experimental and computational approaches. Data that are deposited in the form of database entries need to fulfil several formal and informal standards. This contributes to both human and computer readability. Most of the databases listed in

Database	Purpose	Comments
BiGG models [45]	genome-scale metabolic models	collaboration with other databases (KEGG, PubChem, etc.)
KEGG [7]	general purpose	support for KGML (KEGG Markup Langage), which enables model visualisation
BioCyc [9]	general purpose	set of specific databases (e.g. HumanCyc database for human metabolic pathways)
MetaCyc [9]	metabolic pathways	part of BioCyc database focused on metabolic networks
BioModels [10]	models of biological processes	comprehensive database, which includes models of metabolic and non-metabolic pathways and networks

Table 1: Publicly available databases that are commonly used for the genome-scale metabolic models reconstruction and analysis. General purpose databases include a vast scope of different categories, such as metabolic pathways, reactions, enzymes and drugs.

Table 1 also provide an Application Programming Interface (API), which allows a straightforward access to provided data to different computational tools developed globally by the scientific community.

4. Reconstruction of genome-scale metabolic models

GEMs systematically incorporate multi-omic data into an unified representation [46, 47, 48]. These data are often referred to as BiGG (Biochemical, Genetic and Genomic) data and represent the metabolic network of a specific organism [38]. GEM reconstruction has to describe every enzyme and its corresponding metabolic reactions within the metabolic network. The reconstruction must contain information about (1) substrates and products of each enzymatic reaction, (2) stoichiometric coefficients, (3) reaction directionalities, and (4) their compartmentalisation [49]. GEMs must also include additional reactions that are not present in the observed metabolic network per se, but describe the transfer of metabolites between different compartments in the model as well as transfer of metabolites between the cell and its external media (so called *exchange reactions*) [38].

Before the availability of genome annotations, reconstructions were performed solely with experimental data, research literature and biochemical characterization of enzymes. These data, however, rarely describe the enzymatic behaviour in a living organism accurately. Modern reconstruction protocols on the other hand include four main steps: (1) genome annotation, (2) automated metabolic network reconstruction, (3) metabolic network re-

finement and, (4) evaluation and validation of metabolic reconstruction with additional experimental data [38, 46, 47, 49].

4.1. Genome annotation

Annotation defines biological functions of certain parts of genome sequences and classifies them into standardized gene ontology classes [38, 47]. Automated gene search algorithms are used to annotate a specific genome sequence. These algorithms often search for sequence homology to other, already defined and characterized genes, and consequently proteins [50]. Annotations are later verified manually with the goal to define unknown genes and their functions. This is achieved with different sequence alignment tools such as BLAST and other phylogenetic grouping tools [47]. Annotations with low confidence values can be verified by performing additional experiments [38]. Many annotated genomes are publicly available in the web databases, such as CHOGenome [51].

4.2. Automated metabolic network reconstruction

Using annotated genome as an input, tools like Pathway tools [22] and metaSHARK [52] can automatically generate GEMs [38], which can be represented by a list of genes, reactions and enzymes [47]. Furthermore, cell-line- and tissue-specific metabolic models can be obtained from GEMs using model extraction methods, such as GIMME [27]. Two recent surveys of reconstruction and model extraction methods together with their systematic evaluation are available in [53] and [28]. Even though these algorithms are able to drastically enhance the reconstruction process, the reconstructed networks may still contain numerous errors. They may include genes that do not participate in a given metabolic pathway, while omit genes that present important connections between metabolic pathways.

4.3. Metabolic network refinements

Manual refinements of the reconstructed network need to be performed additionally to obtain a valid *in silico* representation of metabolic networks. Automatically generated GEMs have to be refined with the potential corrections of reaction stoichiometries, thermodynamics and energy constraints [38, 49, 50]. Refinements are performed in a systematic order. A confidence value is assigned to every gene in the network [47]. This is followed by the reaction analysis, which is the most time consuming and critical process. Its main goal is to verify the occurrence and the parameters describing specific

reactions within the cell. In order to obtain a high quality metabolic reconstruction, organism compartmentalisation and exchange reactions, gene processing, and thermodynamic and biomass accuracy are being emphasised [38, 47, 48].

4.4. Evaluation and validation of metabolic reconstruction

Validation methods compare the experimental results with the data obtained from the metabolic reconstruction. Metabolic reconstructions can be to some extent validated with additional *in vivo* and *in vitro* experiments. Here, we are often limited to the observation of phenotype and cell growth. Phenotype of cells can be observed in different conditions. The observations are performed together with additional measurements of, e.g. energy sources uptake and secretion, and excretion of metabolic products into the media [47]. Modern approaches are able to asses the distribution of metabolites and estimate the reaction fluxes *in vivo* using methods, such as fluorescent or 13C labelling [36, 47]. There are many studies that try to validate the metabolic reconstruction making comparison among *in silico* predictions and *in vivo* results of specific gene knock-outs [38, 50]. GEMs can also be validated by observing if they correctly predict known auxotrophic features of the cell as well. In this case the organic compounds that the organism is able to synthetize by its own are observed. For example, a known feature of CHO cell lines is that they are unable to synthetize cysteine and arginine on their own, thus making these two amino acids essential [33]. Another example of computational approaches that can be applied during the process of validation are *gap filling algorithms*. Missing metabolic functions are usually caused by blocked metabolic reactions, which are caused by missing metabolic reactions, i.e. *gaps*. These gaps may be to some extent filled in automatically with algorithms, such as fastGapFill [54], Meneco [55] and GAUGE [56].

5. Model visualisation

Visualisation of metabolic networks is important for the interpretation and understanding of their composition and comparison with similar networks. Visualisation is included within the majority of the existing pathway databases such as KEGG [7]. Visualised metabolic pathways and networks are however mostly manually drawn and stored in a static form [57]. During data updates, these images have to be modified manually. Moreover, manual

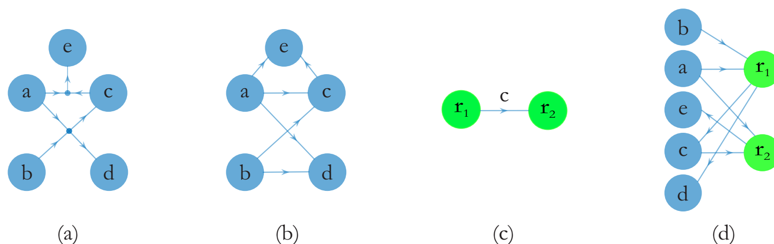

Figure 2: Different visualisation approaches for reactions r_1: a+b $\rightarrow$ c + d and r_2: a + c $\rightarrow$ e. Fig. (a) illustrates the approach where the metabolites are visualised as vertices and reactions as directed hyperedges. In Fig. (b) hyperedges are transformed to regular edges. Here, each reaction is represented with a set of edges. Every substrate in a reaction has a directed edge to each of the products. Fig. (c) illustrates the approach where each vertex represents one reaction, and the vertices are linked when two reactions share an intermediate. Fig. (d) illustrates the bipartite approach where the first type of vertices represents the reactions and the second type the metabolites. Blue vertices represent the metabolites and green vertices the reactions.

visualisation of new metabolic pathways is extremely time consuming. Several efforts to automate or at least aid the visualisation of metabolic networks emerged in the last twenty years. Some of the most commonly used tools that support the (semi-)automatic visualisation are presented in Table 2.

Computer aided visualisation mostly relies on a directed graph (digraph) representation of metabolic networks. Here, the vertices correspond to metabolites and the edges to reactions. When reactions have more than one reactants and/or products these have to be visualised as hyperedges (see Fig. 2(a)) [57]. Alternative approaches transform all hyperedges to regular edges. Reactions may be decomposed into multiple edges, i.e. one for each reactant (see Fig. 2(b)). Another approach depicts reactions as vertices and metabolites as edges. The vertices are linked when two reactions share an intermediate (see Fig. 2(c)). Algorithms, such as path-finding algorithms, are hard to be used on such graphs. A more complex but effective approach is to use a bipartite graph representation, where reactions are represented with different types of vertices than metabolites (see Fig. 2(d)) [58].

Different tools for automatic visualisation of GEMs have been developed recently (see Table 2). These include Cytoscape [59] in a combination with different plugins, such as CySBML [60], and MetDraw [61]. Certain aesthetic criteria are expected from the automatic visualisation of metabolic pathways. Some criteria are common in graphical representation of graphs, such as

Tool name	Automated visualisation	Supported formats	Comments
Cytoscape and CySBML [59, 60]	Yes	SIF, NNF, GML, XGMNL, SBML, BioPAX	general purpose tool for biomolecular networks, many algorithms, wide array of settings
Escher [31]	No	COBRA JSON	manual adding of vertices and edges, dynamic analysis in the background
MetDraw [61]	Yes	SBML	simple visualisation of SBML models, use of vector images
Paint4Net [63]	Yes	COBRA JSON	immediate COBRA file visualisation, semi-automatic removal of redundant elements

Table 2: Computational tools for the visualisation of genome-scale metabolic models.

minimization of intersecting edges [62], maximum symmetry and minimum drawing surface [57]. Other criteria are specific to metabolic pathways and arise from the conventions used in biochemistry textbooks, e.g. directionality of hierarchic components [62]. These render most of the graph visualisation algorithms unsuitable for the purpose.

Manual and semi-automatic visualisation tools are therefore still used and developed nowadays. For example, Escher [31] does not support full automatic visualisation, but aids the user in the manual construction of the network visualisation. The tool also supports the interaction with constraint-based methods that run in the background. This allows the users to perform perturbations on the network graphically and interactively observe their consequences.

6. Computational modelling of CHO metabolism

CHO cells have become prevalent in the production of recombinant proteins for clinical applications [64]. These proteins should be therapeutically active, human-compatible, and target-specific [65]. In contrast to bacterial or yeast cells, mammalian cells are able to provide the proper protein folding, assembly and post-translational modifications, which are necessary in order to achieve high quality products [16]. In the last decades CHO cells have been widely applied to the production of biopharmaceuticals such as monoclonal antibodies, hormones, cytokines, and blood coagulation factors. Large attention has been devoted to the development of computational models of different CHO cell lines to improve the quality of biopharmaceuticals (therapeutics efficiency, specific targeting and immunogenicity), and to increase the product yield.

Figure 3: The evolution of CHO metabolic models. The complexity and accuracy of metabolic models has been increasing ever since the publication of the first CHO metabolic model in 1999. Colours of the nodes represent the models' types: constrained based approaches are coloured green and dynamical approaches red. Sizes of the nodes represent the complexity of metabolic models that range from approximately ten to a few thousand of metabolic reactions.

6.1. Overview of CHO models

The focus of the first computational model of CHO metabolism was the analysis of central carbon metabolism [66]. This model was introduced in 1999. The progress of CHO metabolic models in the context of their biological relevance and accuracy has been continuing ever since. We illustrate the evolution of CHO metabolic models in Fig. 3.

The majority of the models were developed for the optimisation of a specific segment of bioproduction. Different models were introduced to analyse the influence of media composition, e.g., to determine key metabolites linked to the accumulation of adverse products [67] or to analyse the effects of amino acid composition of media on the growth and productivity of CHO cells [68, 69]. Models were used to analyse the metabolic switch from lactate production to lactate uptake [70], to study the differences in the metabolism at growth and non-growth phases [71, 72], and to quantify intracellular metabolic fluxes of producing CHO cells supplemented with different lots of wheat hydrolysates [73]. The spatial localisation of metabolic reactions was also taken into account to predict the medium composition on CHO cells growth more accurately [74, 75].

Most of the reported computational approaches, simplify the whole me-

tabolism to specific metabolic pathways, such as central carbon metabolism [66, 76, 77], or lump the pathways to a small set of reactions which prove to significantly affect the analysed segment of the system [78]. Constraint-based approaches based on FBA and its derivations have enabled us to reconstruct large-scale metabolic models even in the case of missing quantitative data [64]. These approaches still require genome annotation data to perform the GEM reconstruction. The CHO cells genome was not annotated until 2011 [79]. The first GEM model of CHO cells was reconstructed in 2012 by using the mouse GEM as template [67]. Recently, several research groups combined efforts to construct the first consensus GEM of CHO cell metabolism, namely *iCHO1766* [33]. We describe this model in more details in section 6.2.

CHO GEMs are not only used to predict the cell growth and the yield of recombinant protein production, but can also be used to determine the differences between different cell lines. Yusufi et al. identified the changes in metabolism between the wild and the recombinant antibody-producing CHO cell lines with the integration of transcriptomics data into the iCHO1766 GEM [80]. They used the upgraded models to analyse the differences in metabolic pathways and to identify the *transcriptional hotspots*. Moreover, omics-based GEMs can be used to discover new potential biomarkers for selecting CHO clones with stable genomes and high productivity, and thus to improve current selection methods such as DHFR amplification process [80].

Even though CHO GEMs have a vast scope of successful applications in the field of biopharmaceutical production, there are still several drawbacks that need to be addressed in the future [33]. (1) GEMs need to describe compartmentalisation of metabolic reactions more accurately, whereas reaction directionalities may differ in different compartments [67]. Moreover, the exchange reactions between compartments can play a vital role in the cell metabolism and can even present a rate-limiting step. (2) The natural selection and genetic drift need to be regarded. The accuracy of the models can be strongly diminished by the accumulated mutations in the cell lines. (3) Accurate measurements of metabolite concentrations are required in order to predict accurate protein yield and cell growth. Due to large amount of different metabolites within various compartments, such measurements are yet to be improved. (4) We are still unable to completely understand the mammalian metabolism. Various connections between metabolic pathways that may be vital are consequently still missing. (5) Biomass objective functions are based on experimental measurements and usually presume static

environment. In reality this is not the case as biomass composition differs under different conditions. The models are also currently unable to predict the dynamics of the cell growth. (6) The models are not able to account for the influence of inhibitory metabolites (such as lactate or ammonia in the cell medium) on the cell growth. (7) Finally, the metabolism is just a part of the whole system that controls the production of recombinant proteins. GEMs would need to be integrated with other models, such as secretory pathways and glycosylation models. A lot of effort has been devoted to the construction of such models, but none have been sufficiently integrated with CHO GEMs up to date [33].

6.2. iCHO1766 model

The iCHO1766 model was reconstructed on the basis of the annotation of *Cricetulus griseus* (the Chinese hamster) genome [79]. The model represents to date the most accurate computational prediction of the growth rate and the rate of recombinant protein production. The whole model and its reconstruction together with the reference data for genes, reactions and metabolites is available in the BiGG database [45].

The establishment of the model followed a community approach. Different research groups independently reconstructed GEMs, which they integrated into a consensus model. They established initial models on the basis of experimental results (transcriptomic and proteomic data) and literature data [67, 71, 81, 82, 83]. Models were also partially derived from the GEM Recon [5, 6, 84, 85] on the basis of CHO homologies to human genes. The community built three cell line specific GEMs, namely for CHO-K1, -S and -DG44 producing cell lines, with the GIMME (Gene Inactivity Moderated by Metabolism and Expression) algorithm [27]. They generated CHO-K1 and -S models with the RNA sequencing and proteomic data. Microarray data were used for the establishment of CHO-DG44 model [33].

In order to accurately describe the cell growth and protein production *in silico*, it is necessary to determine the relative amounts of metabolites needed by the cell to synthesise all cellular components and recombinant proteins. These data were obtained from Feist et al. [86]. The literature data for recombinant erythropoietin (EPO) and immunoglobulin G (IgG) production were used to find the difference between calculated values in non-producing cell lines and measured values for IgG-producing hybridoma cell lines. Thereafter, two biomass reactions were formulated and included into the model [33].

The iCHO1766 computational models accurately predict the growth phenotypes and known auxotrophies. They are able to predict the protein synthesis capacities and to quantify a potential increase in product yield after specific perturbations are introduced. They were used to show that the cell engineering may affect the product yield more significantly than bioprocessing treatments. Furthermore, the cell growth and product yield were shown not to be proportionally related [33].

7. Case study: computational analysis of iCHO1766 model

We can use the FBA and its alternatives to predict the optimal cell growth and optimal product formation in dependency of different conditions. These include cell culture media composition and activity of enzymes catalysing the metabolic reactions within the network. Here we demonstrate the application of constraint-based approaches on three different analyses of iCHO1766 model, namely (1) basic FBA of metabolic network to assess the reference state of the network, (2) qualitative perturbation analysis to assess the activity of selected biosynthesis reactions before and after the perturbations, and (3) medium analysis to assess the effects of amino acid medium composition on cell growth. We performed the analyses on the CHO-S cell line model [33].

We visualised the metabolic reactions with the software developed by our group, which we made publicly available at `http://lrss.fri.uni-lj.si/bio/material/viz_tool_matlab_v02.zip`. The software is written in Matlab and supports the automatic visualisation of GEMs using bipartite graph approach. The visualisation is performed in the neighbourhood of the metabolites of interest that are specified by the user. Size of the neighbourhood is defined with the maximal distance from the specified metabolites. User can also specify compartments, which should be included in the visualisation. The software runs FBA in the background, which allows the user to visualise only the reactions with certain flux values. Moreover, user can trigger perturbations with the modification of flux boundaries. The visualisation of the results of the specified perturbations can be performed in three different ways, i.e. (1) visualisation of reactions that become active after the perturbations, (2) visualisation of reactions that become inactive after the perturbations, and (3) visualisation of reactions that remain active, but change reaction fluxes after the perturbations.

7.1. Flux balance analysis of unperturbed model

We can use the FBA to assess the steady-state metabolic fluxes that optimise the objective function under the given constraints. We performed the FBA on the CHO-S cell line model using the maximisation of metabolic flux through the biomass reaction for a producing cell line as an optimisation criterion. We used the constraints as provided within the original paper of Hefzi et al. and in the model iCHO1766 for the selected cell line [33]. Obtained solution presented a reference point for our further analyses described in Sections 7.2 and 7.3.

We will demonstrate the FBA assessment of steady-state metabolic fluxes on the reactions that are directly connected to the metabolism of *cytosolic asparagine*. We illustrate this segment in Fig. 4. The stoichiometric matrix $\boldsymbol{N}$ represents a mathematical description of the visualised segment:

$$
\begin{array}{l}
\\
\mathit{asn_L}[c]\\ \mathit{asn_L}[e]\\ \mathit{ala_L}[e]\\ \mathit{ala_L}[c]\\ \mathit{cys_L}[c]\\ \mathit{cys_L}[e]\\ \mathit{gln_L}[c]\\ \mathit{gln_L}[e]\\ \mathit{gly}[e]\\ \mathit{gly}[c]\\ \mathit{hom_L}[e]\\ \mathit{hom_L}[c]\\ \mathit{ile_L}[c]\\ \mathit{ile_L}[e]\\ \mathit{leu_L}[c]\\ \mathit{leu_L}[e]\\ \mathit{met_L}[c]\\ \mathit{met_L}[e]\\ \mathit{pro_L}[e]\\ \mathit{pro_L}[c]\\ \mathit{ser_L}[c]\\ \mathit{ser_L}[e]\\ \mathit{thr_L}[c]\\ \mathit{thr_L}[e]\\ \mathit{trp_L}[c]\\ \mathit{trp_L}[e]\\ \mathit{tyr_L}[c]\\ \mathit{tyr_L}[e]\\ \mathit{val_L}[c]\\ \mathit{val_L}[e]
\end{array}
\begin{array}{ccccccccccccccc}
r1637 & r1643 & r1648 & r1653 & r1654 & r1655 & r1656 & r1553 & r1567 & r1580 & r1592 & r1603 & r1613 & r1630 & r2532\\
\end{array}
$$

$$
\boldsymbol{N} = \left(\begin{array}{ccccccccccccccc}
-1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1\\
1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0\\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0\\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right). \tag{1}
$$

Here, rows represent the metabolites and columns represent the metabolic reactions. We can describe the steady-state fluxes through the observed

Figure 4: Visualisation of the reactions that directly affect the cytosolic asparagine concentrations (asn_L[c]). We assessed the fluxes through the reactions with the FBA. We used biomass reaction for a producing cell line as an optimisation criterion. Blue nodes correspond to metabolites, red nodes to the reactions consuming and green nodes to the reactions producing cytosolic asparagine. Numbers assigned to each of the corresponding nodes describe the fluxes through the metabolic reactions. We explain all abbreviations in the Supplementary text.

metabolic reactions with the vector

$$\boldsymbol{v} = \begin{pmatrix} -1000 \\ -1000 \\ -1000 \\ -1000 \\ -1000 \\ -1000 \\ -1000 \\ -1000 \\ 1000 \\ 1000 \\ -1000 \\ 0.0034 \\ -1000 \\ -1000 \\ -999.9966 \end{pmatrix}. \tag{2}$$

Multiplication of the matrix $\boldsymbol{N}$ with the vector $\boldsymbol{v}$ should yield zero for the observed metabolite (note that due to the simplicity, we omitted the reactions not directly connected to cytosolic asparagine from the visualisation as well as from the stoichiometric matrix). The product of the matrix with the vector of fluxes is another vector with the first element equal to 0. That element corresponds to the rate of change of cytosolic asparagine. Therefore, the concentration of this metabolite is in steady-state when this assumed flux distribution holds in the network.

7.2. Analysis of model perturbations and their visualisation

One of the main benefits of constraint-based approaches is their capability to assess the consequences of perturbations on the metabolic networks. We can perform the perturbations with the modification of constraints defining the viable fluxes through observed metabolic reactions on an arbitrary segment of the model. We are thus able to simulate the consequences of knock-downs of genes encoding specific enzymes. Moreover, we can use these perturbations to assess the relevance of specific substrates within the cellular medium.

We can describe the availability of substrates within the medium with so-called uptake reactions and their corresponding flux boundaries. In our next example we observed the consequences of the removal of two non-essential amino acids from the medium, namely asparagine and glutamine. This was

achieved by setting the upper and lower flux boundaries through the corresponding uptake reactions to zero. FBA was performed on the additionally constrained model and the results obtained were compared with the results of the reference analysis (see Section 7.1). We observed the consequences of the perturbations in a qualitative manner with the visualisation of the reactions that became active after a perturbation was introduced.

Asparagine and glutamine serve as important sources of nitrogen and energy in the mammalian cells. They highly affect the cell growth, protein production, by-product release and energy metabolism [87]. CHO cells are however able to synthesize these two amino acids when in demand [33].

Asparagine and glutamine biosynthesis reactions are inactive when asparagine and glutamine are available in the medium. We observed this *in silico* with the FBA on the unperturbed GEM (see Section 7.1). In our first perturbation analysis we simulated the removal of asparagine from the medium. This caused the activation of asparagine biosynthesis reaction (see upper left Fig. 5). Obtained results correspond with the experimental results reported in the literature [88]. Asparagine is synthesized from aspartate with the reaction catalysed by the asparagine synthase. When asparagine is not present in the medium, the asparagine synthase gene expression is increased. In our second perturbation analysis we simulated the removal of glutamine from the medium. In this case two reactions became active, i.e. asparagine degradation which presents an additional source of ammonia (see middle left Fig. 5) and glutamine biosynthesis, which consumes ammonia and glutamate, and is catalysed by the glutamine synthetase (see midle right Fig. 5). This again corresponds with the experimental results reported in the literature [89]. When cells are exposed to growth in medium without glutamine, the glutamine synthetase gene expression is increased. In our third perturbation analysis both amino acids were removed. In this case, asparagine as well as glutamine biosynthesis reactions were active (see bottom row in Fig. 5).

7.3. Evaluating the effects of medium composition on cell growth

The optimisation of amino acid composition of CHO cell culture media is important for optimal cell growth and efficient production of recombinant proteins [68]. We quantified the influence of selected amino acids on cell growth. We perturbed the original iCHO1766 CHO-S cell line model [33] with the constraints that reduce the uptake of selected amino acids. The original model presumes the medium availability of arginine, asparagine, aspartate, cysteine, glutamine, histidine, isoleucine, leucine, lysine, methionine,

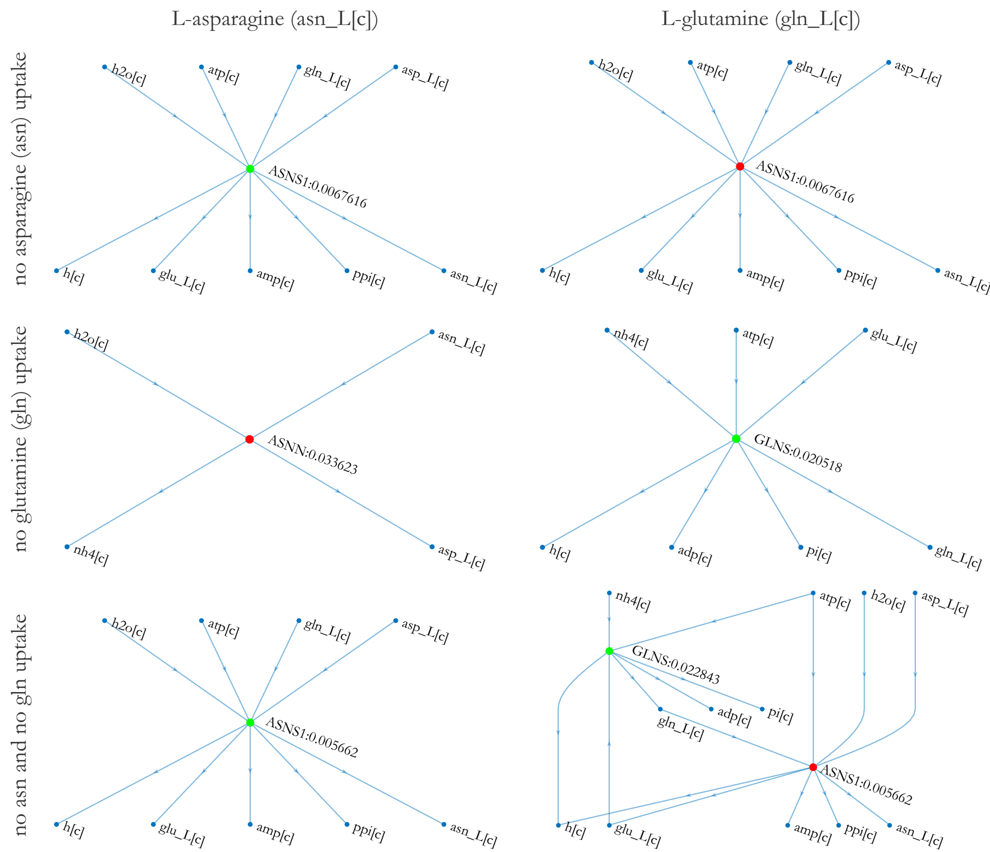

Figure 5: Visualisation of the reactions that become active after we removed the selected amino acids from the medium. Figs. in the left column represent the reactions that directly affect the cytosolic asparagine concentrations (asn_L[c]). Figs. in the right column represent the reactions that directly affect the cytosolic glutamine concentrations (gln_L[c]). Upper row represents the reactions that become active after the elimination of asparagine from the medium, middle row the reactions that become active after the elimination of glutamine from the medium, and bottom row the reactions that become active after the elimination of both amino acids from the medium. Blue nodes correspond to metabolites, red nodes to the consuming and green nodes to the producing reactions. Numbers assigned to each of the corresponding nodes describe the fluxes through the metabolic reactions. We use the following abbreviations for reaction names: ASNS1 – asparagine synthase (glutamine-hydrolysing), GLNS – glutamine synthetase, ASNN – L-asparaginase. We explain the remaining abbreviations in the Supplementary text.

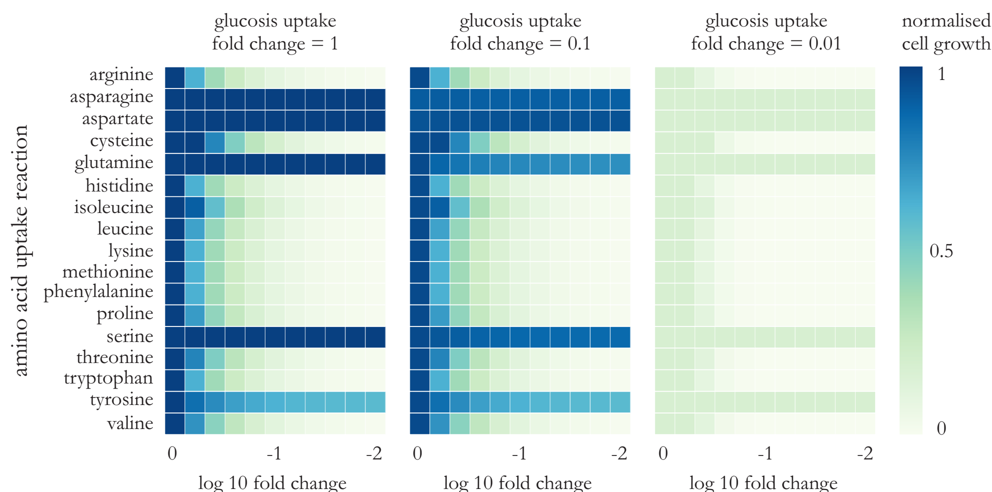

Figure 6: Dependence of cell growth on amino acid availability in medium. We performed the simulations with unperturbed (left), with 10- (middle) and with 100-fold reduction of glucosis uptake (right).

phenylalanine, proline, serine, threonine, tryptophan, tyrosine and valine. We performed our analysis with the additional constraints imposed on their uptake reactions as well as on the uptake reaction of glucosis, by the reduction of the reaction flux boundaries. FBA was performed after the flux boundary reductions were introduced and the results were compared with the results of the reference analysis (see Section 7.1).

Cells can compensate the removal of non-essential amino acids with the increased glucose uptake and increased activity of amino acid biosynthesis reactions (see Section 7.2). We conducted the perturbations of amino acid uptake reactions together with the glucose uptake reaction. We performed three different analyses, i.e. with unperturbed, with 10- and with 100-fold reduction of glucosis uptake. We present our results in Fig. 6. Arginine, cysteine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, threonine, tryptophan and valine present essential amino acids in selected CHO cell lines. When their availability was decreased by 10-fold, cells stopped growing. Remaining amino acids, i.e. asparagine, aspartic acide, glutamine, serine and tyrosine, are non-essential. Their removal from the medium was compensated with their biosynthesis reactions. When glucose was in excess only tyrosine affected the cell growth (see left Fig. 6). When glucose was

decreased by 10-fold, cell growth was still comparable to the unperturbed growth. In this scenario only glutamine and tyrosine showed to noticeably influence the cell growth (see middle Fig. 6), which coincides with the literature data [87]. When glucose was decreased by 100-fold, cell growth was decreased more drastically. Cells were, however, still able to grow without non-essential amino acids (see right Fig. 6).

8. Conclusion

Although the response of metabolic networks is mainly derived from simple enzymatic reactions, they possess complex and rich dynamical properties. Their study requires complex systems approaches. The dynamics of metabolic networks can be partially reproduced with the GEMs in combination with computational approaches we described. Even though the accuracy of these reconstructions are far from being perfect, our journey does not stop with the flawless GEMs. Combining computational models of metabolic networks with gene regulatory, protein interaction and signalling networks (for example, see [90, 91]) as well as with other cellular processes into whole-cell models [92] promises a whole new perspective. This will bring us to the capabilities of building representative virtual tissues and virtual organs with the long term goal to computationally reconstruct the whole human body [93]. Even though a perfect *in silico* reconstruction of the human body currently seems very far from the reality, different state-of-the-art computational models already serve as an excellent basis for the extraction of novel knowledge. In many cases they drastically reduce the amount of experimental and clinical work, improve diagnostic tools and increase our understanding of complex biological phenomena [94, 95].

Here we described some of these approaches that have been vastly applied in recent years to the fields not directly related to computational modelling, such as metabolic engineering, systems medicine and production of biopharmaceuticals. Even though many details are omitted, this review should serve as a good introduction to the computational reconstruction and analysis of GEMs.

Acknowledgements

The research was partially supported by the scientific-research programme Pervasive Computing (P2-0359) financed by the Slovenian Research Agency in the years from 2009 to 2017 and by the basic research and application

project Designed cellular logic (J1-6740) financed by the Slovenian Research Agency in the years from 2014 to 2017. We acknowledge also resources of FP7 CASyM (Coordinating Action Systems Medicine Europe, Grant no. 305033), and the Slovenian Research Agency grants P1-0390 and the infrastructure grant ELIXIR.

[1] F. R. Blattner, The complete genome sequence of escherichia coli K-12, Science 277 (5331) (1997) 1453–1462.

[2] E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh et al., Initial sequencing and analysis of the human genome, Nature 409 (6822) (2001) 860–921.

[3] D. McCloskey, B. O. Pallson, A. M. Feist, Basic and applied uses of genome-scale metabolic network reconstructions of Escherichia coli, Molecular Systems Biology 9 (661) (2013).

[4] E. J. OBrien, J. M. Monk, B. O. Palsson, Using genome-scale models to predict biological capabilities, Cell 161 (5) (2015) 971–987.

[5] N. C. Duarte, S. A. Becker, N. Jamshidi, I. Thiele, M. L. Mo, T. D. Vo, R. Srivas, B. Ø. Palsson, Global reconstruction of the human metabolic network based on genomic and bibliomic data, Proceedings of the National Academy of Sciences 104 (6) (2007) 1777–1782.

[6] N. Swainston, K. Smallbone, H. Hefzi, P. D. Dobson, J. Brewer, M. Hanscho, D. C. Zielinski, K. S. Ang, N. J. Gardiner, J. M. Gutierrez, et al., Recon 2.2: from reconstruction to model of human metabolism, Metabolomics 12 (7) (2016) 1–7.

[7] M. Kanehisa, S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki, M. Hirakawa, From genomics to chemical genomics: new developments in KEGG, Nucleic Acids Research 1 (34) (2006) D354–7.

[8] J. Barthelmes, C. Ebeling, A. Chang, I. Schomburg, D. Schomburg, BRENDA, AMENDA and FRENDA: the enzyme information system in 2007, Nucleic Acids Research 35 (2007) D511–4.

[9] R. Caspi, R. Billington, L. Ferrer, H. Foerster, C. A. Fulcher, I. M. Keseler, A. Kothari, M. Krummenacker, M. Latendresse, L. A. Mueller, Q. Ong, S. Paley, P. Subhraveti, D. S. Weaver, P. D. Karp, The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases, Nucleic Acids Research 44 (D1) (2016) D471–D480.

[10] N. Juty, R. Ali, M. Glont, S. Keating, N. Rodriguez, M. J. Swat, S. M. Wimalaratne, H. Hermjakob, N. Le Novère, C. Laibe, V. Chelliah, BioModels: Content, features, functionality and use., CPT: Pharmacometrics & Systems Pharmacology 4 (2) (2015) 55–68.

[11] Z. A. King, J. Lu, A. Dräger, P. Miller, S. Federowicz, J. A. Lerman, A. Ebrahim, B. O. Palsson, N. E. Lewis, BiGG models: A platform for integrating, standardizing and sharing genome-scale models, Nucleic Acids Research 44 (D1) (2015) D515–D522.

[12] D. Jullesson, F. David, B. Pfleger, J. Nielsen, Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals, Biotechnology Advances 33 (7) (2015) 1395–1402.

[13] C. L. Barrett, T. Y. Kim, H. U. Kim, B. O. Palsson, S. Y. Lee, Systems biology as a foundation for genome-scale synthetic biology, Current Opinion in Biotechnology 17 (5) (2006) 488–492.

[14] Z. A. Irani, E. J. Kerkhoven, S. A. Shojaosadati, J. Nielsen, Genome-scale metabolic model of Pichia pastoris with native and humanized glycosylation of recombinant proteins, Biotechnology and Bioengineering 113 (5) (2016) 961–969.

[15] C. Chen, H. Le, C. T. Goudar, Integration of systems biology in cell line and process development for biopharmaceutical manufacturing, Biochemical Engineering Journal 107 (2016) 11–17.

[16] K. P. Jayapal, K. Wlaschin, W. Hu, M. G. S. Yap, Recombinant protein therapeutics from CHO cells – 20 years and counting, Chemical Engineering Progress 103 (10) (2007) 40–47.

[17] A. Naik, D. Rozman, A. Belic, SteatoNet: The first integrated human metabolic model with multi-layered regulation to investigate liver-

associated pathologies, PLoS Computational Biology 10 (12) (2014) e1003993.

[18] A. Mardinoglu, R. Agren, C. Kampf, A. Asplund, M. Uhlen, J. Nielsen, Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease, Nature Communications 5 (2014).

[19] T. Cvitanović, M. C. Reichert, M. Moškon, M. Mraz, F. Lammert, D. Rozman, Large-scale computational models of liver metabolism: How far from the clinics?, Hepatology (2017).

[20] F. Gatto, H. Miess, A. Schulze, J. Nielsen, Flux balance analysis predicts essential genes in clear cell renal cell carcinoma metabolism, Scientific Reports 5 (2015) 10738.

[21] P. Ghaffari, A. Mardinoglu, J. Nielsen, Cancer metabolism: A modeling perspective, Frontiers in Physiology 6 (2015).

[22] P. Karp, M. Latendresse, S. M. Paley, M. Krummenacker, Q. Ong, , R. Billington, A. Kothari, D. Weaver, T. Lee, P. Subhraveti, A. Spaulding, C. Fulcher, I. Keseler, R. Caspi, Pathway Tools version 19.0: Integrated software for pathway/genome informatics and systems biology,, Briefings in Bioinformatics 17 (5) (2016) 877–890.

[23] R. Agren, L. Liu, S. Shoaie, W. Vongsangnak, I. Nookaew, J. Nielsen, The RAVEN toolbox and its use for generating a genome-scale metabolic model for Penicillium chrysogenum, PLoS Computational Biology 9 (3) (2013) e1002980.

[24] S. A. Becker, A. M. Feist, M. L. Mo, G. Hannum, B. O. Palsson, M. J. Herrgard, Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox, Nature Protocols 2 (2007) 727–738.

[25] J. Schellenberger, R. Que, R. M. Fleming, I. Thiele, J. Orth, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis, S. Rahmanian, J. Kang, D. Hyduke, B. O. Palsson, Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0, Nature protocols 6 (9) (2011) 1290–1307.

[26] J. D. Orth, I. Thiele, B. O. Palsson, What is flux balance analysis?, Nature Biotechnology 28 (3) (2010) 245–248.

[27] S. A. Becker, B. O. Palsson, Context-specific metabolic networks are consistent with experiments, PLoS Comput Biol 4 (5) (2008) e1000082.

[28] S. Opdam, A. Richelle, B. Kellman, S. Li, D. C. Zielinski, N. E. Lewis, A systematic evaluation of methods for tailoring genome-scale metabolic models, Cell Systems 4 (3) (2017) 318–329.

[29] A. P. Burgard, P. Pharkya, C. D. Maranas, Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization, Biotechnol. Bioeng. 84 (6) (2003) 647–657.

[30] N. E. Lewis, H. Nagarajan, B. O. Palsson, Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods, Nature Reviews Microbiology 10 (4) (2012) 291–305.

[31] Z. A. King, A. Dräger, A. Ebrahim, N. Sonnenschein, N. E. Lewis, B. O. Palsson, Escher: A web application for building, sharing, and embedding data-rich visualizations of biological pathways, PLOS Computational Biology 11 (8) (2015) e1004321.

[32] A. Noronha, A. D. Danielsdottir, P. Gawron, F. Johannsson, S. Jonsdottir, S. Jarlsson, J. P. Gunnarsson, S. Brynjolfsson, R. Schneider, I. Thiele et al., ReconMap: an interactive visualization of human metabolism, Bioinformatics 33 (4) (2017) 605–607.

[33] H. Hefzi, K. S. Ang, M. Hanscho, A. Bordbar, D. Ruckerbauer, M. Lakshmanan, C. A. Orellana, D. Baycin-Hizal, Y. Huang, D. Ley, et al., A consensus genome-scale reconstruction of chinese hamster ovary cell metabolism, Cell Systems 3 (5) (2016) 434–443.

[34] N. Le Novere, Quantitative and logic modelling of molecular and gene networks, Nat. Rev. Genet. 16 (3) (2015) 146–158.

[35] B. Palsson, Systems Biology: Properties of Reconstructed Networks, Cambridge University Press, 2006.

[36] N. Zamboni, S.-M. Fendt, M. Rühl, U. Sauer, 13C-based metabolic flux analysis, Nature Protocols 4 (6) (2009) 878–892.

[37] N. D. Price, J. L. Reed, B. Ø. Palsson, Genome-scale models of microbial cells: evaluating the consequences of constraints, Nature Reviews Microbiology 2 (11) (2004) 886–897.

[38] I. Thiele, B. O. Pallson, A protocol for generating a high-quality genome-scale metabolic reconstruction, Nature Protocols 5 (2010) 93–121.

[39] R. Mahadevan, J. S. Edwards, F. J. Doyle, Dynamic flux balance analysis of diauxic growth in Escherichia coli, Biophysical journal 83 (3) (2002) 1331–1340.

[40] X. Feng, Y. Xu, Y. Chen, Y. J. Tang, Integrating flux balance analysis into kinetic models to decipher the dynamic metabolism of Shewanella oneidensis MR-1, PLOS Computational Biology 8 (3) (2012).

[41] M. W. Covert, C. H. Schilling, B. O. Palsson, Regulation of gene expression in flux balance models of metabolism, Journal of Theoretical Biology 213 (1) (2001) 73–88.

[42] N. E. Lewis, K. K. Hixson, T. M. Conrad, J. A. Lerman, P. Charusanti, A. D. Polpitiya, J. N. Adkins, G. Schramm, S. O. Purvine, D. Lopez-Ferrer et al., Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models, Molecular Systems Biology 6 (390) (2010).

[43] M. Lakshmanan, G. Koh, B. K. S. Chung, D.-Y. Lee, Software applications for flux balance analysis, Briefings in Bioinformatics 15 (1) (2014) 108–122.

[44] E. Klipp, W. Liebermeister, C. Wierling, A. Kowald, Systems Biology: A Textbook, 2nd Edition, Wiley, 2016.

[45] Z. A. King, J. Lu, A. Dräger, P. Miller, S. Federowicz, J. A. Lerman, A. Ebrahim, B. O. Palsson, N. E. Lewis, BiGG models: A platform for integrating, standardizing and sharing genome-scale models, Nucleic acids research 44 (D1) (2016) D515–D522.

[46] Ó. Rolfsson, B. O. Palsson, Decoding the jargon of bottom-up metabolic systems biology, BioEssays 37 (6) (2015) 588–591.

[47] C. R. Haggart, J. A. Bartell, J. J. Saucerman, J. A. Papin, Whole-genome metabolic network reconstruction and constraint-based modeling?, Methods in Enzymology 500 (2011) 411–433.

[48] G. J. Baart, D. E. Martens, Genome-scale metabolic models: reconstruction and analysis, Neisseria meningitidis: Advanced Methods and Protocols (2012) 107–126.

[49] A. M. Feist, M. J. Herrgård, I. Thiele, J. L. Reed, B. Ø. Palsson, Reconstruction of biochemical networks in microorganisms, Nature Reviews Microbiology 7 (2) (2009) 129–143.

[50] M. Durot, P.-Y. Bourguignon, V. Schachter, Genome-scale models of bacterial metabolism: reconstruction and applications, FEMS Microbiology Reviews 33 (1) (2009) 164–190.

[51] B. G. Kremkow, J. Y. Baik, M. L. MacDonald, K. H. Lee, CHOgenome.org 2.0: Genome resources and website updates, Biotechnology journal 10 (7) (2015) 931–938.

[52] J. W. Pinney, metaSHARK: software for automated metabolic network prediction from DNA sequence and its application to the genomes of Plasmodium falciparum and Eimeria tenella, Nucleic Acids Research 33 (4) (2005) 1399–1409.

[53] M. P. Pacheco, T. Pfau, T. Sauter, Benchmarking procedures for high-throughput context specific reconstruction algorithms, Frontiers in Physiology 6 (2016).

[54] I. Thiele, N. Vlassis, R. M. T. Fleming, fastGapFill: efficient gap filling in metabolic networks, Bioinformatics 30 (17) (2014) 2529–2531.

[55] S. Prigent, C. Frioux, S. M. Dittami, S. Thiele, A. Larhlimi, G. Collet, F. Gutknecht, J. Got, D. Eveillard, J. Bourdon et al., Meneco, a topology-based gap-filling tool applicable to degraded genome-wide metabolic networks, PLOS Computational Biology 13 (1) (2017) e1005276.

[56] Z. Hosseini, S.-A. Marashi, Discovering missing reactions of metabolic networks by using gene co-expression data, Scientific Reports 7 (2017) 41774.

[57] M. Y. Becker, I. Rojas, A graph layout algorithm for drawing metabolic pathways, Bioinformatics 17 (5) (2001) 461–467.

[58] J. van Helden, L. Wernisch, D. Gilbert, S. Wodak, Graph-based analysis of metabolic networks, in: Bioinformatics and genome analysis, Springer, 2002, pp. 245–274.

[59] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome research 13 (11) (2003) 2498–2504.

[60] M. König, A. Dräger, H.-G. Holzhütter, CySBML: a Cytoscape plugin for SBML, Bioinformatics 28 (18) (2012) 2402–2403.

[61] P. A. Jensen, J. A. Papin, MetDraw: automated visualization of genome-scale metabolic network reconstructions and high-throughput data, Bioinformatics 30 (9) (2014) 1327–1328.

[62] R. Bourqui, L. Cottret, V. Lacroix, D. Auber, P. Mary, M.-F. Sagot, F. Jourdan, Metabolic network visualization eliminating node redundance and preserving metabolic pathways, BMC systems biology 1 (29) (2007).

[63] A. Kostromins, E. Stalidzans, Paint4net: COBRA toolbox extension for visualization of stoichiometric models of metabolism, Biosystems 109 (2) (2012) 233–239.

[64] S. N. Galleguillos, D. Ruckerbauer, M. P. Gerstl, N. Borth, M. Hanscho, J. Zanghellini, What can mathematical modelling say about CHO metabolism and protein glycosylation?, Computational and Structural Biotechnology Journal 15 (2017) 212–221.

[65] H. F. Kildegaard, D. Baycin-Hizal, N. E. Lewis, M. J. Betenbaugh, The emerging CHO systems biology era: harnessing the omics revolution for biotechnology, Current opinion in biotechnology 24 (6) (2013) 1102–1107.

[66] G. B. Nyberg, R. R. Balcarcel, B. D. Follstad, G. Stephanopoulos, D. I. C. Wang, Metabolism of peptide amino acids by Chinese hamster

ovary cells grown in a complex medium, Biotechnology and bioengineering 62 (3) (1999) 324–335.

[67] S. Selvarasu, Y. S. Ho, W. P. K. Chong, N. S. C. Wong, F. N. K. Yusufi, Y. Y. Lee, M. G. S. Yap, D.-Y. Lee, Combined in silico modeling and metabolomics analysis to characterize fed-batch CHO cell culture, Biotechnology and Bioengineering 109 (6) (2012) 1415–1429.

[68] Z. Xing, B. Kenty, I. Koyrakh, M. Borys, S.-H. Pan, Z. J. Li, Optimizing amino acid composition of CHO cell culture media for a fusion protein production, Process Biochemistry 46 (7) (2011) 1423–1429.

[69] Z. Sheikholeslami, M. Jolicoeur, O. Henry, Elucidating the effects of postinduction glutamine feeding on the growth and productivity of CHO cells, Biotechnology Progress 30 (3) (2014) 535–546.

[70] V. S. Martnez, S. Dietmair, L.-E. Quek, M. P. Hodson, P. Gray, L. K. Nielsen, Flux balance analysis of CHO cells before and after a metabolic switch from lactate production to consumption, Biotechnology and Bioengineering 110 (2) (2013) 660–666.

[71] W. S. Ahn, M. R. Antoniewicz, Metabolic flux analysis of CHO cells at growth and non-growth phases using isotopic tracers and mass spectrometry, Metabolic engineering 13 (5) (2011) 598–609.

[72] N. Templeton, J. Dean, P. Reddy, J. D. Young, Peak antibody production is associated with increased oxidative metabolism in an industrially relevant fed-batch CHO cell culture, Biotechnology and bioengineering 110 (7) (2013) 2013–24.

[73] H. W. Lee, A. Christie, J. A. Starkey, E. K. Read, S. Yoon, Intracellular metabolic flux analysis of CHO cells supplemented with wheat hydrolysates for improved mAb production and cell-growth, Journal of Chemical Technology and Biotechnology 90 (2) (2015) 291–302.

[74] A. Nicolae, J. Wahrheit, J. Bahnemann, A.-P. Zeng, E. Heinzle, Non-stationary 13C metabolic flux analysis of Chinese hamster ovary cells in batch culture using extracellular labeling highlights metabolic reversibility and compartmentation, BMC Systems Biology 8 (1) (2014) 50.

[75] J. Wahrheit, J. Niklas, E. Heinzle, Metabolic control at the cytosolmitochondria interface in different growth phases of CHO cells, Metabolic Engineering 23 (2014) 9–21.

[76] N. Chen, M. H. Bennett, C. Kontoravdi, Analysis of Chinese hamster ovary cell metabolism through a combined computational and experimental approach, Cytotechnology 66 (6) (2014) 945–966.

[77] A. Ghorbaniaghdam, O. Henry, M. Jolicoeur, A kinetic-metabolic model based on cell energetic state: study of CHO cell behavior under Nabutyrate stimulation, Bioprocess and Biosystems Engineering 36 (4) (2013) 469–487.

[78] R. P. Nolan, K. Lee, Dynamic model of CHO cell metabolism, Metabolic engineering 13 (1) (2011) 108–124.

[79] X. Xu, H. Nagarajan, N. E. Lewis, S. Pan, Z. Cai, X. Liu, W. Chen, M. Xie, W. Wang, S et al. Hammond, The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line, Nature Biotechnology 29 (8) (2011) 735–741.

[80] F. N. K. Yusufi, M. Lakshmanan, Y. S. Ho, B. L. W. Loo, P. Ariyaratne, Y. Yang, S. K. Ng, T. R. M. Tan, H. C. Yeo, H. L. Lim et al., Mammalian systems biotechnology reveals global cellular adaptations in a recombinant CHO cell line, Cell Systems 4 (5) (2017) 530–542.e6.

[81] W. S. Ahn, M. R. Antoniewicz, Parallel labeling experiments with [1,2-13C]glucose and [U-13C]glutamine provide new insights into CHO cell metabolism, Metabolic engineering 15 (2013) 34–47.

[82] N. Carinhas, T. M. Duarte, L. C. Barreiro, M. J. Carrondo, P. M. Alves, A. P. Teixeira, Metabolic signatures of GS-CHO cell clones associated with butyrate treatment and culture phase transition, Biotechnology and bioengineering 110 (12) (2013) 3244–3257.

[83] V. S. Martínez, M. Buchsteiner, P. Gray, L. K. Nielsen, L.-E. Quek, Dynamic metabolic flux analysis using B-splines to study the effects of temperature shift on CHO cell metabolism, Metabolic Engineering Communications 2 (2015) 46–57.

[84] L.-E. Quek, S. Dietmair, M. Hanscho, V. S. Martínez, N. Borth, L. K. Nielsen, Reducing Recon 2 for steady-state flux analysis of HEK cell culture, Journal of biotechnology 184 (2014) 172–178.

[85] I. Thiele, N. Swainston, R. M. Fleming, A. Hoppe, S. Sahoo, M. K. Aurich, H. Haraldsdottir, M. L. Mo, O. Rolfsson, M. D. Stobbe, et al., A community-driven global reconstruction of human metabolism, Nature biotechnology 31 (5) (2013) 419–425.

[86] A. M. Feist, B. O. Palsson, The biomass objective function, Current opinion in microbiology 13 (3) (2010) 344–349.

[87] L.-X. Zhang, W.-Y. Zhang, C. Wang, J.-T. Liu, X.-C. Deng, X.-P. Liu, L. Fan, W.-S. Tan, Responses of CHO-DHFR cells to ratio of asparagine to glutamine in feed media: cell growth, antibody production, metabolic waste, glutamate, and energy metabolism, Bioresources and Bioprocessing 3 (1) (2016) 5.

[88] S. M. Arfin, D. R. Simpson, C. Chiang, I. L. Andrulis, G. W. Hatfield, A role for asparaginyl-tRNA in the regulation of asparagine synthetase in a mammalian cell line, Proceedings of the National Academy of Sciences 74 (6) (1977) 2367–2369.

[89] A. Sanfeliu, G. Stephanopoulos, Effect of glutamine limitation on the death of attached Chinese hamster ovary cells, Biotechnology and bioengineering 64 (1) (1999) 46–53.

[90] S. Lee, C. Zhang, M. Kilicarslan, B. Piening, E. Bjornson, B. Hallstrm, A. Groen, E. Ferrannini, M. Laakso, M. Snyder et al., Integrated network analysis reveals an association between plasma mannose levels and insulin resistance, Cell Metabolism 24 (1) (2016) 172–184.

[91] S. Lee, A. Mardinoglu, C. Zhang, D. Lee, J. Nielsen, Dysregulated signaling hubs of liver lipid metabolism reveal hepatocellular carcinoma pathogenesis, Nucleic Acids Research 44 (12) (2016) 5529–5539.

[92] J. R. Karr, K. Takahashi, A. Funahashi, The principles of whole-cell modeling, Current Opinion in Microbiology 27 (2015) 18–24.

[93] P. Kohl, D. Noble, Systems biology and the virtual physiological human, Molecular Systems Biology 5 (292) (2009).

[94] H. Fouladiha, S.-A. Marashi, Biomedical applications of cell- and tissue-specific metabolic network models, Journal of Biomedical Informatics 68 (2017) 35–49.

[95] J. Y. Ryu, H. U. Kim, S. Y. Lee, Reconstruction of genome-scale human metabolic models using omics data., Integrative Biology 7 (8) (2015) 859–68.