Journal of Biotechnology 261 (2017) 221–228

Contents lists available at ScienceDirect

Journal of Biotechnology

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

Use of *CellNetAnalyzer* in biotechnology and metabolic engineering

Axel von Kamp, Sven Thiele, Oliver Hädicke, Steffen Klamt*

Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany

ARTICLE INFO

Keywords:
Metabolic networks
Metabolic flux analysis
Flux balance analysis
Elementary modes
Strain optimization
Minimal cut sets

ABSTRACT

Mathematical models of the cellular metabolism have become an essential tool for the optimization of biotechnological processes. They help to obtain a systemic understanding of the metabolic processes in the used microorganisms and to find suitable genetic modifications maximizing the production performance. In particular, methods of stoichiometric and constraint-based modeling are frequently used in the context of metabolic and bioprocess engineering. Since metabolic networks can be complex and comprise hundreds or even thousands of metabolites and reactions, dedicated software tools are required for an efficient analysis. One such software suite is *CellNetAnalyzer*, a MATLAB package providing, among others, various methods for analyzing stoichiometric and constraint-based metabolic models. *CellNetAnalyzer* can be used via command-line based operations or via a graphical user interface with embedded network visualizations. Herein we will present key functionalities of *CellNetAnalyzer* for applications in biotechnology and metabolic engineering and thereby review constraint-based modeling techniques such as metabolic flux analysis, flux balance analysis, flux variability analysis, metabolic pathway analysis (elementary flux modes) and methods for computational strain design.

1. Introduction

Industrial (white) biotechnology uses cells or parts of cells for the production of chemicals, biofuels, pharmaceuticals, nutraceuticals, enzymes or other industrially relevant products. The design and optimization of biotechnological processes usually involves genetic modifications in the metabolism of the used production organisms to maximize their production performance. For the targeted (rational) metabolic engineering of cell factories, mathematical modeling of the cellular metabolism has become an essential tool. Various theoretical methods of metabolic modeling have been developed to analyze the capabilities of metabolic networks, to study the behavior of the metabolism under different growth and production conditions, to discover potential bottlenecks and to eventually identify targets for genetic modifications redirecting metabolic fluxes to a desired compound. In particular, methods of stoichiometric and constraint-based modeling have been successfully applied in metabolic and bioprocess engineering (Gutierrez and Lewis, 2015; King et al., 2015; Maia et al., 2015; Simeonidis and Price, 2015; Kim et al., 2015; Machado and Herrgard, 2015). These methods include, for example, metabolic flux analysis (MFA; characterization of metabolic fluxes under controlled conditions), flux balance analysis (FBA; analysis of optimal flux distributions), flux variability analysis (FVA; analysis of feasible ranges of metabolic fluxes), metabolic pathway analysis (discovery and analysis of metabolic pathways) and methods for computational strain design (computation of metabolic engineering strategies optimizing the production behavior of the organism).

Since stoichiometric and constraint-based metabolic models may involve hundreds or even thousands of metabolites and reactions, dedicated software tools are required to support an efficient analysis of (up to genome-scale) metabolic networks. Accordingly, several software packages for constraint-based modeling have been developed in the past years, including, for example, the COBRA toolbox (Schellenberger et al., 2011; Ebrahim et al., 2013), OPTFLUX (Rocha et al., 2010), OMIX (Droste et al., 2011, 2013), MUFINS (Wu et al., 2016), RAVEN (Agren et al., 2013) and *CellNetAnalyzer* (Klamt et al., 2007). COPASI (Hoops et al., 2006), presented in another article in this special issue, also supports analysis of basic stoichiometric features of metabolic networks but focuses more on kinetic modeling of biochemical systems.

CellNetAnalyzer is a package for MATLAB providing various (partially unique) algorithms for analyzing structure and function of biological networks. Metabolic networks can be studied based on stoichiometric and constraint-based models whereas signaling and

Abbreviations: CNA, *CellNetAnalyzer*; SCBM, Stoichiometric and constraint-based modeling; CR(s), Conservation relation(s); MFA, Metabolic flux analysis; FBA, Flux balance analysis; FVA, Flux variability analysis; EFM(s), Elementary flux modes(s); MCS(s), Minimal cut set(s)

* Corresponding author.
E-mail address: klamt@mpi-magdeburg.mpg.de (S. Klamt).

http://dx.doi.org/10.1016/j.jbiotec.2017.05.001
Received 26 January 2017; Received in revised form 28 April 2017; Accepted 3 May 2017
Available online 10 May 2017

regulatory networks can be explored by qualitative and semi-quantitative modeling approaches. The development of the software started more than 15 years ago and since then its scope and functionality has grown steadily. *CellNetAnalyzer* can be used via command-line based operations or via a graphical user interface with embedded network visualizations. Herein we will present key functionalities of *CellNetAnalyzer* for applications in biotechnology and metabolic engineering and thereby briefly review constraint-based modeling techniques.

2. Fundamentals of stoichiometric and constraint-based modeling

We start with a short introduction to the mathematical foundations of stoichiometric and constraint-based modeling (SCBM); detailed descriptions can be found elsewhere (Maarleveld et al., 2013; O'Brien et al., 2015; Klamt et al., 2014). SCBM methods require as (minimal) input the $m \times q$ stoichiometric matrix $\mathbf{N}$ capturing the structure of the metabolic network (columns: q reactions; rows: m metabolites with their reaction stoichiometries). Central to all SCBM methods is the assumption of steady state (concentrations of intracellular metabolites do not change) which implies the metabolite balancing equation

$$\mathbf{Nr} = \mathbf{0} \tag{1}$$

where $\mathbf{r}$ is the vector of net reaction rates (also called flux vector or flux distribution). Usually, several biochemical reactions are known to be irreversible which are collected in the index set *Irr*. These reactions can only proceed in forward direction posing sign restrictions on their rates:

$$r_i \geq 0 \; \forall \, i \in Irr \; . \tag{2}$$

Mathematically, the set of flux vectors $\mathbf{r}$ satisfying (1) and (2) form a convex polyhedral cone ("flux cone"). This cone is often analyzed by means of elementary flux modes (Section 4.5). For some reactions we might additionally know minimum/maximum flux capacities (e.g., maximal substrate uptake rates)

$$\alpha_i \leq r_i \leq \beta_i \tag{3}$$

and for some fluxes we might even have measurements

$$r_k = m_k. \tag{4}$$

Combining constraints (1) and (2) with (3) and/or (4) changes the solution space from a cone to a bounded or unbounded (flux) polyhedron. SCBM is based on Eqs. (1)–(4) and employs techniques from linear algebra, linear programming, and computational geometry to analyze the flux space and properties of feasible flux vectors (see also Section 4).

3. Overview of *CellNetAnalyzer*

CellNetAnalyzer (CNA) is a MATLAB toolbox for analyzing biological networks on the basis of topological, stoichiometric, qualitative (logical) and semi-quantitative modeling approaches requiring no or only few (kinetic) parameters. In particular, CNA includes various methods that facilitate an in-depth analysis of metabolic networks based on techniques of SCBM as detailed in Section 4. Functions to study signaling and regulatory networks via interaction (influence) graphs, logical (Boolean) networks, or logic-based ODEs are also included, however, this type of analysis will not be described herein (we refer the reader to Klamt et al., 2007).

The internal architecture of CNA is depicted in Fig. 1. User-created network project(s) are the central objects in CNA. A network project can be of type "mass-flow" (metabolic) or "signal–flow" (signaling/regulatory). Every network project consists of a formal network representation (model) and, optionally, of a graphical user interface (GUI) with one or several *interactive network maps* visualizing the network and allowing interactive input and output of calculated results (Fig. 2). The user can endow a network project with a GUI by providing suitable graphics (bitmap images) of the network, e.g. by using appropriate drawing programs (such as OMIX (Droste et al., 2011, 2013); see also Supplementary Info) or by using maps from other sources such as KEGG. In this way, CNA is very flexible regarding the visual representation of the network. The connection between model and network graphics − resulting in the interactive network maps − is then established by placing input fields (small text boxes) on the network graphics (Fig. 2). Each input field is associated with one network element, for instance, a reaction. The position of the input fields in the network map can be intuitively defined by the user by clicking on the respective position in the network map once a new reaction is defined. The abstract model of a metabolic network is constructed by declaring metabolites and reactions and their respective properties (names, ID, external/internal metabolites, reaction equation, minimal/maximal reaction rates; coefficients in linear objective function, notes, etc.). CNA supports the convenient definition of biomass constituents (proteins, RNA, DNA, etc.). Prior to computations, the biomass composition can be specified by the percentages of the biomass constituents which enables quick adaptation of the stoichiometry of the growth reaction in the metabolic network model. Within the GUI, models can be constructed and edited via a network composer. A new project can also be instantiated by providing the network's stoichiometric matrix, indices of irreversible reactions (Eq. (2)), and flux constraints (Eq. (3)). Furthermore, stoichiometric and constraint-based models can be imported and exported in SBML format (Hucka et al., 2003; including the recently established flux balance constraint package (Olivier and Bergmann, 2015)) or be converted from or to COBRA (Schellenberger et al., 2011) and Metatool (Pfeiffer et al., 1999) models.

Created network projects can be analyzed by the comprehensive toolbox provided with *CellNetAnalyzer* (major functions are described in Section 4). In the GUI, the user may enter, for example, known reaction rates into the respective input fields and then start the calculation by choosing a function from the CNA's menu bar installed in the interactive network maps (Fig. 2). In return, results of calculations are displayed in the network maps. CNA also provides an Application Programming Interface (API) which supports model/GUI access, command-line mode, and batch calculations (Klamt and von Kamp, 2011). In particular, this allows model analysis without the necessity to have a GUI. Most functions provided in the GUI are also supported in command-line mode via the API. In fact, some functions, where a GUI-based workflow is not practical, are only accessible via API. Furthermore, the API allows access and modifications of the model and if the project is endowed with a GUI it can also be used to read/write values from/to the GUI. In this way, the user may program own calculation routines that make use of the abstract model and then display results of these calculations within the network maps. Generally, for some of its calculations, CNA utilizes external packages including linear programming solvers (CPLEX, glpk) and elementary modes calculation routines (efmtool (Terzer and Stelling, 2008) and Metatool (von Kamp and Schuster, 2006)), to which it interfaces via Java and MEX code.

4. Metabolic network analysis with *CellNetAnalyzer*

4.1. Basic network properties

CNA calculates a number of basic network properties, which is especially useful when a new metabolic network model has been created. This includes *conservation relations* and *coupled or blocked reactions*. Conservation relations (CRs) are weighted sums of metabolite concentrations that remain constant in a metabolic reaction network, irrespective of the chosen reaction kinetics. A typical example for a conservation relation in certain metabolic network models is [NADH] + [NAD^+] = CONST. CRs correspond to linearly dependent rows in the stoichiometric matrix $\mathbf{N}$ and CRs can be represented by vectors of