BIOINFORMATICS APPLICATIONS NOTE Vol. 27 no. 1 2011, pages 142–143
doi:10.1093/bioinformatics/btq607

Systems biology

Advance Access publication November 28, 2010

von Bertalanffy 1.0: a COBRA toolbox extension to thermodynamically constrain metabolic models

Ronan M. T. Fleming[1,*] and Ines Thiele[2]

[1]Science Institute and Center for Systems Biology and [2]Center for Systems Biology and Faculty of Industrial Engineering, Mechanical Engineering and Computer Science, University of Iceland, Reykjavik, Iceland

Associate Editor: Trey Ideker

ABSTRACT

Motivation: In flux balance analysis of genome scale stoichiometric models of metabolism, the principal constraints are uptake or secretion rates, the steady state mass conservation assumption and reaction directionality. Here, we introduce an algorithmic pipeline for quantitative assignment of reaction directionality in multi-compartmental genome scale models based on an application of the second law of thermodynamics to each reaction. Given experimental or computationally estimated standard metabolite species Gibbs energy and metabolite concentrations, the algorithms bounds reaction Gibbs energy, which is transformed to *in vivo* pH, temperature, ionic strength and electrical potential.

Results: This cross-platform MATLAB extension to the COnstraint-Based Reconstruction and Analysis (COBRA) toolbox is computationally efficient, extensively documented and open source.

Availability: http://opencobra.sourceforge.net

Contact: ronan.mt.fleming@gmail.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received and revised on September 14, 2010; accepted on October 25, 2010

1 INTRODUCTION

Ludwig von Bertalanffy (1901–1972) sought a theory encompassing 'universal principles applying to systems in general' (von Bertalanffy, 1973) and emphasized the role of thermodynamics in understanding biological systems. In modeling the overall behavior of a genome scale stoichiometric model of metabolism, one of the key constraints is reaction directionality, which is required to follow the second law of thermodynamics. At constant temperature, pressure, pH and ionic strength, this requirement means that the stoichiometrically weighted sum of transformed Gibbs energy of net substrates must exceeds that of net products (Alberty, 2003). With experimentally derived or computationally estimated standard reactant Gibbs energy of formation (Jankowski *et al.*, 2008), if one assumes a typically narrow physiological range of reactant concentrations, one may constrain reaction directionality on a genome scale (Feist *et al.*, 2007; Fleming *et al.*, 2009).

The Constraint-Based Reconsruction and Analysis (COBRA) toolbox (Becker *et al.*, 2007) is a growing suite of MATLAB functions for quantitative prediction of cellular metabolism. Here, we introduce an extension of the COBRA toolbox for quantitative assignment of stoichiometric model reaction directionality via integration of experimentally derived (Alberty, 2003) and group contribution estimates of standard Gibbs energies of formation (Jankowski *et al.*, 2008) and reactant concentrations. Encoded in a single extensively documented pipeline is a thorough thermodynamic treatment of the necessary transformation of such data (Fleming *et al.*, 2009) to *in vivo* temperature as well as compartment specific pH, ionic strength and electrical potential for multi-compartment models.

The pipeline automatically generates extensive statistics and figures, focusing on quantitative reaction directions that disagree with reconstruction directions, manually assigned from literature (Thiele and Palsson, 2010). The COBRA toolbox supports model exchange via the SBML and libSBML (Bornstein *et al.*, 2008), thereby the extension reported herein may be applied to an arbitrary mass and charge-balanced metabolic reconstruction.

Table 1. Data required to establish a thermodynamically constrained stoichiometric model of metabolism

Data	Ref./range
Metabolic reconstruction	(Thiele and Palsson, 2010)
Experimental $\Delta_i G^o$ and $\Delta_i H^o$	(Alberty, 2006)
+/- Group contribution estimate of $\Delta_i G^o$	(Jankowski *et al.*, 2008)
Metabolite structure data file	.mol
Temperature	273–313 K
pH	5 to 9
Electrical potential	(mV)
Ionic strength	0–0.35 M
Metabolomic data	

A group contribution estimate of $\Delta_i G^o$ is necessary if no experimental data are available.

2 METHODS

We have previously documented the necessary biophysical chemistry theory while thermodynamically constraining a genome scale model of *Escherchia coli* metabolism (Fleming *et al.*, 2009). The necessary prerequisite data are indicated in Table 1. Alberty (Alberty, 2006) has established thermodynamic tables of experimentally derived *metabolite species* (i) standard Gibbs energy $\Delta_f G_i^o$, and standard enthalpy $\Delta_f H_i^o$ for 135 *reactants* (j), at 298.15 K, atmospheric pressure and zero ionic strength (See Supplementary Material 1). A reactant is a group of metabolite species differing only in state of protonation. We interface the COBRA toolbox extension with a web-based implementation of a group contribution method tailored to organic molecules (Jankowski *et al.*, 2008). The estimated $\Delta_f G_i^o$, for the most predominant

*To whom correspondence should be addressed.

protonated form of a metabolite species at a particular pH, is used to augment the experimental data.

Initiation of the pipeline is followed by algorithmic checks on the quality of the reconstruction, such as mass and charge balance, identification of reactions that exchange metabolite species with the environment. Certain adjustments to reaction stoichiometry are made in order to maintain thermodynamic consistency. For example, CO_2 dissolved in water is represented thermodynamically as H_2CO_3 so each reaction involving CO_2 must have H_2O added to the opposite side (Alberty, 2003)

Experimentally derived $\Delta_f G_j^o$, augmented by group contribution estimates, are assigned to each compartment-specific reactant. Standard Gibbs energy of formation is usually reported at 298.15 K. We adjust this to a temperature between 298 and 313 K using the van't Hoff equation (Alberty, 2003). The effect of ionic strength on the activity of each metabolite species is taken into account with the extended Debye–Hückle equation (Alberty, 2003). Within a given compartment, Legendre transformations of metabolite species standard Gibbs energy of formation, for specified pH and electrical potential, allow a group of metabolite species to be thermodynamically represented as a single reactant. The hydrogen ion stoichiometric coefficient is adjusted to balance the number of hydrogen atoms in each reaction based on thermodynamic calculation of the average number of hydrogen ions bound by each reactant. In doing so, a compartment-specific adjustment is made to recover H^+ concentration from a measurement of H^+ *activity*. For reactions that transport between compartments at different pH, or electrical potential, thermodynamic consistency is maintained by adjustment to the standard Gibbs energy for the reaction, depending on the number of hydrogen atoms, or the charge, of the transported metabolite species.

The transformed reaction Gibbs energy, and thereby the directionality of each reaction, is assigned by combining standard transformed Gibbs energy with metabolomic data. Due to uncertainty in estimated standard Gibbs energy or lack of metabolomic data, typically many reactions seem quantitatively reversible considering the possible upper and lower bounds on transformed reaction Gibbs energy alone. However, the uncertainty associated with group contribution estimates of standard metabolite species Gibbs energy is normally distributed (Jankowski *et al.*, 2008), so a calculation of the cumulative probability that a reaction is irreversible can be used to stratify such reactions by a measure of confidence that the reaction operates in a particular direction.

Allowing many reactions, irreversible *in vivo*, to be reversible *in silico,* under-constrains metabolism, therefore, in the pipeline, reactions that are irreversible with sufficient confidence, above a certain cutoff, are set to irreversible. The trade-off between under- and over-constraining the feasible set of steady state solutions may be identified by decreasing the cutoff until a reaction direction essential for steady state flux is reversed. Ultimately, quantitative assignment of reaction directionality requires comparison with experimental literature. To aid this process, the pipeline generates a *directionality report*, for each quantitatively assigned reaction direction, which conflicts with a reconstruction direction. This report also includes details on the relative contributions of uncertainty in estimation or concentration for each metabolite involved. Moreover, to further aid comparison with literature, numerous figures are automatically generated that illustrate the stratification of directionality.

3 IMPLEMENTATION

The pipeline is implemented as approximately 14 000 lines of cross platform MATLAB (Mathworks Inc., Natick, MA, USA) code supplemented by approximately 4500 comment lines. The COBRA toolbox is required in addition if it is desired to test for a feasible steady state before and after establishing a thermodynamically constrained model.

4 CONCLUSION

This pipeline facilitates the quantitative thermodynamic assignment of reaction directionality using experimentally derived standard metabolite species Gibbs energies of formation (Alberty, 2003) together with group contribution estimates (Jankowski *et al.*, 2008). The code is appropriate for reaction thermodynamics at biochemically relevant pH, ionic strength and electrical potential ranges. The stoichiometric metabolic reconstruction can be multi-compartmental, but should pass certain quality controls (Thiele and Palsson, 2010), such as elemental and charge balance of each reaction.

Funding: This work was supported by the National Institute of Health, grant 5R01GM057089-11, and United States Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery Through Advanced Computing program, grant DE-SC0002009.

Conflict of Interest: none declared.

REFERENCES

Alberty,R.A. (2003) *Thermodynamics of Biochemical Reactions*. Wiley-Interscience, Hoboken, New Jersey.

Alberty,R.A. (2006) *Biochemical Thermodynamics: Applications of Mathematica*. Wiley-Interscience, Hoboken, New Jersey.

Becker,S.A. *et al*. (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat. Protoc.*, **2**, 727–738.

Bornstein,B.J. *et al*. (2008) Libsbml: an api library for sbml. *Bioinformatics*, **24**, 880–881.

Feist,A.M. *et al*. (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.*, **3**, e121.

Fleming,R.M.T. *et al*. (2009) Quantitative assignment of reaction directionality in constraint-based models of metabolism: application to *Escherichia coli*. *Biophys. Chem.*, **145**, 47–56.

Jankowski,M.D. *et al*. (2008) Group contribution method for thermodynamic analysis of complex metabolic networks. *Biophys. J.*, **95**, 1487–1499.

Thiele,I. and Palsson,B.Ø. (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.*, **5**, 93–121.

von Bertalanffy,L. (1973) *General System Theory: Foundations, Development, Applications*. G. Braziller, New York.