

Supporting Material

1.1 Problem formulation

Thermodynamics Flux Analysis (TFA) adds constraints on top of a classic Flux Balance Analysis (FBA) problem to couple reaction directionalities to thermodynamics constraints. In particular, the formulation in (Soh and Hatzimanikatis, 2014) adds metabolite concentrations and Gibbs energy of reactions, and couples the sign of the Gibbs energy of a reaction to its directionality.

These constraints aim to reduce the feasible flux solution space of the problem and increasing the predictive power of the model. This methodology is used to perform Thermodynamics-based Variability Analysis (TVA), a series of TFA maximization and minimization of the variables in the model, such as reaction fluxes, to determine their allowable ranges and directionalities.

Given a model with specified reaction directionalities, it is possible to characterize the thermodynamic states of the underlying physiology by sampling equilibrium displacements and concentrations. We show here the formulation as proposed in (Soh and Hatzimanikatis, 2014):

FBA constraints	Mass balance	$S \cdot v = 0$
	Flux capacity	$\underline{v} \leq v \leq \bar{v}$
TFA constraints	Gibbs energy of reaction	$\Delta_r G'_i = \Delta_{r, \text{tpt}} G'_i + \sum_{j=1}^m n_{i,j} \mu_j$
	Chemical potential	$\mu_j = \Delta_f G_j'^0 + \Delta_{f, \text{err}} G_j'^0 + RT \ln x_j$
	Thermodynamic feasibility	$\Delta_r G'_i - K + K * z_i < 0$
	Coupling constraint	$v_i - K * z_i < 0$

The TFA problem in the table incorporates thermodynamics-based constraints in the original FBA problem in the two first equations.

For biochemical reactions, the transformed Gibbs free energy of the reaction i , $\Delta_r G'_i$, is a function of the transformed Gibbs energy of the chemical potentials μ_j of the reactants j . If the reaction is a transport of the compounds from one compartment to another, the Gibbs free energy of transport $\Delta_{r, \text{tpt}} G'_i$ is also considered, according to the formulation in Jol *et al.* (Jol, et al., 2010). $\Delta_r G'_i$ is calculated in the third equation.

The chemical potential of the reactants is a function of the standard transformed Gibbs free energy of formation of the compounds $\Delta_f G_j'^0$ and the metabolite's activity, as shown in the fourth equation. Activities of the compounds can be expressed directly as concentrations, as we perform Debye-Huckel correction (Debye and Hückel, 1923). $\Delta_{f, \text{err}} G_j'^0$ is the estimated error in the energy of formation.

K is a large (Big-M, $K > \max \Delta_r G'_i$) value, and z is a binary variable. The two last equations enforce the constraint $\Delta_r G'_i < 0 \Leftrightarrow v_j \geq 0$. K should be chosen so that it is bigger by one or two orders of magnitude than the maximal abs ($\Delta_r G'_i$).

This formulation requires net fluxes to be non-negative. To do so, each reaction is separated in two: a net forward and a net backward, and their net fluxes are associated in the following manner:

$$v_{\text{net}} = v_{\text{forward}} - v_{\text{backward}}$$

In that form, the net forward and the net backward reactions are constrained to have non-negative values. Additional constraints are applied to ensure that at most one of these two is active at a time.

1.2 Usage and Example: sampling thermodynamic displacements

We provide a reduced *E. coli* model made with the software presented in (Ataman, et al., 2017), as well as the model it was generated from, iJO1366 (Orth, et al., 2011).

We can sample the natural logarithms of thermodynamic displacement $\ln(\Gamma)$ for each reaction in the genome-scale model.

In a reaction with one product P and one substrate S, the thermodynamic displacement can be defined as such:

$$\Gamma = \frac{S_{eq} P}{P_{eq} S} = \frac{1}{k_{eq}} \frac{P}{S}$$

S_{eq}, P_{eq} are the concentration at equilibrium of the substrate and product according to the notations in (Heinrich and Schuster, 2012). k_{eq} is the associated equilibrium constant. In that context, we can also write the Gibbs energy of the reaction:

$$\begin{aligned} \frac{\Delta G}{RT} \\ \Delta G = \Delta G^0 + RT \ln\left(\frac{P}{S}\right) \\ \frac{\Delta G}{RT} = \frac{\Delta G^0}{RT} + \ln\left(\frac{P}{S}\right) \end{aligned}$$

Since $\ln\left(\frac{1}{k_{eq}}\right) = \frac{\Delta G^0}{RT}$:

$$e^{\frac{\Delta G}{RT}} = \frac{1}{k_{eq}} \frac{P}{S}$$

Hence,

$$\Gamma = e^{\frac{\Delta G}{RT}}$$

Then:

$$\Gamma = e^{\frac{\Delta G}{RT}} \Leftrightarrow \ln(\Gamma) = \frac{\Delta G}{RT}$$

and hence the directionality of a reaction is opposite to the sign of $\ln(\Gamma)$.

It is possible to directly sample admissible thermodynamics displacements and calculated $\overline{\ln \Gamma}$ the average thermodynamic displacement for each reaction. The thermodynamic displacements are constrained because of their link to $\Delta_r G^0$, which is defined by metabolite concentrations. Thus, admissible displacements depend directly on the concentration ranges.

It is also useful to sample directly admissible concentrations. As an example, here is a sampling performed for cytosolic ATP using the methods provided with the package:

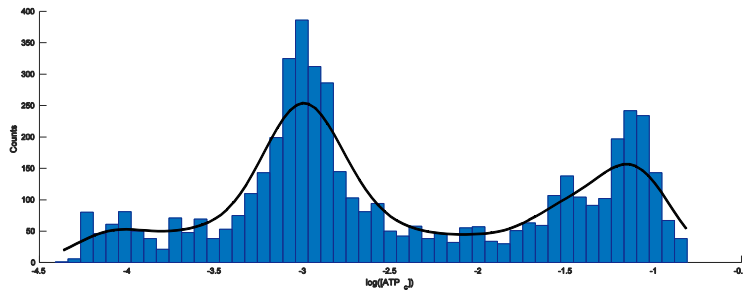


Figure S 1: Example distribution of the sampling of admissible ATP concentrations

1.3 Further Analysis

Given an MILP model with thermodynamics constraints, it is possible to perform several kinds of additional studies.

Thermodynamics-based Variability Analysis (TVA)

TVA can be performed on any variable of the model. These include metabolite concentrations, reaction fluxes, Gibbs free energy of reactions, displacement from equilibrium.

Integration of metabolomics

Integration of metabolomics data is possible because the logarithmic concentration of metabolites is a variable within the model. In particular, it is possible to perform Thermodynamics-based Metabolite Sensitivity Analysis (TMSA) (Kiparissides and Hatzimanikatis, 2017), which allows to define a priority list of metabolites to measure in order to constrain further the model.

Characterization of physiologies

By enumerating bidirectional reactions, and looking at the different solutions spanned by their directionalities, it is possible to characterize the relationship between different flux and physiologies. Additionally, it is possible to observe which reactions are operating close to or far from equilibrium.

Sampling

The resulting constraint-based model is amenable to sampling of any of its variables, such as metabolite concentrations or thermodynamic displacements. pyTFA and matTFA can indeed call COBRA's sampling methods, Artificially-Centered Hit and Run (Schellenberger and Palsson, 2009) and OptGpSampler (Megchelenbrink, et al., 2014). Given a physiology, this allows preparing data for kinetic modeling methods, such as Metabolic Control Analysis-based ORACLE (Miskovic, et al., 2017; Miskovic and Hatzimanikatis, 2010).

1.4 Data

Gibbs free energies of formation ΔG_f^0

Gibbs free energies of formation can be obtained from various data sources, among them:

- ~~Litterature~~Literature, e.g. (Jankowski, et al., 2008)
- eQuilibrator (Flamholz, et al., 2012)
- Databases (eg NIST)
- LCSB also provides support on obtaining these data upon request.

ha formattato: Inglese (Stati Uniti)

1.5 References

- Ataman, M., *et al.* redGEM: Systematic reduction and analysis of genome-scale metabolic reconstructions for development of consistent core metabolic models. *Plos Comput Biol* 2017;13(7).
- Debye, P. and Hückel, E. Zur Theorie der Elektrolyte. I. Gefrierpunktserniedrigung und verwandte Erscheinungen. The theory of electrolytes. I. Lowering of freezing point and related phenomena. *Phys Z* 1923;24:185-206.
- Flamholz, A., *et al.* eQuilibrator—the biochemical thermodynamics calculator. *Nucleic Acids Res* 2012;40(Database issue):D770-775.
- Heinrich, R. and Schuster, S. The regulation of cellular systems. Springer Science & Business Media; 2012.
- Jankowski, M.D., *et al.* Group contribution method for thermodynamic analysis of complex metabolic networks. *Biophys J* 2008;95(3):1487-1499.
- Jol, S.J., *et al.* Thermodynamic Calculations for Biochemical Transport and Reaction Processes in Metabolic Networks. *Biophys J* 2010;99(10):3139-3144.
- Kiparissides, A. and Hatzimanikatis, V. Thermodynamics-based Metabolite Sensitivity Analysis in metabolic networks. *Metab Eng* 2017;39:117-127.
- Megchelenbrink, W., Huynen, M. and Marchiori, E. optGpSampler: an improved tool for uniformly sampling the solution-space of genome-scale metabolic networks. *Plos One* 2014;9(2):e86587.
- Miskovic, L., *et al.* A design–build–test cycle using modeling and experiments reveals interdependencies between upper glycolysis and xylose uptake in recombinant *S. cerevisiae* and improves predictive capabilities of large-scale kinetic models. *Biotechnology for Biofuels* 2017;10(1):166.
- Miskovic, L. and Hatzimanikatis, V. Production of biofuels and biochemicals: in need of an ORACLE. *Trends Biotechnol* 2010;28(8):391-397.
- Orth, J.D., *et al.* A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism-2011. *Mol Syst Biol* 2011;7.
- Schellenberger, J. and Palsson, B.O. Use of randomized sampling for analysis of metabolic networks. *J Biol Chem* 2009;284(9):5457-5461.
- Soh, K.C. and Hatzimanikatis, V. Constraining the flux space using thermodynamics and integration of metabolomics data. *Methods Mol Biol* 2014;1191:49-63.