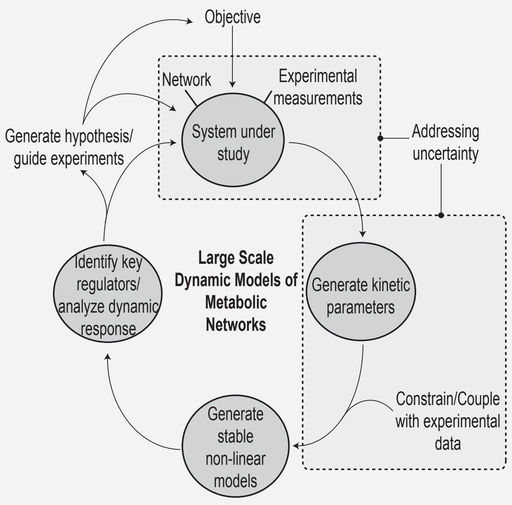
Data, Parameters & Nonlinearities: Development and Applications of Large-scale Dynamic Models of Metabolism

Anirikh Chakrabarti[a,b], Ljubiša Mišković[a,b], Keng Cher Soh[a,b] and Vassily Hatzimanikatis[a,b]

[a] *Laboratory of Computational Systems Biotechnology, EPFL, Lausanne, Switzerland.*

[b] *Swiss Institute of Bioinformatics, Switzerland.*

Abstract



Dynamic nonlinear models of metabolism offer a significant advantage as compared to constraint-based stoichiometric descriptions. However, progress in the development of large-scale nonlinear models has been hindered by both structural and quantitative uncertainties. In particular, the knowledge about kinetic rate laws and their parameters is till today still very limited when compared to the number of stoichiometric reactions known to be present in a large-scale metabolic model. In addition, strategies to systematically identify and implement large-scale dynamic models for metabolism are still lacking. In this contribution, we propose a novel methodology for development of dynamic nonlinear models for metabolism. Using the ORACLE[1] (Optimization and Risk Analysis of Complex Living Entities) framework, we integrate thermodynamics and available omics and kinetic data into a large-scale stoichiometric model. The resulting set of log-linear kinetic models is used to compute kinetic parameters of the involved enzymatic reactions such as the maximal velocities and Michaelis constants. These kinetic parameters are in turn used to compute populations of stable, nonlinear, dynamic models sharing the same stable steady-state as the log-linear ones. The computed models offer unprecedented possibilities for system analysis, e.g. to study the responses of metabolism upon large perturbations; to investigate time course evolutions in and around the steady state; and to identify multiple steady-states and their basins of attraction. We illustrate the features of the generated models in the case of optimally grown *E. coli*, where our analysis of the estimated maximal reaction rates highlights the significance of network thermodynamics in constraining the variability of these quantities.

ORACLE methodology

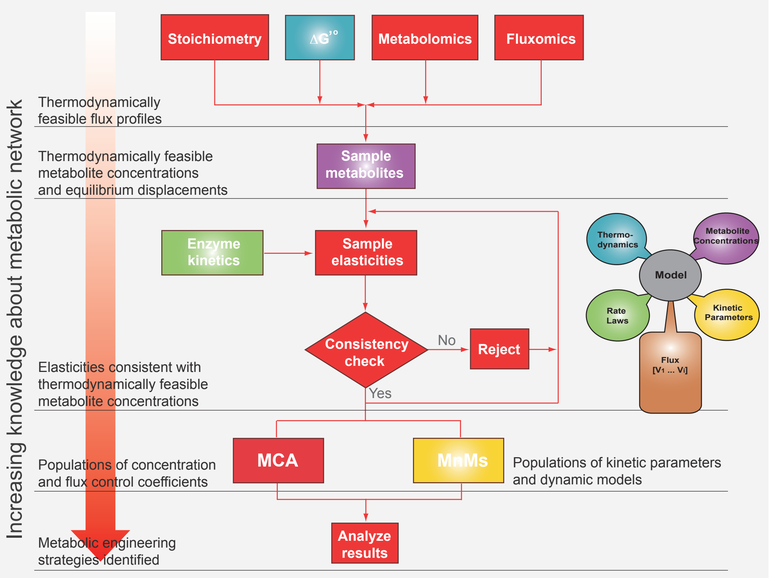


Figure 1: Flowchart of the computational procedure for uncertainty analysis of metabolic networks within the ORACLE framework. The successive application of computational procedures integrates biological information from different levels and sources thus refining kinetic models and providing guidance for metabolic engineering.

ORACLE[1] used to assemble the key aspects defining a non-linear model: thermodynamics, rate laws, metabolite concentrations and kinetic parameters, and coupling with partial/complete data.

Consistently reduced *E. coli* model

Derived from a genome scale *E. coli* reconstruction. Consists of 133 reactions and 77 metabolites. Used to demonstrate the strategy for developing large scale dynamic models.

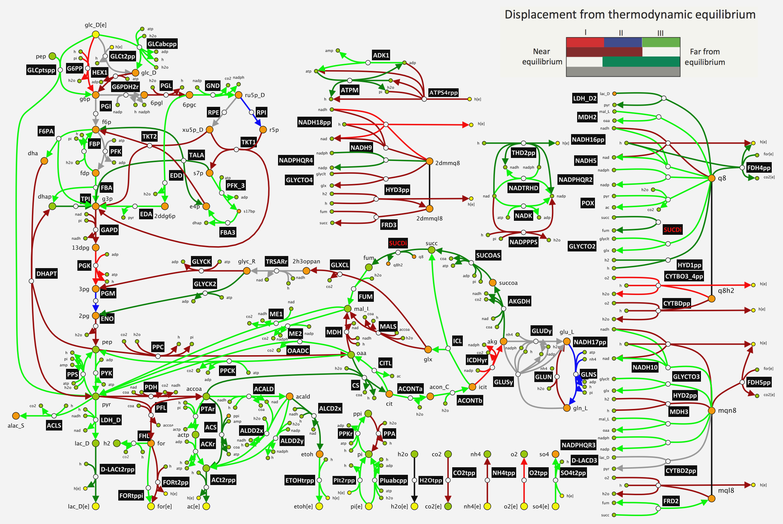


Figure 2: Thermodynamic displacements of the reactions in the consistently reduced *E. coli.* network. The current network included 133 reactions and 77 metabolites.

Role of thermodynamics

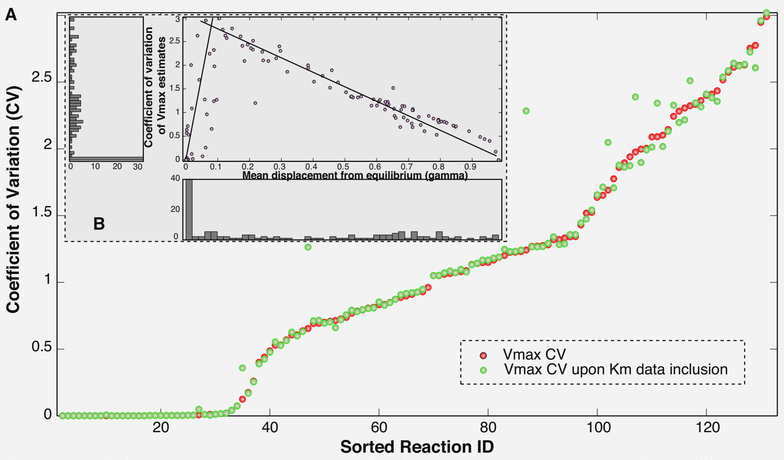


Figure 5: (A) Coefficient of variation (CV) of the V_{max} estimates. For 50% of the reactions CV was < 1. (B) Strong dependence of the CV to the mean of the equilibrium displacement.

We see that, for the optimally grown *E. coli*, the variability of the maximal velocity estimates depends on the displacement from the thermodynamic equilibrium of the involved enzymes.

Kinetic parameter estimation

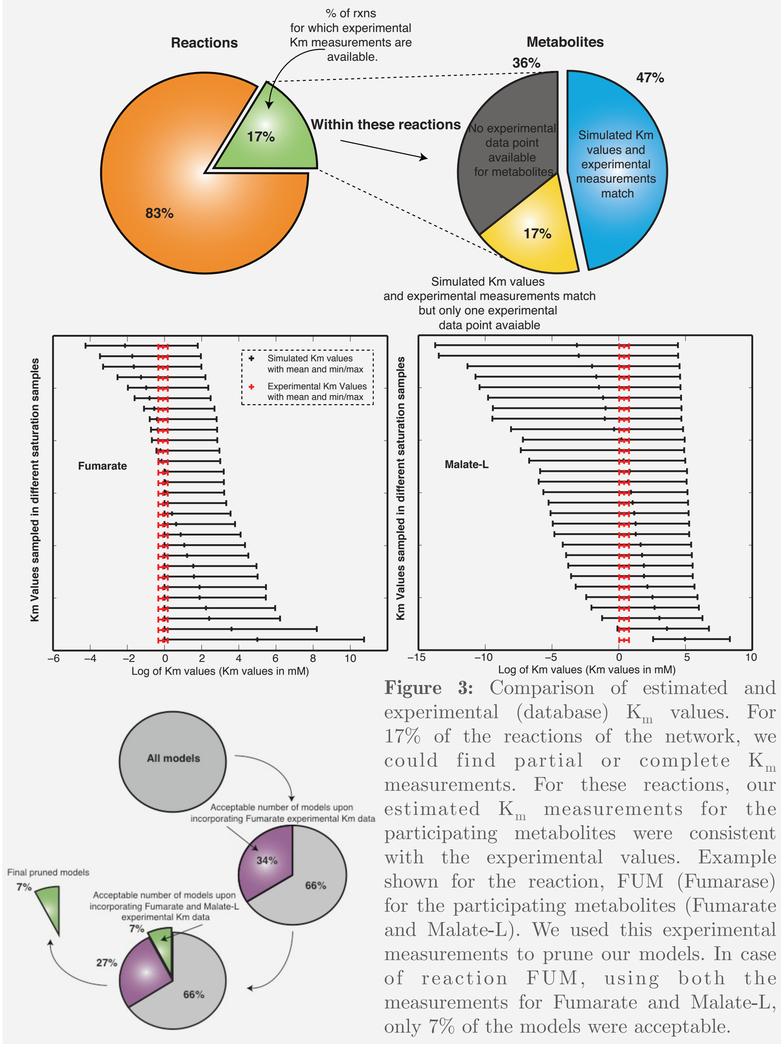


Figure 3: Comparison of estimated and experimental (database) K_m values. For 17% of the reactions of the network, we could find partial or complete K_m measurements. For these reactions, our estimated K_m measurements for the participating metabolites were consistent with the experimental values. Example shown for the reaction, FUM (Fumarase) for the participating metabolites (Fumarate and Malate-L). We used this experimental measurements to prune our models. In case of reaction FUM, using both the measurements for Fumarate and Malate-L, only 7% of the models were acceptable.

Estimated kinetic parameters, e.g. K_m values, are comparable to experimental measurements in databases like BRENDA. We use this partial data whenever available, to prune/refine our parametric estimates.

Parametric relationship mining

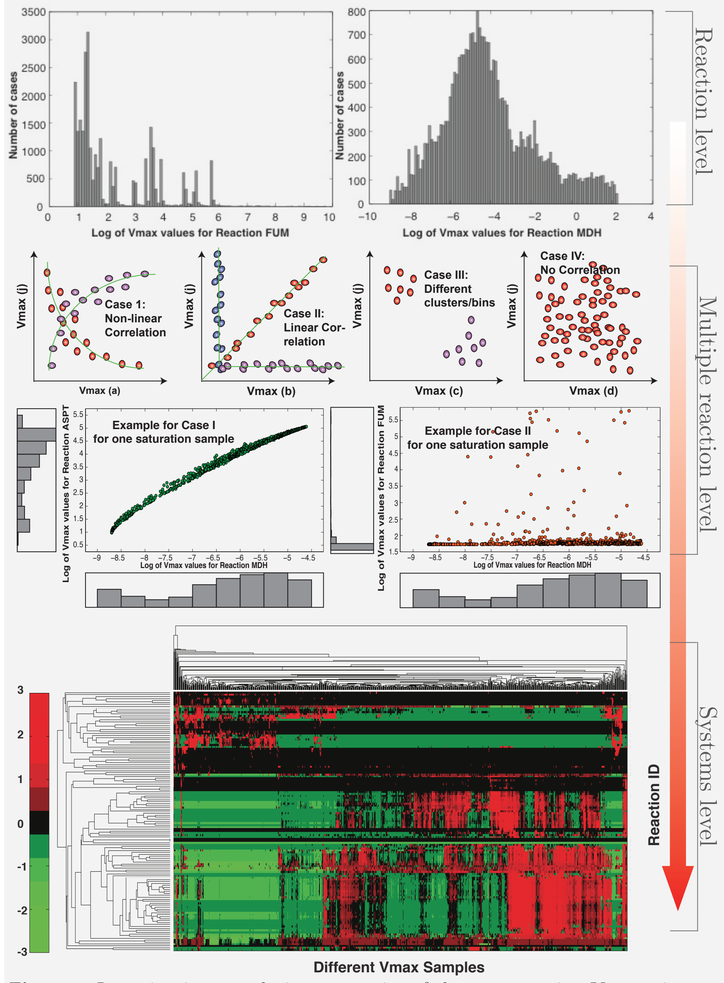


Figure 4: Investigating correlative properties of the system using V_{max} estimates. At a singular reaction level, we clearly see distributions for the parameters. At a multiple reaction level, inherent correlations and relations amongst each other become evident. Example shown for reactions MDH, ASPT and FUM. At a systems level, we can identify potentially correlated parts of the system.

V_{max} estimates are used in dynamic model generation but also to highlight not so evident correlative/co-regulated properties of the network.

Model analysis using dynamic models of metabolic networks

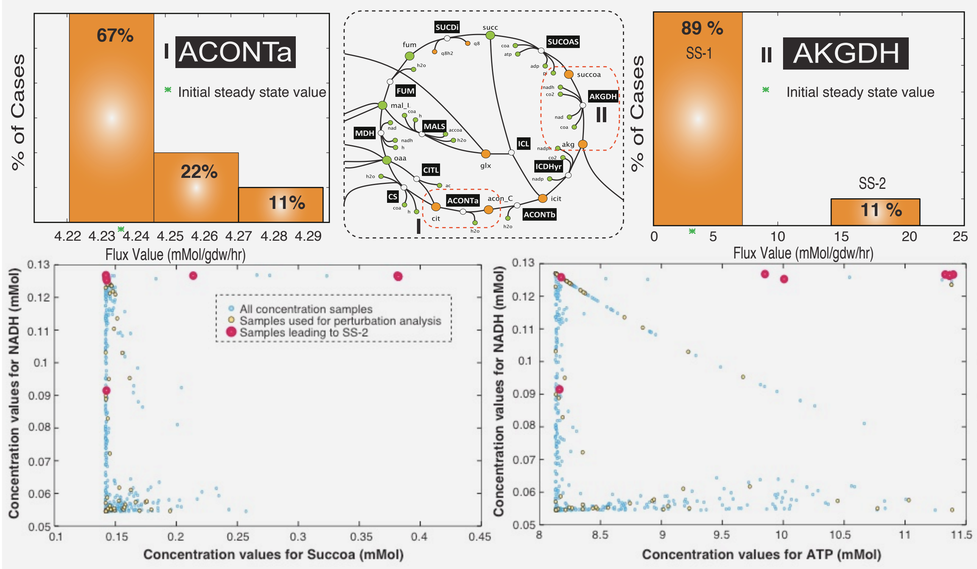


Figure 6: Investigating basins of attraction upon perturbations to the models. In case of reaction ACONTa, the flux value for almost 100% of the cases returns back to the initial steady state flux value of 4.23 mMol/gdw/hr. While for others like AKGDH, there are two basins of attraction; one around 4 mMol/gdw/hr (89% of cases) and the other at 17 mMol/gdw/hr (11% of cases). We can further investigate the concentrations of metabolites (which were the only parameters that were changed) that gives rise to the other steady states.

The non-linear estimations about the stable state can be used to analyze diverse properties of the system upon large perturbations and investigate time course evolutions in and around this steady state.

Conclusion

- ✧ The estimated parameters are consistent with BRENDA and other databases.
- ✧ These parametric estimates are used to investigate correlative/co-regulated properties of the system.
- ✧ The partial kinetic data is incorporated to further refine/constrain the kinetic parameter estimates.
- ✧ The kinetic parameters used to systematically develop populations of stable dynamic models having the same steady-state as the log-linear ones.
- ✧ These non-linear estimations around the stable state are used to analyze diverse properties of the system upon large perturbations and investigate time course evolutions in and around this steady state.

References

[1] Miskovic & Hatzimanikatis. Production of biofuels and biochemicals: in need of an ORACLE. *Trends in Biotech.*, 2010.