Research Article

Towards kinetic modeling of genome-scale metabolic networks without sacrificing stoichiometric, thermodynamic and physiology constraints

Anirikh Chakrabarti [1,2] , Ljubisa Miskovic [1,2] , Keng Cher Soh [1,2] and Vassily Hatzimanikatis [*1,2]

[1] Laboratory of Computational Systems Biotechnology (LCSB),

Swiss Federal Institute of Technology (EPFL),

CH-1015 Lausanne, Switzerland

[2] Swiss Institute of Bioinformatics, CH-1015 Switzerland.

[*] Corresponding author:

Vassily Hatzimanikatis,

Laboratory of Computational Systems Biotechnology (LCSB),

Swiss Federal Institute of Technology (EPFL),

CH-1015 Lausanne, Switzerland

Email: vassily.hatzimanikatis@epfl.ch

Phone: + 41 (0)21 693 98 70

Fax: +41 (0)21 693 98 75

Abstract

Mathematical modeling is an essential tool for a comprehensive understanding of cell metabolism and its interactions with the environmental and process conditions. Recent developments in the construction and analysis of stoichiometric models made it possible to define limits on steady-state metabolic behavior using flux balance analysis. However, detailed information about enzyme kinetics and enzyme regulation is needed to formulate kinetic models that can accurately capture the dynamic metabolic responses. The use of mechanistic enzyme kinetics is a difficult task due to uncertainty in the kinetic properties of enzymes. Therefore, the majority of recent works consider only the mass action kinetics for the reactions in the metabolic networks. In this work, we applied the ORACLE framework and constructed a large-scale, mechanistic kinetic model of optimally grown *E. coli*. We investigated the complex interplay between stoichiometry, thermodynamics, and kinetics in determining the flexibility and capabilities of metabolism. Our results indicate that enzyme saturation is a necessary consideration in modeling metabolic networks and it extends the feasible ranges of the metabolic fluxes and metabolite concentrations. Our results further suggest that the enzymes in metabolic networks have evolved to function at different saturation states to ensure greater flexibility and robustness of the cellular metabolism.

Keywords: Thermodynamics | Mass-action | Saturation | Elasticity | Stability.

Abbreviations:

ORACLE – Optimization and Risk Analysis of Complex Living Entities

TFBA – Thermodynamic Flux Balance Analysis

MA – Mass Action

FDPA – Flux Directionality Profile Analysis

PC – Principal Component

PCA – Principal Component Analysis

TCA – Tri Carboxylic Cycle

CV - Coefficient of Variation

1 Introduction

The rational metabolic engineering of microorganisms for the production of fuels and commodity chemicals requires reliable metabolic network models. These models allow for the interpretation and integration of the biological knowledge and experimental data in the form of mathematical expressions that can be used to identify rate-limiting steps and guide the engineering of carbon, energy, and redox flows in metabolic networks [1]. Furthermore, we can use these models to predict and optimize the dynamic behavior of cells in culture [2]. Nowadays, well-curated genome-scale metabolic reconstructions are available for an ever-growing number of organisms [3-6]. However, these models are primarily used for constraint-based analyses under steady state conditions and the capability to transform them into kinetic models is still being perfected [7-9].

Constraints imposed during modeling must relate to biologically relevant phenotypes. The process of identification and analysis of these constraints itself can enhance our understanding of evolution and capabilities of the modeled organisms. Although there has been significant progress in the field of development of large and genome scale kinetic models recently [9-14], these efforts are hindered by incomplete or missing information about the kinetic properties of enzymes. To overcome this difficulty, some recent studies utilize approximations of kinetic mechanisms by considering only the mass action (MA) term in their model formulation [10, 15].

In a seminal paper by Varma and Palsson, they postulate that it is possible to define limits on metabolic behavior using flux balances, but for more detailed description of metabolic responses, information about enzyme kinetics and their regulation is needed [16]. Motivated by this postulate, we investigated how thermodynamics and *kinetics* can constrain the space of feasible metabolic behavior, represented through feasible metabolic flux and concentration profiles. For example, the space of metabolite concentrations can be constrained by imposing physiological bounds determined by experimental methods [17]. By virtue of the second law of thermodynamics [18-20], which states that a reaction can only take place if the Gibbs free energy of the reaction is negative, there is *thermodynamic coupling* between the metabolite concentrations that further reduces the allowable space. As an additional level of information, we incorporated knowledge about enzyme kinetics. By considering only the mass action term in modeling of the kinetic mechanisms, we show that the effects of enzyme saturation as observed in biological systems cannot be captured. In order to address this issue and explain possible pitfalls in considering only the mass action term, we performed our analyses by considering kinetic models with both: i) simple mass action rate laws and ii) rate laws that consider also the detailed enzyme mechanisms.

Theoretical derivations provided in this study allowed us to compare the differences between mass action and mechanistic enzyme kinetics. We performed this comparison by evaluating the feasible consistent concentration profiles with two types of

kinetics for a representative flux state of the metabolic network. This was followed by a thorough statistical comparison of these two cases for the entire metabolic flux and concentration space. Our results demonstrated that enzyme saturation is an important and necessary consideration in modeling metabolic networks. Our analysis further suggests that the kinetic mechanisms that determine the enzyme saturations have evolved to support higher flexibility and robustness of the cell metabolism and physiology. Though in the current study we used a large-scale *E. coli* model, the presented methodology is readily scalable to genome-scale models of metabolism.

2 Materials and Methods

2.1 The ORACLE methodology: Workflow

The systematic development of thermodynamically feasible kinetic models using ORACLE [21] consists of the following steps (as outlined in Fig. 1):

1. The first step is to **define the stoichiometry** using information from genome-scale reconstructions [4-6] and the **thermodynamic constraints** based on the available information on the Gibbs free energies of reactions [22-24]. We further **incorporate the available fluxomics and metabolomics** data [25-30]. After integrating the above-mentioned information, we apply Thermodynamic Flux Balance Analysis (TFBA) in order to compute the thermodynamically feasible flux profiles [19, 20, 31, 32]. A thermodynamically feasible flux profile is characterized by a unique directionality for

each reaction within a metabolic network. This profile is contained within a convex hull defined by the flux solution space. We sample this convex space and **derive a population of thermodynamically feasible flux profiles**. We perform the principal component analysis on the obtained samples to **find a representative flux**, v_{rep}, of the studied flux profile [33, 34].

2. Next, we **sample the space of metabolite concentrations** consistent with the flux directionalities computed in Step 1, without violating thermodynamic and directionality constraints. If available, estimates of metabolite concentration ranges obtained from experiments under similar physiological conditions are used as the bounds for the computational sampling of the metabolite levels [35]. We then **compute the displacement of the enzymatic reactions from the thermodynamic equilibrium,** which are consistent with the metabolite levels and flux profiles from Step 1.

3. Next, we **integrate kinetic properties of enzymes** from the literature or databases [36, 37]. For those enzymes with incomplete or missing information, we sample either the enzyme states [38], or the degree of the saturation of the enzyme active site [39].

4. We **reject/prune samples** that are either inconsistent to the experimentally measured responses of the metabolic network to gene perturbations and those that do not pass stability test [39].

5. We **compute the populations of control (sensitivity) coefficients** (log-linear, MCA models) [39, 40] of the accepted samples. The control coefficients quantify the sensitivity of the metabolic fluxes and intracellular metabolite concentrations to the activities of each enzyme in the network, to the concentrations of extracellular metabolites or to other parameters.

6. We **analyze the populations of control coefficients** from Step 5 in order to postulate hypotheses about the possible responses of engineered metabolic networks, the couplings within the network, and to identify alternative metabolic engineering strategies.

2.2 The ORACLE methodology: Components

Sampling of the thermodynamically feasible metabolite space

The metabolite concentration space can be defined by the constraints imposed by experimentally measured metabolite concentration levels [27, 30, 41, 42]. However, with additional constraints stemming from the thermodynamic coupling within metabolic network, the allowable metabolite space is further confined. This guarantees the consistency of the metabolite concentrations levels with values of the Gibbs free energy ΔG and with the directionality of the reactions. To illustrate this, consider the example of a simple metabolic network shown in Fig. 2, Panel A. We assumed that the

concentrations of intracellular metabolites, A, B and C, were constrained within the ranges observed in the cell [17, 19]:

$$\underline{A} < lnA < \overline{A} \quad \underline{B} < lnB < \overline{B} \quad \underline{C} < lnC < \overline{C}, \tag{1}$$

where the *underbars* and *overbars* denote lower and upper bounds on the logarithms of the concentration levels, respectively. Thermodynamics also imposes relations (i) between the concentration levels of the intracellular metabolites A, B and C, and of the extracellular metabolites S_e, E_e and L_e,

$$lnA < lnS_e - \frac{\Delta G_1^\circ}{RT} \quad lnB > lnE_e + \frac{\Delta G_3^\circ}{RT} \quad lnC < lnL_e + \frac{\Delta G_5^\circ}{RT} \tag{2}$$

and (ii) between the concentration levels of intracellular metabolites

$$lnA > lnB + \frac{\Delta G_2^\circ}{RT} \quad lnA > lnC + \frac{\Delta G_4^\circ}{RT} \tag{3}$$

where ΔG_i° denotes the standard Gibbs free energy of reactions i ($i = 1 \cdots 5$). The resulting metabolite space was thus reduced to a great extent as compared to the one defined in Eq. 1 (Fig. 2, Panel B). Furthermore, given the fixed directionality of the metabolic fluxes, the constraints imposed by thermodynamics are convex in the logarithmic space of metabolite concentrations, and therefore, the allowable metabolite concentration space remains convex. The convexity of this space allowed us to use the Artificial-Centering Hit-and-Run sampler in the COBRA Toolbox for the efficient sampling of concentrations within this space [43, 44].

Modeling and Simulation of Enzyme Kinetics

In the current study, we assessed the feasibility of metabolic states by verifying the stability of their corresponding kinetic models. As the type of employed kinetic laws affects the shape and the volume of the feasible space of metabolite concentrations and metabolic fluxes, we implemented two general types of kinetics laws: (i) a mass action rate law [9, 45-48]**;** and (ii) mechanistic enzyme kinetics, which considers mass action rate law for each elementary step, and it results in saturation kinetics [36, 47-49]. The employed mechanistic kinetics laws comprised reversible Michaelis-Menten kinetics, ordered Bi-Bi, Bi-Ter, and Ter-Bi, etc. [49]. For more complex reaction mechanisms we used generalized approximations of enzymatic mechanisms such as generalized reversible Hill [50], or convenience kinetics [51].

We examined the local stability of kinetic models by inspecting the eigenvalues of the Jacobian matrix of the system [39]. The elasticities, defined as the sensitivity of reaction rates to changes in parameters or metabolite concentration levels, are the integral component of the Jacobian matrix [39]. Based on this fact, we identified analytically the contributions of the mass action and the enzyme saturation components to the feasibility of metabolic states as discussed below.

Mass Action and Enzyme Saturation

The rate law of a reaction can be expressed in the following form [48, 52]:

$$v = V_{max}\omega\phi \tag{4}$$

where V_{max} is the maximal velocity of the reaction, ω is the mass action component and ϕ the saturation component. A partial derivative with respect to a metabolite x, gives us:

$$\frac{\partial v}{\partial x} = V_m \frac{\partial \omega}{\partial x}\phi + V_m \omega \frac{\partial \phi}{\partial x} \tag{5}$$

Scaling Eq. 5 by x/v, we obtain:

$$\frac{x\partial v}{v\partial x} = \frac{\partial ln\omega}{\partial lnx} + \frac{\partial ln\phi}{\partial lnx} \tag{6}$$

The left hand side of Eq. 6, i.e., the scaled sensitivity of the enzyme rate, is called *enzyme elasticity* [21, 39, 48] and it quantifies the strength of interaction of the enzyme with the corresponding metabolite, x. It can be shown that the magnitude of the elasticity corresponds also to enzyme saturation [38, 39], and it approaches a high number for low saturation and for most common kinetics, it approaches to a small finite number for high saturation. This number depends on the specific mechanism of the enzyme kinetics, and it is equal to 1 for the classical case of irreversible Michaelis-Menten rate law.

The first term on the right hand side of Eq. 6, $\frac{\partial ln\omega}{\partial lnx}$, we define as the *mass-action elasticity*, ε_m, and the second term, $\frac{\partial ln\phi}{\partial lnx}$, as the *saturation elasticity*, ε_s. Thus, the elasticity of a reaction, ε, is a simple sum of the mass-action and saturation components:

$$\varepsilon = \varepsilon_m + \varepsilon_s \tag{7}$$

Illustration of calculation of Mass Action and Saturation elasticities

For a simple Bi-Bi reaction of the form:

$$\mathbf{S_1} + \mathbf{S_2} \Leftrightarrow 2\mathbf{P} \tag{8}$$

where $\mathbf{S}_1$ and $\mathbf{S_2}$ are the substrates, and $\mathbf{P}$ is the product, we can write the corresponding rate equation as:

$$v = \frac{V_{max}(1-\Gamma)\tilde{s}_1\tilde{s}_2}{1+\tilde{s}_1+\tilde{s}_2+2\tilde{p}} \tag{9}$$

where,

$$\tilde{s}_1 = \frac{\mathbf{S_1}}{K_{mS_1}}, \quad \tilde{s}_2 = \frac{\mathbf{S_2}}{K_{mS_2}}, \quad \text{and} \quad \tilde{p} = \frac{\mathbf{P}}{K_{mP}}, \tag{10}$$

and $\Gamma = \frac{1}{K_{eq}}\frac{\mathbf{P}^2}{\mathbf{S_1}\mathbf{S_2}}$ denotes the displacement from thermodynamic equilibrium [38, 48].

Restructuring Eq. 9 in the form of Eq. 4, and using Eq. 6, we obtain the elasticities with respect to $\mathbf{S_1}, \mathbf{S_2}$ and $\mathbf{P}$ as:

$$\frac{\mathbf{S}_1}{v}\frac{\partial v}{\partial \mathbf{S}_1} = \underbrace{\frac{1}{(1-\Gamma)}}_{\varepsilon_m^{S_1}} - \underbrace{\frac{\tilde{s}_1}{1+\tilde{s}_1+\tilde{s}_2+2\tilde{p}}}_{\varepsilon_s^{S_1}} \tag{11}$$

$$\frac{\mathbf{S}_2}{v}\frac{\partial v}{\partial \mathbf{S}_2} = \underbrace{\frac{1}{(1-\Gamma)}}_{\varepsilon_m^{S_2}} - \underbrace{\frac{\tilde{s}_2}{1+\tilde{s}_1+\tilde{s}_2+2\tilde{p}}}_{\varepsilon_s^{S_2}} \tag{12}$$

$$\frac{\mathbf{P}}{v}\frac{\partial v}{\partial \mathbf{P}} = \underbrace{-2\frac{\Gamma}{(1-\Gamma)}}_{\varepsilon_m^P} - \underbrace{\frac{2\tilde{p}}{1+\tilde{s}_1+\tilde{s}_2+2\tilde{p}}}_{\varepsilon_s^P} \tag{13}$$

where $\varepsilon_m^{S_1}$, $\varepsilon_m^{S_2}$ and ε_m^P are the *mass-action elasticity terms* with respect to $\mathbf{S_1}, \mathbf{S_2}$ and $\mathbf{P}$, respectively, whereas $\varepsilon_s^{S_1}$, $\varepsilon_s^{S_2}$ and ε_s^P are the corresponding *saturation terms*. So, the elasticities with respect to metabolites for the mechanistic enzyme kinetics are a function of the displacement from thermodynamic equilibrium, the molecularity of involved metabolites, and the relative concentration of metabolites with respect to their binding affinities (saturation constants).

Enzyme saturation in mechanistic enzyme kinetics

For most of enzymatic reactions, the degree of the saturation of the enzymes active site, A, with the respect of a substrate, S, can be expressed as [39]:

$$\sigma_A = \frac{[AS]}{A_T} = \frac{[S]/K_m}{1+[S]/K_m} \tag{14}$$

where operator $[\cdot]$ denotes the concentration value of the corresponding species, A_T the total active site concentration, AS is the enzyme-substrate complex formed with an on-rate and off-rate constants, k_{on} and k_{off}, respectively, with $K_m = k_{off}/k_{on}$, which is also called "saturation constant". The degree of the saturation σ_A ranges from 0 to 1. So, when an enzyme is operating close to 0% of saturation, we have $\sigma_A \approx 0$, i.e. $[S] << K_m$.

In contrast, when an enzyme is operating close to 100% of saturation, we have $\sigma_A \approx 1$, i.e. $[S] >> K_m$. For $\sigma_A = 0.5$, an enzyme is operating at 50% of saturation and $[S] = K_m$. In general, as we consider complex mechanisms, the exact saturation of the enzymes will be described by more complex expressions [38, 48, 49]. In the current study, we used Eq. 14 to describe the extent of saturation of enzymes by substrates, metabolites, and effectors for complex mechanisms as well.

Metabolic network kinetics: Integration of kinetics into metabolic stoichiometric models

For a metabolic network under study, at a steady state we have:

$$\underline{\underline{N}} \cdot \underline{v} = \underline{0} \tag{15}$$

where $\underline{\underline{N}}$ is the stoichiometric matrix and $\underline{v}$ is the rate vector. This is the set of mass balance constraints used in the Flux Balance Analysis. The main objective here is to identify kinetic models that are consistent with flux profiles that are in turn consistent with these mass balance constraints.

From Eq. 15, the Jacobian, $\underline{\underline{J}}$, for the metabolic system can be derived as follows [53-55]:

$$\underline{\underline{J}} = \underline{\underline{N}} \frac{\partial \underline{v}}{\partial \underline{x}} = \underline{\underline{N}} \underline{\underline{V}} \underbrace{\underline{\underline{V}}^{-1} \frac{\partial \underline{v}}{\partial \underline{x}} \underline{\underline{X}}}_{\underline{\underline{E}}} \underline{\underline{X}}^{-1} \tag{16}$$

where $\underline{\underline{V}}$ and $\underline{\underline{X}}$ are diagonal matrices of the reaction rates and the metabolic concentrations. In Eq. 16, the $\underline{\underline{V}}^{-1}\frac{\partial \underline{v}}{\partial \underline{x}}\underline{\underline{X}}$ term is defined as the *elasticity matrix* $\underline{\underline{E}}$. From Eq. 7, $\underline{\underline{E}}$ can be expressed into the sum of two matrices with the mass-action and the saturation components of the individual reactions: $\underline{\underline{E}} = \underline{\underline{E}}_m + \underline{\underline{E}}_s$. Hence, Eq. 16 becomes:

$$\underline{\underline{J}} = \underline{\underline{N}}\underline{\underline{V}}\underline{\underline{E}}_m\underline{\underline{X}}^{-1} + \underline{\underline{N}}\underline{\underline{V}}\underline{\underline{E}}_s\underline{\underline{X}}^{-1} \quad (17)$$

Therefore, in assessing the stability of kinetic models, we are able to evaluate the stability of the models with the mass action kinetics only as well as the models with mechanistic enzyme kinetics, i.e. the ones that also include saturation components.

The stability test is one of the most important components in the ORACLE workflow. Since we start with a flux profile that is consistent with the mass balance constraints, we implicitly assume that this flux profile is *dynamically stable and at a steady state*. This means the flux values do not change over time for the time range of observation. Therefore, any set of kinetic parameters that is consistent with the metabolic flux profiles and concentration values must also be *asymptotically stable* [48, 56, 57].

Stability scores and bining

The ORACLE methodology as outlined earlier provides us with thermodynamically feasible flux profiles and metabolite concentration samples consistent with the flux profiles. We randomly generated 1000 flux profiles and 1000 metabolite concentration samples for our further analysis. When simulating the case of pure mass action kinetics,

the stability of the kinetic models is evaluated over all the combinations of concentration and flux samples, resulting in an evaluation of 1 million cases. In the case of mass action kinetics, the information of flux distribution and metabolites alone are enough to estimate the stability for different flux and metabolite concentration profiles. In the case of the mechanistic enzyme kinetics, for each of aforementioned 1 million cases it was necessary to evaluate the stability for different realizations of enzyme saturations in order to ensure unbiased sampling of the kinetic space. In the results section, for each of performed analyses we discussed how many enzyme saturation samples are tested per one flux-concentration combination.

In order to evaluate the stability of such a complex and multidimensional flux-concentration space, we defined the following scores of stability:

C-score: Given a flux sample, we evaluated the stability of the kinetic models across all concentration samples (Concentration dependent -score). For example, if 55% of all concentration samples generate stable kinetic models, then the C-score of the analyzed flux sample is 55%. In the case of mechanistic enzyme kinetics, each combination flux-concentration is evaluated across all concentration samples *and* for a number of different realizations of enzyme saturations. In these cases, we reported either *mean* C-score, i.e. C-score averaged over the realizations of enzyme saturations, or *maximal* C-score, i.e. the maximal attained C-score by any of the realizations of enzyme saturations.

V-score: Given a concentration sample, we evaluated the stability of the kinetic models across all flux samples (Flux (V) dependent -score). The V-score, mean V-score and maximal V-score are defined analogously to the C-score, mean C-score and maximal C-score. For example, if 40% of all flux samples give stable kinetic models for a given concentration sample, then the V-score of the analyzed concentration sample would be 40%.

We further categorized the aforementioned flux samples into 10 regions *(as* $\#V1$*,* $\#V2$ *,…,* $\#V10$), based on the L_2 norm of the distance from the representative flux, v_{rep} , in order to analyze specific regions in the flux and concentration space during subsequent studies. The order of the classification was such that region $\#V1$ denoted the hyper-sphere with the radius H_R, enclosing the space surrounding the representative flux sample. Region $\#V2$ was a hyper-layer surrounding region $\#V1$, with the radius $2H_R$, and so on, with region $\#V10$, with the radius $10H_R$, being the farthest from the representative flux sample and surrounding region $\#V9$ (Fig. 3A). Similar categorization into 10 regions was performed for the concentration level samples, based on the L_2 norm of the distance from the representative concentration, C_{rep}, computed in an analogous manner as v_{rep}, with the categorized regions denoted, in the order of distance from C_{rep}, as $\#C1$, $\#C2$,..., $\#C10$. The number of flux samples in each region $\#V1$*,* $\#V2$ *,…,* $\#V10$ were not necessarily the same. The same held true for the concentration samples in each region $\#C1$*,* $\#C2$ *,…,* $\#C10$. In the flux-concentration space, a combination of

flux samples from any of regions $\#V1$–$\#V10$ with concentration samples from any of regions $\#C1$–$\#C10$ defines a *flux-concentration bin*. For example, Bin $\#V6\#C2$ contains the flux samples from region $\#V6$ and the concentration samples from $\#C2$. In fact, the studied flux-concentration space is split into 100 bins.

B-score: We evaluate the stability for all combinations of the concentration and flux samples contained within a bin (Bin-score). For the models with mass action kinetics, B-score represents a percentage of aforementioned combinations giving stable kinetic models. For the models with mechanistic enzyme kinetics, mean B-score represents a percentage of the combinations concentrations-fluxes contained within a bin giving stable kinetic models averaged over the different realizations of enzyme saturations.

3 Results

3.1 Thermodynamically feasible model of *E. coli*

The stoichiometric model used in this study is a core metabolic model of *E. coli.* It was derived using a model reduction algorithm proposed in the thesis of K. C. Soh [34], and based on the latest *E. coli* genome-scale reconstruction by Orth et al. [58]. The reduction algorithm is a consistent and systematic approach to identify the core metabolites and reactions around the central carbon metabolism of *E. coli.* The algorithm starts with the core subsystems such as glycolysis, TCA cycle and pentose-

phosphate pathways and then expands outwards to identify other metabolites and reactions that should be taken into consideration as part of central metabolism. The biomass reaction is then reformulated accordingly based on this set of core metabolites. The resulting core model used in this study, comprised of 146 intracellular reactions (including the biomass reaction) and 90 metabolites (Fig. 4). The model was configured with a fixed glucose uptake of 10 mmol/gDW-hr with a maximum oxygen uptake rate of 20 mmol/gDW-hr as specified in Varma and Palsson [16] and the minimum biomass flux was set to 80% of the optimal solution. We allowed the product fluxes to be redistributed to different byproducts based on sampling of the solution flux space. We further integrated the fluxes from experimental data for the optimal growth conditions from [27] and the metabolite concentrations (including uncertainty) using Thermodynamics-based Flux Balance Analysis (TFBA) [31, 32] (Supporting material, excel file, mat files and text files). As both TFBA and flux balance analysis can yield alternate optimal solutions, we further applied Flux Directionality Profile Analysis (FDPA) to systematically identify the possible flux and corresponding thermodynamic profiles [34]. Out of the two flux directionality profiles obtained, differing only in the directionality of the fumarase (FUM) reaction, we chose the profile with FUM having net flux in the direction of conversion of fumarate and water into malate for subsequent analysis. We continued by generating the feasible flux samples for the chosen flux directionality profile. Then, we performed the principal component analysis (PCA) on the generated set of flux samples to find the representative flux, v_{rep} [33]. The

representative flux vector, v_{rep}, is collinear with the first principal component (PC1) and its magnitude is equal to the mean of magnitudes of sampled flux vectors projections onto PC1. The representative flux, v_{rep}, was used in the subsequent studies.

3.2 Stability landscapes for pure mass-action and mechanistic enzyme kinetics

First, we investigated the stability properties of the models with mass action kinetics. Next, we investigated whether or not the saturation components of the enzymatic mechanisms, within metabolic networks, have an impact on the overall stability of the resulting large-scale kinetic models. For that purpose, the kinetic mechanisms of the reactions in the whole network were modeled using: (i) pure mass-action kinetics; (ii) mechanistic enzyme kinetics wherein it was assumed that all enzymes in the network operate at 50% of their saturation, referred subsequently as *half-saturation* (details in Materials and Methods).

We evaluated the proportion of concentration samples out of the 1000 profiles sampled (see Materials and Methods) that gave stable kinetic models for the representative flux, v_{rep}, and for both types of kinetics. In the case of pure mass action kinetics, 410 concentration samples gave stable kinetic models, i.e. C-score of 41%, whereas in the case of mechanistic enzyme kinetics all 1000 concentration samples resulted in stable kinetic models, i.e., we obtained a C-score of 100% (see Materials and Methods). This

suggests that **the saturation kinetics contribute significantly to the stability of certain flux *realizations*.**

We further compared these two sets of models by exploring the stability landscape for all one million combinations of 1000 flux and 1000 concentration space samples (see Materials and Methods). In the case of mechanistic enzyme kinetics, since all enzymes operate at half-saturation, only one sample of enzyme saturation was used in simulations. Therefore, we assessed the stability of 1 million models for each of the enzyme kinetics cases. We found that for the case of pure mass action, $\approx 97\%$ of the flux samples had a C-score $\leq 5\%$ (Fig. 5, panel A). This implies that for any of the approximately 970 flux profiles, we found less than 50 concentration profiles that made that flux profile stable. In Fig. 5, we chose two reactions and two metabolite concentrations based on their coefficient of variation (CV) values to illustrate the findings. The CV values denote a variation of estimates of flux/metabolite samples [59]. Each of the chosen flux and concentration samples belonged to the top or the bottom quartile of the corresponding CV ranges. More specifically, for the chosen fluxes, CV value of fructose-biphosphate aldolase was 0.23, representing a low variation of these estimates, whereas it was 0.85 for that of fructose 6-phosphate aldolase, representing a high variation of the estimates. Similarly, for the chosen metabolite concentration levels, the CV of citrate was 0.02, while CV of fumarate was 0.34 (Supplementary Fig. 1).

The C-score quantifies for a flux sample the percentage of stability of the kinetic models across all concentration samples (see Materials and Methods). In the case of pure mass

action kinetics, the maximal attained C-score of all flux samples was approximately 52%. For the kinetic models with mechanistic enzyme kinetics, the stability increased considerably and also consistently for each and every flux sample (Fig. 5 B). We found that 7% of the flux samples had a C-score ≥ 50%, with around 5% of the flux samples having a C-score of 100%.

We further analyzed stability as a function of the concentration level samples. When all enzyme kinetics followed pure mass action law, it was observed that all of the concentration samples had a V-score ≤ 5% (Fig. 5C). This implies that for every concentration profile, we found less than 50 stable flux profiles. Similar to our analysis of the flux space, when we used mechanistic enzyme kinetics for all enzymes, the stability increased for all the concentration samples. Specifically, 100% of the concentration samples had a V-score ≥ 5%, and 2% of concentration samples had a V-score above 10%.

While we observed consistently higher stability scores for both the flux and the concentration samples in the case of models with mechanistic enzyme kinetics, we have not been able to identify contiguous iso-stability regions (connected regions in the concentration or flux space having similar stability scores) by inspecting 2-dimensional cross-plots (Fig. 5 A, C). While the main reason for this phenomenon is the high dimensionality of the flux and concentration space, in depth future investigation could yield significant insights into the function and physiology of the metabolic pathways.

To identify the aforementioned iso-stability regions in the high-dimensional flux and/or concentration space, we clustered the flux and the concentration samples into 100 bins and quantified the stability B-score (details in Materials and Methods). B-score quantifies the percent of stability for all combinations of the concentration and flux samples contained within a bin of flux and concentrations samples. Analysis of the stability of flux-concentration bins indicated very limited variability of B-scores as a function of flux samples as compared to a function of concentration samples (Fig. 3C). This observation was true for both types of kinetic models (Fig. 3B, C). More precisely, it appears that certain concentration regions had a decisive influence on the stability scores. In the case of pure mass action kinetics models, the highest B-scores were obtained for bins that involved the concentration region $\#C8$. Similarly, for the models with mechanistic enzyme kinetics, the highest mean B-scores were obtained for bins involving the concentration regions $\#C3$, $\#C6$ and $\#C7$. The highest B-score obtained for the case of mass action kinetics was always $< 2\%$. In comparison, for mechanistic enzyme kinetics models, the mean B-scores obtained were significantly superior with the highest one being $\approx 12\%$, whereas the lowest one being $\approx 4\%$. Once again, these results suggest that the saturation kinetics of enzymes contribute to the stability of the system, and they allow a wider range of kinetically feasible concentration profiles.

3.3 Relationships between enzyme efficiency and stability

We established in the previous section that kinetic models of metabolic networks, with all their enzymes operating at half-saturation, had higher stability than the ones with the pure mass action. In the following section, we studied the consistency of this observation through different degrees of enzyme saturation. We constructed kinetic models with enzymes operating in ranges of 0-10%, 10-20%, ..., 90-100% of their saturation. Near 100% saturation here represented a case of an enzyme being completely saturated and near 0% saturation represented an enzyme operating in the linear range. While such extreme cases might not be physiologically meaningful, as enzymatic reactions in a biological system generally do not operate within such a narrow range of saturation, this analysis provides a theoretical understanding of the effect of enzyme saturations on stability. The comparison was performed on the same set of 1000 flux samples and 1000 concentration samples as considered previously. Therefore, for the mechanistic enzyme kinetics models and for each range of enzyme saturation, one hundred enzyme saturation realizations were generated. Therefore, for the mechanistic enzyme kinetics models, we analyzed 100 million alternative kinetic models, whereas for the mass action kinetics 1 million models were generated and analyzed.

We observed an increase in stability upon inclusion of enzyme saturation (Fig. 6 and Supplementary Fig. 2). Specifically, as the enzyme saturation in the network increased, we had a clear and consistent increase in the stability over the entire surface of mean B-scores. Ultimately, for the range of 90-100% of enzyme saturation, mean B-scores values of 30% were attained. This observation is in line with the findings reported by Bennett

et. al., where they report that large majority (83%) of enzymes are more than 50% saturated for the case of glucose-fed, exponentially growing *E. coli* [27]. Furthermore, by inspecting the contour plots of Fig. 6, we found that, as already observed in Figs. 3B and 3C, the stability was dominantly affected by the concentrations as witnessed by the contour lines mostly parallel to the flux axis. However, the maxima of the mean B-score surfaces were not coinciding, implying that stability is a complex interplay between kinetics, concentrations and fluxes. In other words, a precise knowledge about metabolite concentrations within the network alone does not necessarily contribute to the identification of kinetic models consistent with the data.

3.4 Stability of enzymes operating through the entire range of saturation

In nature, enzymes can operate in a wide range of regimes, ranging from the linear phase to the one with enzyme being completely saturated. In the subsequent analysis, we generated randomly 1,000 vectors of enzyme saturation regimes for the entire system. For each enzyme saturation vector, we computed the stability for 1 million combinations of flux and concentration samples, i.e., in total we analyzed 1 billion kinetic models. For all the flux samples, maximal C-scores of $\geq 95\%$ were attained (Supplementary Fig. 3). On an average, stability of the flux space increased considerably when enzymes operated through the entire range of saturation. This observation was based on the fact that 40% of flux samples had a C-score of $\geq 30\%$, as

compared to the case with pure mass action laws, where 0.8% of flux samples had a C-score of ≥ 30% (compare Fig. 7A and Fig. 5A).

Overall, we obtained significantly higher stability for the entire concentration space upon inclusion of full range of enzyme saturation states (compare Fig. 7C and Fig. 5C). We obtained 12% of concentration samples, with a mean V-score of ≥ 30% averaged over 1,000 enzyme saturation regimes, as compared to none for the mass action. This implies that, for each of the 12% of the concentration profiles, we could find more than 300 kinetic models for the whole network that could explain/realize the corresponding concentration and flux profiles. In comparison, none of these profiles could be explained using mass action laws for each enzyme. Even more striking, the minimum observed V-score in the case of mechanistic enzyme kinetics was 10% over the entire set of 1 million combinations of concentration and flux profiles, as compared to maximum observed V-score of 5% in case of pure mass action.

Upon the inclusion of full range of enzyme saturations, the stability increased significantly within individual bins/ranges of flux and concentration profiles, with B-scores reaching above 70% (Fig. 7D). The maximal stability scores (B-scores) were obtained for Bins $\#V4\#C7$, and $\#V6\#C7$, $\#V6\#C8$ with the values of 72%, 73.4% and 73.3% respectively. This observation suggests that certain combinations of fluxes and metabolites are more probable to appear in this biological system. Their high stability ranking implies that there exist wider ranges of kinetic parameters that could realize these fluxes and concentrations.

3.5 Consideration of experimental kinetic information improves stability statistics

The ORACLE methodology allows us to incorporate experimentally available kinetic information of enzymes from literature databases such as BRENDA. We integrated the available Michaelis constants, K_m, from BRENDA and for the enzymes whose information was missing, we sampled the corresponding enzyme saturation space (see Materials and Methods). While we observed several samples with low stability scores, in general the stability scores of the flux and concentration samples increased compared to the case without inclusion of available kinetic data (Fig. 7B). For approximately 2% of the flux samples, we obtained a *mean C-score* of $\geq$ 40% and for almost 100% of the flux samples, a *maximum C-score* of 100% was obtained. This is an important observation, as it suggests that enzymes might have evolved to operate within a broad range of stable concentration and flux profiles. This flexibility can be achieved with the proper tuning of the enzyme activities through genetic regulation.

4 Discussion

Varma and Palsson suggested in their work that, in addition to flux balances, kinetic information was needed to determine feasible metabolic capabilities that a cell possesses [16]. In this study, we present for the first time an approach to determine the metabolic

capabilities defined by kinetic properties, using our large-scale kinetic models that take into account *all* the stoichiometric and thermodynamic constraints of the flux balance models. We further investigated, whether or not there are clear boundaries between feasible and infeasible metabolic states, using concepts of nonlinear systems stability.

We studied the kinetic properties of metabolic networks, using metabolic modeling which accounts for *all* the stoichiometric and thermodynamic constraints of metabolic networks, without imposing any simplifying assumption that could involve reduction in the number of reactions and metabolites, and that would violate thermodynamic constraints. We used our modeling framework ORACLE that allowed us to efficiently and consistently explore a vast range of concentration and flux profiles, and consider *every* possible kinetic mechanism for the enzymes in the metabolic network.

We found that using kinetic models with mass action kinetics not only makes a rough approximation of real kinetic mechanisms, but in addition it results in a very low incidence of stability, which could potentially lead to erroneous conclusions. This implies that assessing the feasibility of concentration and flux profiles using mass action kinetics, can lead to overly conservative assessments thus neglecting concentration and flux profiles that are likely to correspond to a physiological condition of the system. On the contrary, we observed that upon incorporation of enzyme saturation terms, the feasibility region in the flux and concentration space of resulting kinetic models was consistently bigger. Moreover, these stable regions included the concentration and flux profiles that were found to be feasible with the pure mass action kinetics as well. In

addition, upon inclusion of experimentally measured Michaelis constants, K_m, the region of the flux-concentration space providing stable kinetic models was consistently bigger as well. The fact that enzyme saturation terms and inclusion of experimentally observed kinetic data consistently increased the stability of the kinetic models, i.e. their feasible space of fluxes and concentrations increased, indicates that enzymes have evolved in a way to increase the flexibility and thus the viability and adaptability of living organisms. We further examined whether or not there were clear boundaries between feasible and infeasible metabolic states, and we investigated if we could transition continuously within flux and concentration spaces through stable states. Although we were unable to find any distinct iso-stability regions in the flux or the concentration space, we found smaller continuous regions of concentration and fluxes with a *higher probability* of having high stability scores. For example, the region enclosed by Bin $\#V4\#C7$ or Bin $\#V6\#C7$ in Fig. 7D is likely to have a high stability score.

Different incidence of stability throughout the flux-concentration space, as witnessed by the different B-scores in the defined bins, indicated that there might exist subspaces, not necessarily enclosed by the categorized flux and concentrations regions that have a high stability score. Identification of these regions will require the use of methods for parameter continuation [60] and nonlinear optimization [61]. As a first attempt, in the current study, we split the flux/concentration space using hyper-planes and assessed the stability scores of the resulting sub-regions to uncover these subspaces (Fig. 8). We identified one sub-region with a very high stability score. Using this technique of cutting

planes it could be possible to narrow down the regions having a high probability of high stability. Having information about regions in the flux-concentration space which are more stable, i.e. the ones that might have a higher probability of being experimentally observed, would allow us to postulate and test hypotheses about the physiology of the studied metabolic networks.

The work presented here, is the first attempt to systematically explore the kinetic space of large-scale and genome-scale metabolic networks, without any simplifying assumptions that omit metabolic stoichiometry and thermodynamic constraints. Our results demonstrated that such exploration is possible using the ORACLE methodology and high-performance computations. The studies presented here represent the first building block of a method for the development of large-scale nonlinear metabolic models. Such models will allow us to analyze metabolic networks without sacrificing accuracy in their representation, and it will enable us to design metabolic engineering strategies based on the analysis of the responses of metabolism to large changes in enzyme activities.

Acknowledgements

A.C. was supported by SystemsX.ch, The Swiss Initiative in Systems Biology, through project BattleX (Contract number 2009/004 Nr 4232). K.C.S. was supported by the Swiss National Science Foundation and by SystemsX.ch, The Swiss Initiative in Systems

Biology, through project MetaNetX (Contract number 2009/009 Nr 4202). L.M and V.H. were supported by funding from the Ecole Polytechnique Fédérale de Lausanne (EPFL).

The authors have declared no conflict of interest.

5 References

[1] Cameron, D. C., Tong, I. T., Cellular and Metabolic Engineering - An Overview. *Appl Biochem. Biotech*. 1993, 38, 105-140.

[2] Gerdtzen, Z., Non-linear reduction for kinetic models of metabolic reaction networks. *Metab. Eng.* 2004, 6, 140-154.

[3] Rocha, I., Forster, J., Nielsen, J., Design and application of genome-scale reconstructed metabolic models. *Methods Mol. Biol.* 2008, 416, 409-431.

[4] Henry, C. S., Zinner, J. F., Cohoon, M. P., Stevens, R. L., iBsu1103: a new genome-scale metabolic model of Bacillus subtilis based on SEED annotations. *Genome Biol.* 2009, 10.

[5] Feist, A. M., Herrgard, M. J., Thiele, I., Reed, J. L., Palsson, B. O., Reconstruction of biochemical networks in microorganisms. *Nat. Rev. Microbio*. 2009, 7, 129-143.

[6] Mo, M. L., Palsson, B. O., Herrgard, M. J., Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC Syst. Biol.* 2009, 3, 37.

[7] Terzer, M., Maynard, N. D., Covert, M. W., Stelling, J., Genome-scale metabolic networks. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 2009, 1, 285-297.

[8] Price, N., Reed, J., Palsson, B., Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.* 2004, 2, 886-897.

[9] Jamshidi, N., Palsson, B., Formulating genome-scale kinetic models in the post-genome era. *Mol. Syst. Biol.* 2008, 4.

[10] Jia, G., Stephanopoulos, G., Gunawan, R., Ensemble Kinetic Modeling of Metabolic Networks from Dynamic Metabolic Profiles. *Metabolites* 2012, 2, 891-912.

[11] Haverkorn van Rijsewijk, B. R., Nanchen, A., Nallet, S., Kleijn, R. J., Sauer, U., Large-scale 13C-flux analysis reveals distinct transcriptional control of respiratory and fermentative metabolism in Escherichia coli. *Mol. Syst. Biol.* 2011, 7, 477.

[12] Smallbone, K., Simeonidis, E., Swainston, N., Mendes, P., Towards a genome-scale kinetic model of cellular metabolism. *BMC Syst. Biol.* 2010, 4.

[13] Varner, J., Ramkrishna, D., Mathematical models of metabolic pathways. *Curr. Opin. Biotechnol.* 1999, 10, 146-150.

[14] Tran, L. M., Rizk, M. L., Liao, J. C., Ensemble Modeling of Metabolic Networks. *Biophys. J.* 2008.

[15] Jamshidi, N., Palsson, B. O., Mass Action Stoichiometric Simulation Models: Incorporating Kinetics and Regulation into Stoichiometric Models. *Biophys. J.* 2010, 98, 175-185.

[16] Varma, A., Palsson, B. O., Metabolic capabilities of Escherichia coli: I. synthesis of biosynthetic precursors and cofactors. *J. Theor. Biol.* 1993, 165, 477-502.

[17] Westerhoff, H. V., Welch, G. R., Enzyme Organization and the Direction of Metabolic Flow - Physicochemical Considerations. *Curr. Top. Cell. Regul.* 1992, 33, 361-390.

[18] Kummel, A., Panke, S., Heinemann, M., Putative regulatory sites unraveled by network-embedded thermodynamic analysis of metabolome data. *Mol. Syst. Biol.* 2006, 2.

[19] Henry, C. S., Broadbelt, L. J., Hatzimanikatis, V., Thermodynamics-based metabolic flux analysis. *Biophys. J.* 2007, 92, 1792-1805.

[20] Beard, D. A., Qian, H., *Chemical biophysics : quantitative analysis of cellular systems.* Cambridge texts in biomedical engineering 2008, xviii, 311 p.

[21] Miskovic, L., Hatzimanikatis, V., Production of biofuels and biochemicals: in need of an ORACLE. *Trends Biotechnol.* 2010, 28, 391-397.
[22] Jankowski, M., Henry, C., Broadbelt, L., Hatzimanikatis, V., Group Contribution Method for Thermodynamic Analysis of Complex Metabolic Networks. *Biophys. J.* 2008, 95, 1487-1499.
[23] Goldberg, R. N., Tewari, Y. B., Bhat, T. N., Thermodynamics of enzyme-catalyzed reactions - a database for quantitative biochemistry. *Bioinformatics 2004*, 20, 2874-2877.
[24] Alberty, R. A., Biochemical thermodynamics. *Biochim. Biophys. Acta* 1994, 1207, 1-11.
[25] Wiechert, W., 13C metabolic flux analysis. *Metab. Eng.* 2001, 3, 195-206.
[26] Otero, J. M., Nielsen, J., Industrial Systems Biology. *Biotechnol. Bioeng.* 2010, 105, 439-460.
[27] Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., et al., Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. Biol.* 2009, 5, 593-599.
[28] Boer, V. M., Crutchfield, C. A., Bradley, P. H., Botstein, D., Rabinowitz, J. D., Growth-limiting Intracellular Metabolites in Yeast Growing under Diverse Nutrient Limitations. *Mol. Biol. Cell* 2010, 21, 198-211.
[29] Garcia, D. E., Baidoo, E. E., Benke, P. I., Pingitore, F., et al., Separation and mass spectrometry in microbial metabolomics. *Curr. Opin. Microbiol.* 2008, 11, 233-239.
[30] Rabinowitz, J. D., Cellular metabolomics of Escherchia coli. *Expert Rev. Proteomics* 2007, 4, 187-198.
[31] Henry, C., Jankowski, M. D., Broadbelt, L. J., Hatzimanikatis, V., Genome-scale thermodynamic analysis of Escherichia coli metabolism. *Biophys. J.* 2006, 90, 1453-1461.
[32] Soh, K. S., Hatzimanikatis, V., Network thermodynamics in the post-genomic era. *Curr. Opin. Microbiol.* 2010, 13, 350-357.
[33] Jolliffe, I., *Principal component analysis*, Springer, New York 2002.
[34] Soh, K. S., *Computational Studies on Cellular Bioenergetics*. Thèse École polytechnique fédérale de Lausanne EPFL 2013, 5560.
[35] Soh, K. C., Miskovic, L., Hatzimanikatis, V., From network models to network responses: integration of thermodynamic and kinetic properties of yeast genome scale metabolic networks. *FEMS Yeast Res.* 2012, 12, 129-143.
[36] Teusink, B., Passarge, J., Reijenga, C. A., Esgalhado, E., et al., Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.* 2000, 267, 5313-5329.
[37] Schomburg, I., Chang, A., Placzek, S., Sohngen, C., et al., BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. *Nucleic Acids Res.* 2013, 41, D764-772.
[38] Miskovic, L., Hatzimanikatis, V., Modeling of uncertainties in biochemical reactions. *Biotechnol. Bioeng.* 2011, 108, 413-423.
[39] Wang, L., Birol, I., Hatzimanikatis, V., Metabolic Control Analysis under Uncertainty: Framework Development and Case Studies. *Biophys. J.* 2004, 87, 3750-3763.
[40] Wang, L., Hatzimanikatis, V., Metabolic engineering under uncertainty. I: Framework development. *Metab. Eng.* 2006, 8, 133-141.
[41] Albe, K. R., Butler, M. H., Wright, B. E., Cellular Concentrations of Enzymes and Their Substrates. *J. Theor. Biol.* 1990, 143, 163-195.

[42] Chassagnole, C., Noisommit-Rizzi, N., Schmid, J., Mauch, K., Reuss, M., Dynamic modeling of the central carbon metabolism ofEscherichia coli. *Biotechnol. Bioeng.* 2002, 79, 53-73.
[43] Becker, S., Feist, A., Mo, M., Hannum, G., et al., Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat. Protoc.* 2007, 2, 727-738.
[44] Schellenberger, J., Richard, Q., Ronan, F., Ines, T., et al., Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* 2011, 6, 1290-1307.
[45] Famili, I., Mahadevan, R., Palsson, B. O., k-Cone analysis: determining all candidate values for kinetic parameters on a network scale. *Biophys. J.* 2005, 88, 1616-1625.
[46] Higgins, J., Theory of Oscillating Reactions. *Ind. Eng. Chem.* 1967, 59, 18-&.
[47] Reich, J. G., Sel'kov, E. i. E., *Energy metabolism of the cell : a theoretical treatise,* Academic Press, London ; New York 1981.
[48] Heinrich, R., Schuster, S., *The Regulation of Cellular Systems.* Chapman & Hall, New York 1996.
[49] Segel, I. H., *Enzyme Kinetics.* John Wiley & Sons, New York 1975.
[50] Hofmeyr, J., Cornish-Bowden, A., The reversible Hill equation: how to incorporate cooperative enzymes into metabolic models. *Comp. Appl. Biosci.* 1997, 13, 377-385.
[51] Liebermeister, W., Klipp, E., Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor. Biol. Med. Modeling* 2006, 3.
[52] Rohwer, J. M., Hofmeyr, J.-H. S., Kinetic and Thermodynamic Aspects of Enzyme Control and Regulation. *J. Phys. Chem.* B 2010, 114, 16280-16289.
[53] Hatzimanikatis, V., Floudas, C. A., Bailey, J. E., Analysis and design of metabolic reaction networks via mixed-integer linear optimization. *Aiche J.* 1996, 42, 1277-1292.
[54] Hatzimanikatis, V., Bailey, J. E., MCA has more to say. *J. Theor. Biol.* 1996, 182, 233-242.
[55] Hatzimanikatis, V., Bailey, J. E., Effects of spatiotemporal variations on metabolic control: Approximate analysis using (log)linear kinetic models. *Biotechnol. Bioeng.* 1997, 54, 91-104.
[56] Strogatz, S. H., *Nonlinear dynamics and Chaos : with applications to physics, biology, chemistry, and engineering,* Addison-Wesley Pub., Reading, Mass. 1994.
[57] Heinrich, R., Rapoport, S. M., Rapoport, T. A., Metabolic-Regulation and Mathematical-Models. *Prog. Biophys. Mol. Biol.* 1977, 32, 1-82.
[58] Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., et al., A comprehensive genome-scale reconstruction of Escherichia coli metabolism. *Mol. Syst. Biol.* 2011, 7, 535.
[59] Lequieu, J., Chakrabarti, A., Nayak, S., Varner, J. D., Computational Modeling and Analysis of Insulin Induced Eukaryotic Translation Initiation. *Plos Comput. Biol.* 2011, 7.
[60] Allgower, E. L., Georg, K., *Introduction to numerical continuation methods,* Society for Industrial & Applied Mathematics, Philadelphia, Pa. 2003.
[61] Floudas, C. A., *Nonlinear and mixed-integer optimization : fundamentals and applications,* Oxford University Press, New York ; Oxford 1995.

Figures

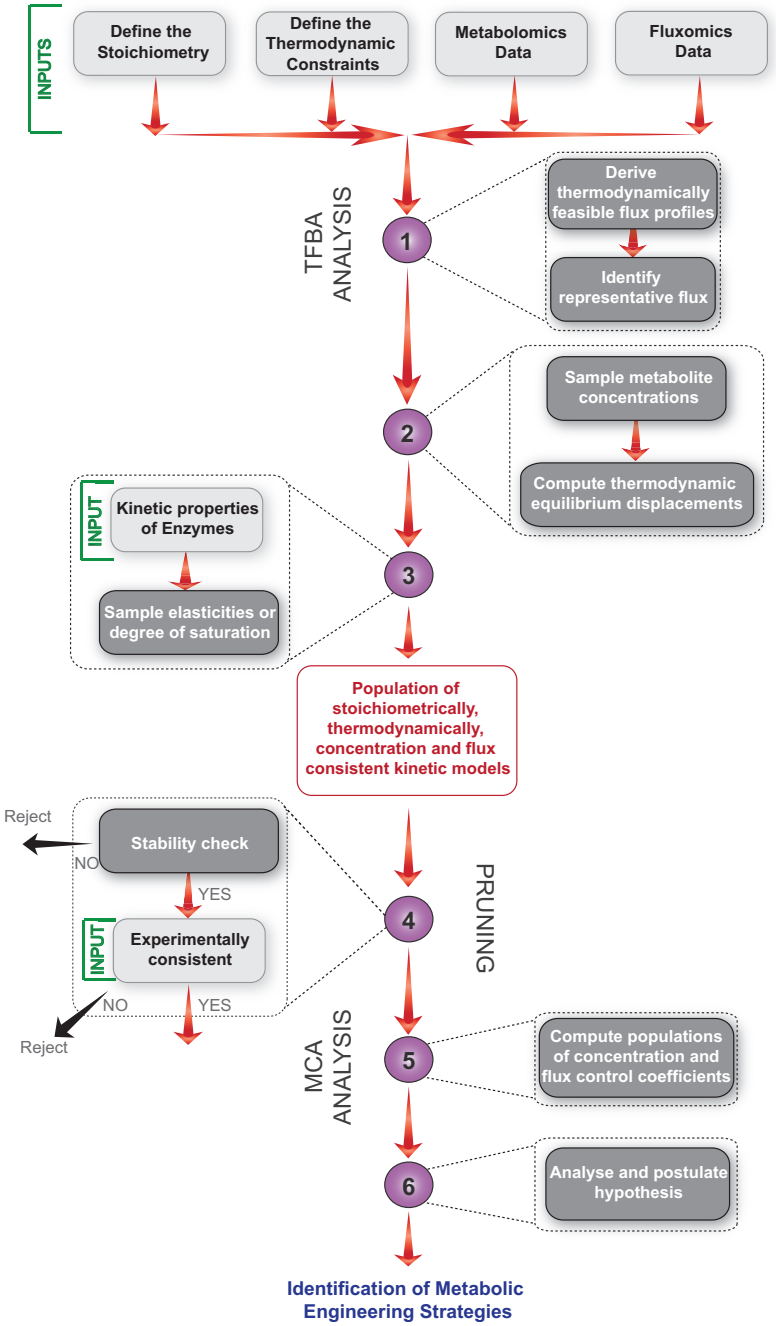

Fig. 1: Workflow of the computational procedure for uncertainty analysis of metabolic networks within the ORACLE framework. Light gray boxes denote the integration of

available experimental and literature data, whereas the dark gray boxes denote the computational procedures.

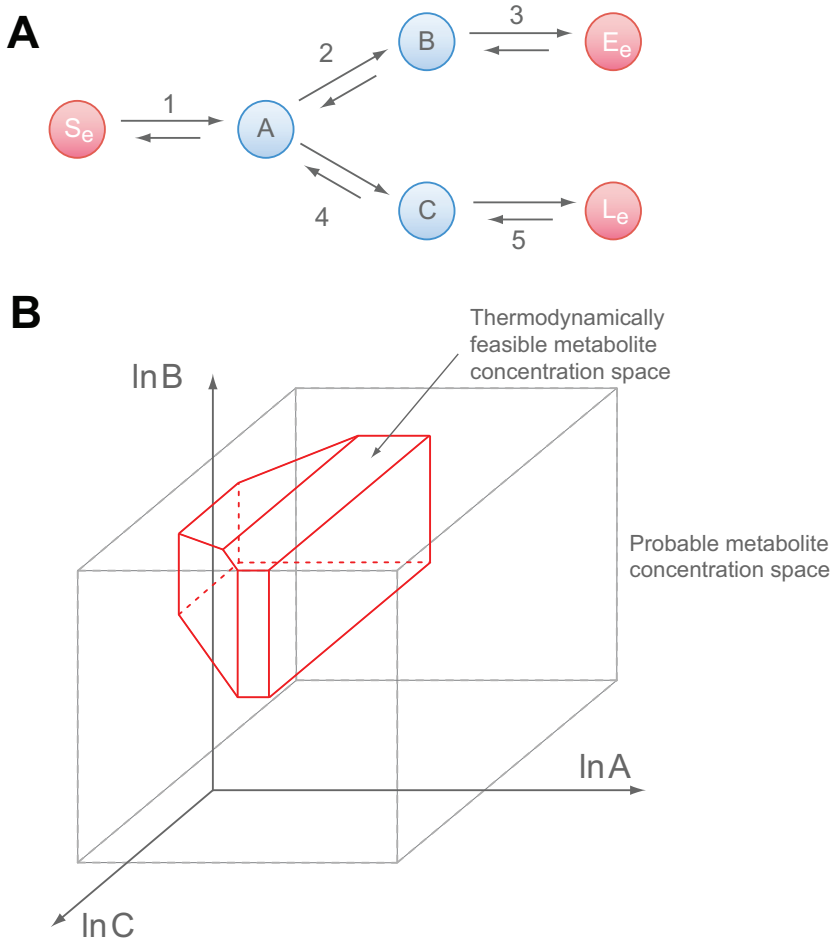

Fig. 2: Effects of integrating the thermodynamic constraints on the space of allowable metabolite concentrations. (A) Branched pathway model where S_e, E_e, and L_e denote the external metabolites, while A, B, and C denote the internal metabolites. (B) Space of metabolite concentrations within bounds experimentally observed under different physiological conditions (dashed line), and within thermodynamically feasible bounds (solid line).

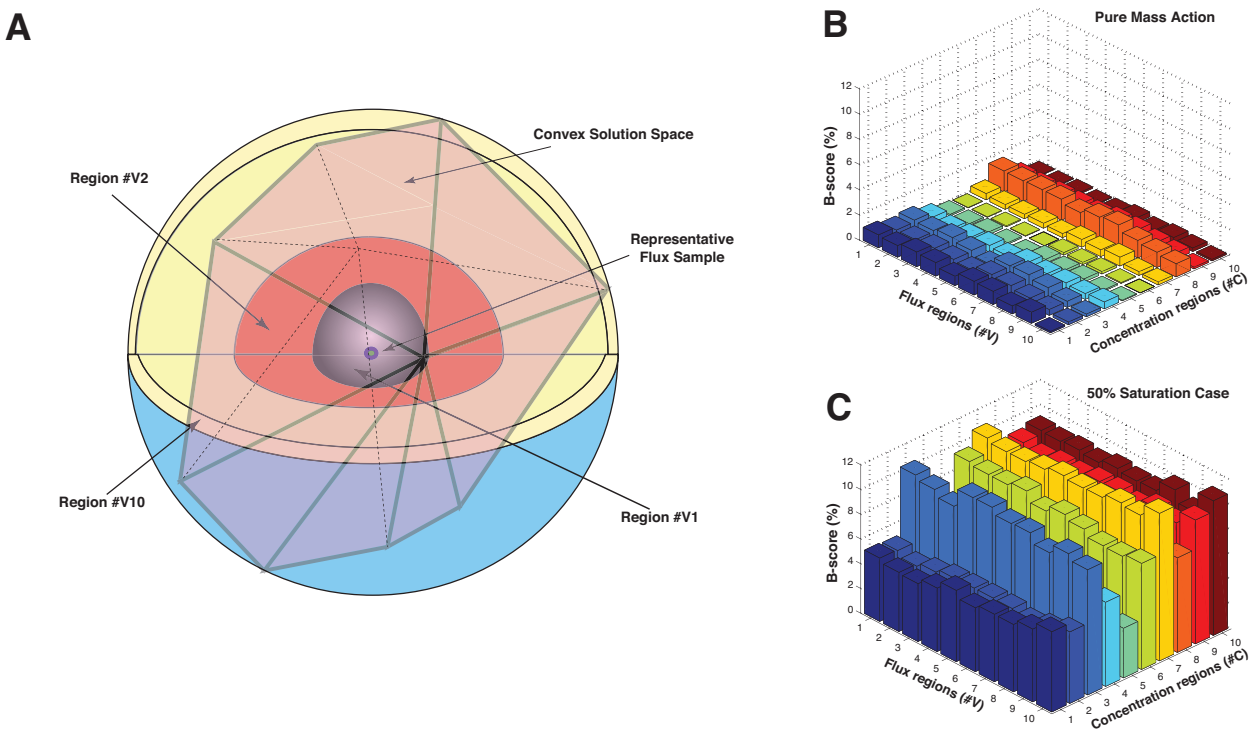

Fig. 3: (A) Bining of flux convex space. (B) Bar graph of the B-score showing the stability in the flux and concentration space for the case of pure mass action. (C) Bar graph of the B-score showing the stability in the flux and concentration space for enzyme kinetics with 50% saturation.

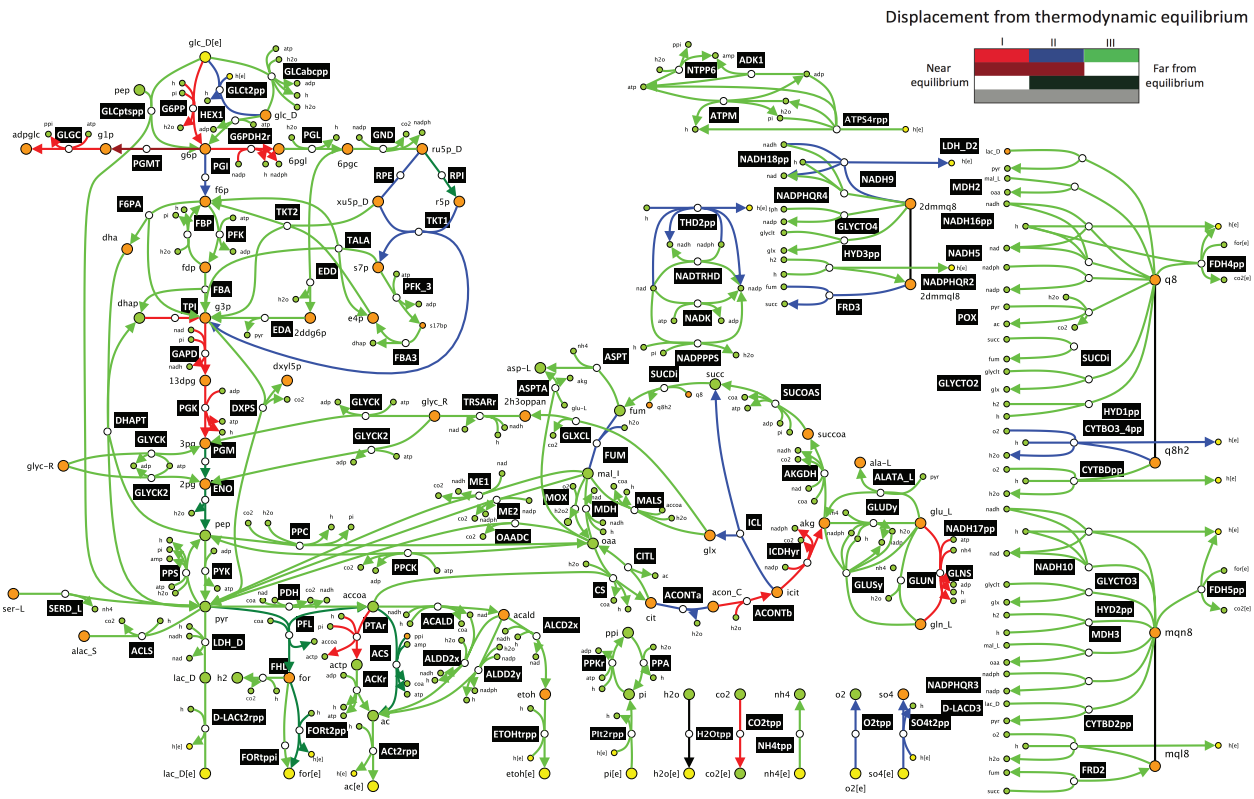

Fig. 4: Metabolic pathways of the consistently reduced *E. coli* network. The network includes 146 reactions and 90 metabolites, and it is fully balanced even with respect to the small molecules such as CO_2, NH_3, and PO_4, and with respect to protons and electrons.

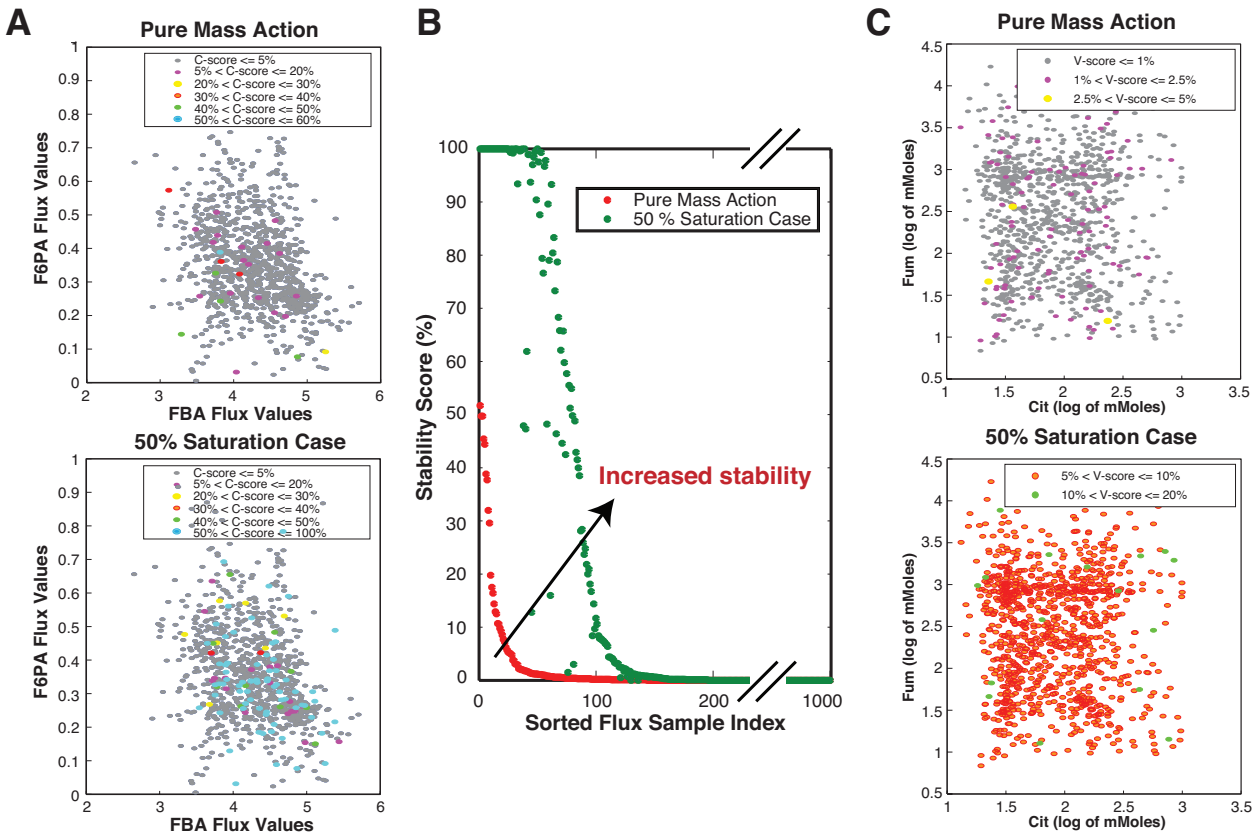

Fig. 5: Stability landscape of fluxes and concentrations in case of pure mass action and 50% enzyme saturation case: (A) Cross-plots of flux samples of fructose-biphosphate aldolase (FBA) versus fructose 6-phosphate aldolase (F6PA). (B) There is a consistent increase in stability for all flux samples for enzyme kinetics with 50% saturation. (C) Cross-plot of concentration samples of citrate (Cit) versus fumarate (Fum).

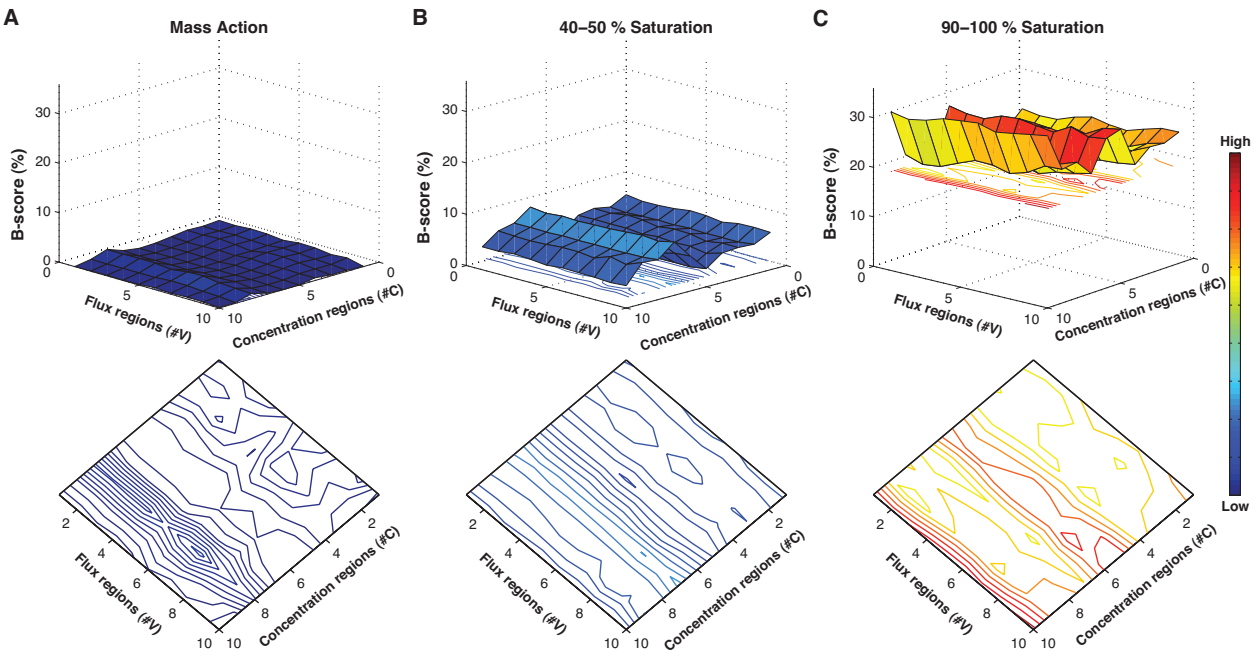

Fig. 6: Stability surfaces and the corresponding contour plots in the flux and concentration space for: (A) pure mass action, (B) 40-50% enzyme saturation levels and (C) 90-100% enzyme saturation levels.

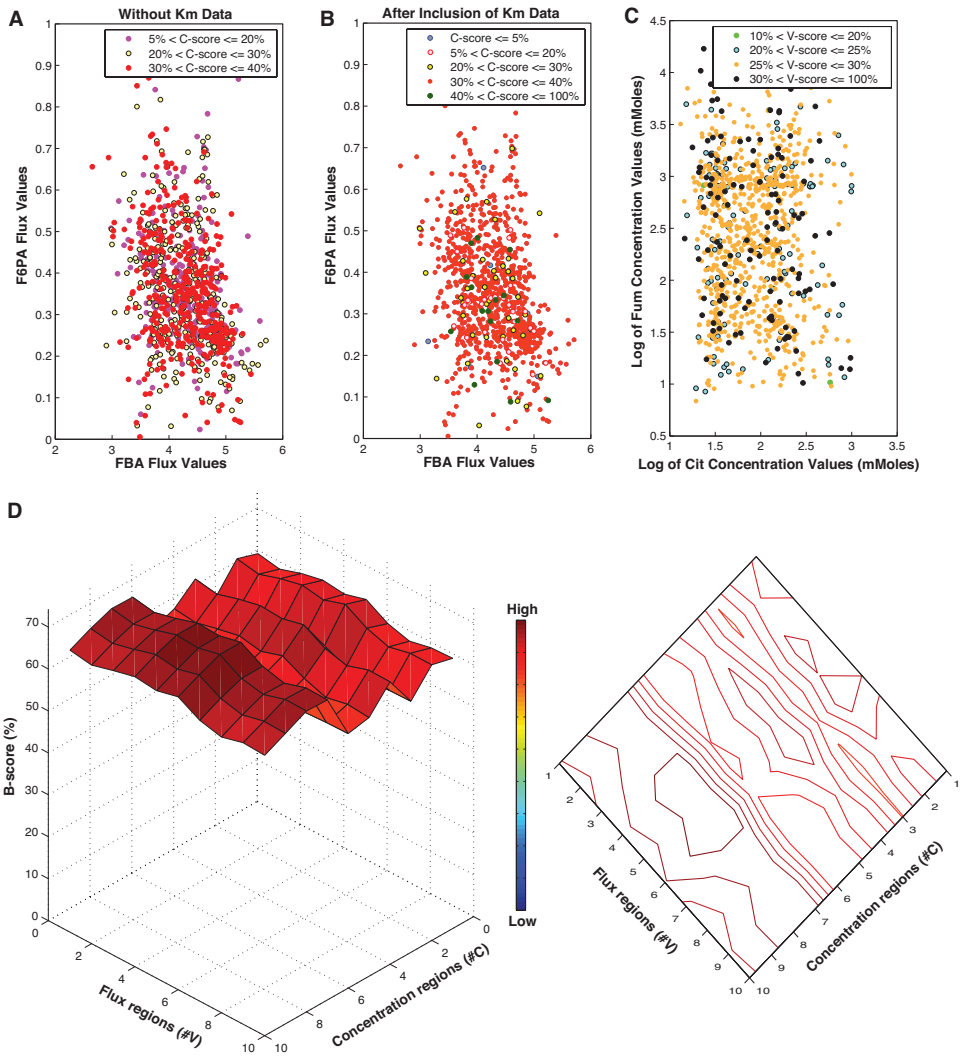

Fig. 7: Stability landscape of flux and concentration samples in case of full range of enzyme saturation levels (0-1). (A) *Mean* C-score cross-plot of flux samples of FBA versus F6PA calculated over 1,000 concentration samples and over 1,000 saturation samples, and without inclusion of any experimental/database K_m measurements. (B) *Mean* C-score cross-plot of the flux samples, with experimental/database kinetic information of K_m measurements incorporated. (C) V-scores cross-plot of concentration samples for citrate versus fumarate. (D) Stability surface, and the corresponding contour plot, of the flux and concentration space.

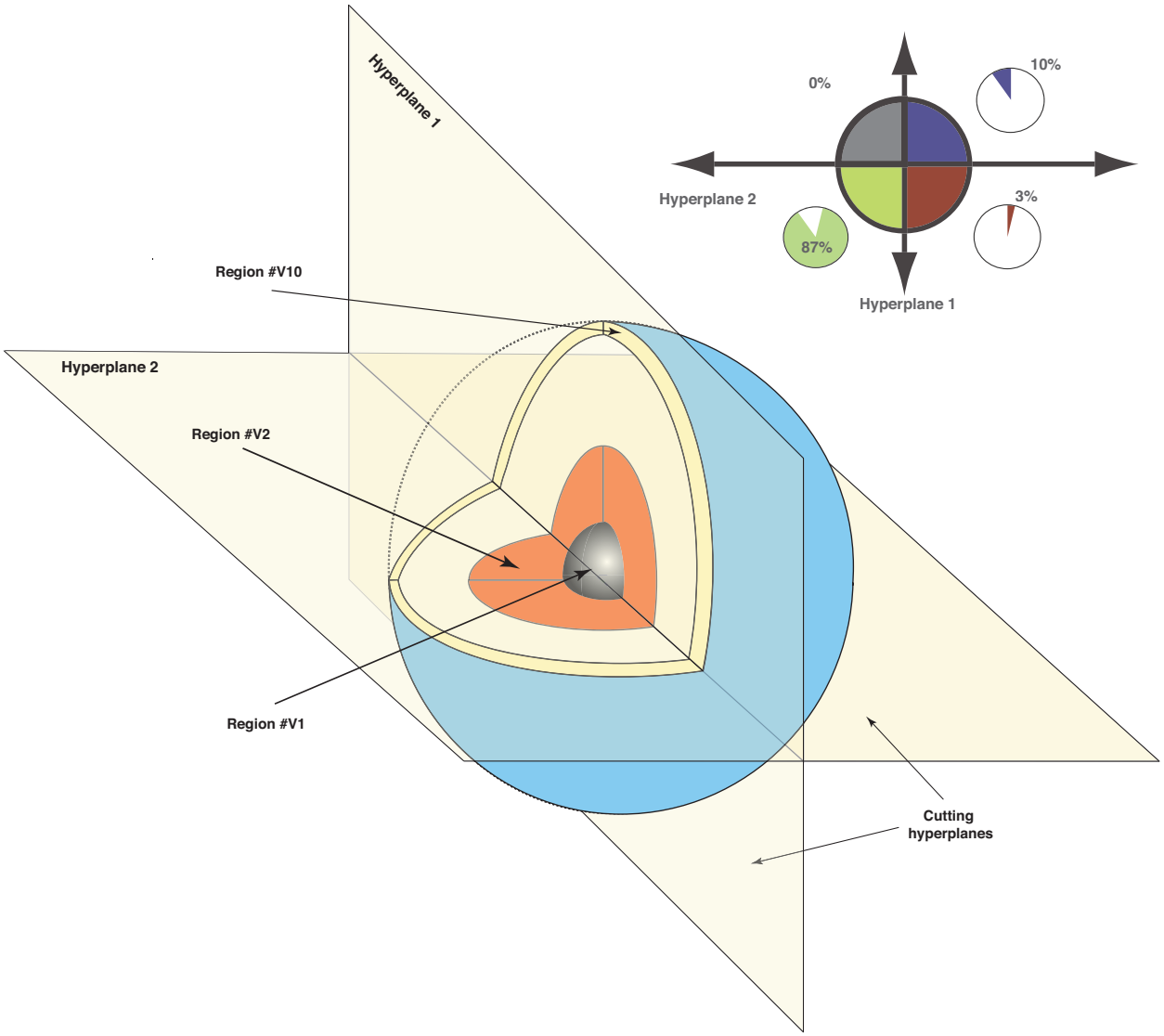

Fig. 8: Partitioning of the flux/concentration space using hyper planes to identify sub-regions within the flux/concentration space with higher probability of stability scores.