

k-OptForce: Integrating Kinetics with Flux Balance Analysis for Strain Design

Anupam Chowdhury, Ali R. Zomorodi, Costas D. Maranas*

Department of Chemical Engineering, Pennsylvania State University, University Park, Pennsylvania, United States of America

Abstract

Computational strain design protocols aim at the system-wide identification of intervention strategies for the enhanced production of biochemicals in microorganisms. Existing approaches relying solely on stoichiometry and rudimentary constraint-based regulation overlook the effects of metabolite concentrations and substrate-level enzyme regulation while identifying metabolic interventions. In this paper, we introduce k-OptForce, which integrates the available kinetic descriptions of metabolic steps with stoichiometric models to sharpen the prediction of intervention strategies for improving the bio-production of a chemical of interest. It enables identification of a minimal set of interventions comprised of both enzymatic parameter changes (for reactions with available kinetics) and reaction flux changes (for reactions with only stoichiometric information). Application of k-OptForce to the overproduction of L-serine in *E. coli* and triacetic acid lactone (TAL) in *S. cerevisiae* revealed that the identified interventions tend to cause less dramatic rearrangements of the flux distribution so as not to violate concentration bounds. In some cases the incorporation of kinetic information leads to the need for additional interventions as kinetic expressions render stoichiometry-only derived interventions infeasible by violating concentration bounds, whereas in other cases the kinetic expressions impart flux changes that favor the overproduction of the target product thereby requiring fewer direct interventions. A sensitivity analysis on metabolite concentrations shows that the required number of interventions can be significantly affected by changing the imposed bounds on metabolite concentrations. Furthermore, k-OptForce was capable of finding non-intuitive interventions aiming at alleviating the substrate-level inhibition of key enzymes in order to enhance the flux towards the product of interest, which cannot be captured by stoichiometry-alone analysis. This study paves the way for the integrated analysis of kinetic and stoichiometric models and enables elucidating system-wide metabolic interventions while capturing regulatory and kinetic effects.

Citation: Chowdhury A, Zomorodi AR, Maranas CD (2014) k-OptForce: Integrating Kinetics with Flux Balance Analysis for Strain Design. PLoS Comput Biol 10(2): e1003487. doi:10.1371/journal.pcbi.1003487

Editor: Daniel A. Beard, University of Michigan, United States of America

Received: August 19, 2013; **Accepted:** January 10, 2014; **Published:** February 20, 2014

Copyright: © 2014 Chowdhury et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors gratefully acknowledge funding from the NSF (<http://www.nsf.gov/>) award no. EEC-0813570 and the DOE (<http://www.energy.gov/>) grant no. DE-SC10822882. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: costas@psu.edu

Introduction

Bio-production is emerging as a competitive strategy for the production of a wide range of chemicals ranging from biofuels, precursor chemicals and bioactive molecules (see [1–3] for detailed reviews). The use of metabolic modeling and computations is increasingly becoming instrumental in deciding how to engineer the production strain [4–11]. Computational strain design generally involves solving an optimization problem which optimizes a specific performance requirement (e.g., maximum flux of desired product) while minimizing the total number of genetic alterations in the metabolic model. Depending on the adopted description of metabolism strain design computational tools could be broadly categorized as based on stoichiometry-alone or kinetic models of metabolism [12].

Kinetic models of metabolism require quantitative expressions that link reaction fluxes and metabolite concentrations. A system of ordinary differential equations (ODEs) is typically solved to obtain the time-dependent variation in metabolite concentrations and reaction fluxes. Different forms of mechanistic expressions have been used extensively such as Michaelis-Menten or Hill Kinetic expressions [13,14]. These expressions require *a priori*

knowledge of detailed enzyme function mechanism and characterization [15,16]. Alternatively, various approximate kinetic forms such as lin-log [17–19] and log-lin [20] kinetics, power law kinetic expressions such as the S-system [21] and Generalized Mass Action [22], and other forms of cooperativity and saturation [23,24] and convenience rate laws [25] have been proposed to reduce the number of kinetic parameters and complexity of the rate expressions. In addition, Varner and Ramkrishna [26–28] pursued the development of kinetic descriptions inspired by cybernetic modeling and optimality concepts. A number of review articles highlight the merits and demerits of various kinetic modeling formalisms [17,29,30]. Uncertainty in the assignment of kinetic parameter values has motivated the development of approaches that do not fix the parameter values but rather sample from a probability distribution [31–34]. Even though the use of kinetic models have led to some successes for strain design [20,35–43] the relative small scope of the employed models, difficulties in obtaining kinetic expressions and questionable portability of kinetic expressions across microbial production platforms have so far limited wide applicability and acceptance.

The introduction of genome-scale models of metabolism [44–46] and the use of Flux Balance Analysis (FBA) to assess their

Author Summary

Computational strain design procedures aim at assisting metabolic engineering efforts by identifying metabolic interventions leading to the targeted overproduction of a desired chemical using network models of cellular metabolism. The effect of metabolite concentrations and substrate-level enzyme regulation cannot be captured with stoichiometry-only metabolic models and analysis methods. Here, we introduce k-OptForce, an optimization-based strain design framework incorporating the mechanistic details afforded by kinetic models, whenever available, into a genome-scale stoichiometric-based modeling formalism. The resulting optimization problems pose significant computational challenges due to the bilevel nature of the formulation and the nonconvex terms in the constraints. A tractable reformulation and solution procedure is introduced for solving the optimization problems. k-OptForce uses kinetic information to (re)apportion reaction fluxes in the network by identifying interventions comprised of both direct enzymatic parameter changes (for reactions with available kinetics) and reaction flux changes (for reactions with only stoichiometric information). Our results show that the introduction of kinetic expressions can significantly alter the identified interventions compared to those identified with stoichiometry-alone analysis. In particular, additional modifications are required in some cases to avoid the violation of metabolite concentration bounds, while in other cases, the kinetic constraints yield metabolic flux distributions that favor the overproduction of the desired product thereby requiring fewer direct interventions.

maximum theoretical yields [47], flux ranges [48] and trade-offs between growth and productivity [49] led to a flurry of computational strain design approaches [50,51] that used a purely stoichiometric description of metabolism. The advantage of using stoichiometry alone supplemented with some regulatory information is that the widest possible range of potentially feasible metabolic phenotypes could be accessed. The linearity of the underlying FBA description also affords significant computational savings and tractability even for genome-scale models. The downside is that identified flux redirection predictions (especially knock up/downs) are sometimes hard to translate into an actionable genetic intervention. For example, it is unclear if a desired metabolic flux up-regulation is achievable or even consistent with enzyme kinetics and/or whether it may lead to physiologically problematic metabolite concentrations.

Stoichiometry-based strain design algorithms are often structured as bilevel mixed integer linear optimization problems (MILP) [4,8,9,11,52]. The outer level optimizes the biotechnological objective (i.e. overproduction of target chemical) through metabolic interventions, while the inner level optimizes the cellular objective that tries to counteract any external imposed genetic or environmental modifications [53,54]. Different fitness functions have been identified to simulate the cellular objective including maximization of biomass yield [4,8], minimization of metabolic adjustment [5,54], regulatory on-off minimization [55,56], worst-case scenario [9,52], or a combination thereof [11,57,58]. Details of these procedures have been reviewed elsewhere [50,51]. Even though they may lack important information on the enzyme kinetics of reactions these procedures have been successfully employed for the strain design of many important chemical products [59–64].

The need to integrate the mechanistic detail (whenever available) of kinetic expressions with the genome-scale scope of

stoichiometric models has been recognized early by the community. Dynamic Flux Balance Analysis (dFBA) [65] integrates uptake kinetic expressions of the carbon substrate while optimizing biomass at every time step to apportion fluxes to the rest of the metabolic network. Several other researchers [66–68] extended this approach to incorporate kinetic expressions of multiple carbon sources and other nutrients into their quasi steady-state formalisms. Zhuang *et al* [69] and Salimi *et al* [70] developed the dynamic multi-species metabolic modeling (DMMM) approach to incorporate uptake kinetics of metabolites in stoichiometric models of a microbial consortium. Alternatively, steady-state flux distributions (from FBA) and stoichiometric information have been used to parameterize genome-scale kinetic models valid for small perturbations [22,71–73]. For example, Fleming *et al* [74] incorporated lin-log kinetic expressions from a small *E. coli* model (comprised of 76 equations) to constrain an FBA simulation. Similarly, Cotton *et al* [75] performed Flux Variability Analysis (FVA) for each flux in a small kinetic model (by allowing the kinetic parameters to vary about their steady-state values) and used the tighter bounds on kinetic parameters to refine flux estimation in genome-scale models. Despite these advancements on the modeling front, the use of hybrid stoichiometric/kinetic models has been left largely unexplored in the context of strain design.

Here, we introduce k-OptForce, which extends the previously developed OptForce procedure [9] by bridging this gap between stoichiometry-only and kinetics-based descriptions of metabolism. This procedure seamlessly integrates the mechanistic detail afforded by kinetic models within a constraint-based optimization framework tractable for genome-scale models. Instead of relying on surrogate fitness functions such as biomass maximization or worst-case simulation for predicting flux re-directions, k-OptForce uses kinetic rate expressions to (re)apportion fluxes in the metabolic network. Using mechanistic models available in literature (for example kinetic models for the central metabolism of *E. coli* [13,76,77] and *S. cerevisiae* [14,78,79]) the allowable phenotype of both the reference and the engineered strain are characterized to be consistent with the allowable kinetic space. Subsequently, alternative genetic intervention strategies consistent with the restrictions imposed by maximum enzyme activity and kinetic regulations, as well as with the worst-case scenario of production of the desired chemical are identified using a bilevel optimization framework. We benchmarked the k-OptForce protocol for the microbial overproduction of L-serine in *E. coli*, and triacetic acid lactone (TAL) in *S. cerevisiae*. For the former, k-OptForce identified key regulatory bottlenecks in upper and lower glycolysis that must be overcome to redirect more flux towards L-serine, which regular OptForce fails to pinpoint. In addition, k-OptForce removed interventions identified by regular OptForce that resulted in kinetically infeasible flux re-distributions. Application of the k-OptForce for the microbial overproduction of TAL in *S. cerevisiae* revealed the impact of additional kinetic constraints in alleviating a severe worst-case simulation of regular OptForce, resulting in a higher prediction of TAL yield from fewer interventions as compared to regular OptForce predictions. The introduction of kinetic expressions in strain design can significantly affect the identified interventions in sometimes non-intuitive ways. In some cases additional modifications are needed to substitute interventions that cause enzyme saturation or concentration bound violations. The mechanism of action of these modifications is often subtle by alleviating substrate inhibition or draining away cofactors from competing pathways. In other cases, kinetic expressions shape flux distributions so as to favor the overproduction of the desired product requiring fewer direct interventions.

Methods

k-OptForce aims at utilizing the available kinetic information within a larger genome-scale stoichiometric model to more accurately characterize all possible metabolic phenotypes of the reference and engineered strains. The procedure builds upon the OptForce procedure [9] by augmenting the metabolic network description with kinetic rate laws whenever available. The reactions in the metabolic network are partitioned into two subsets: reactions with kinetic information $\mathbf{J}^{kin} = \{j | j = 1, 2, \dots, N^{kin}\}$ and reactions with only stoichiometric information $\mathbf{J}^{stoic} = \{j | j = 1, 2, \dots, N^{stoic}\}$ (see Figure 1). Metabolic flux of reactions in \mathbf{J}^{stoic} is constrained only by stoichiometric balances and reaction directionality restrictions whereas flux for reactions in \mathbf{J}^{kin} is fully determined by enzyme activity, metabolite concentrations and kinetic parameter values. The \mathbf{J}^{kin} part of the metabolic network is mathematically described by a system of (usually) nonlinear ordinary differential equations (ODEs) denoting the non-steady-state balance for each metabolite. For a square system of ODEs integration yields steady-state metabolite concentrations and fluxes. The allowable metabolic phenotype (i.e., flux ranges) of the remaining portion of the network \mathbf{J}^{stoic} is inferred so as to be consistent with the predicted fluxes of the reactions in \mathbf{J}^{kin} .

The k-OptForce procedure is composed of the following steps:

Characterization of the reference (e.g., wild-type) strain

The system of ODEs is solved first to obtain steady-state fluxes for reactions in \mathbf{J}^{kin} . The phenotypic space of the reference strain is then identified by iteratively maximizing and minimizing the flux of each reaction in \mathbf{J}^{stoic} while keeping the fluxes of reactions in \mathbf{J}^{kin} fixed at their steady-state values, and restricting the flux of any other reaction for which any experimental data is available at their experimentally determined values or ranges (Figure 2A and Supplementary Text S1).

Characterization of the overproducing strain

The flux ranges of the strain consistent with a desired overproduction target are similarly constructed by successively maximizing and minimizing the flux of network reactions subject to network stoichiometry, overproduction target, while also incorporating the kinetic expressions for reactions in \mathbf{J}^{kin} as additional constraints. The resulting optimization formulation is shown qualitatively in Figure 2B (see Supplementary Text S1 for detailed formulation). Up/down flux regulations within \mathbf{J}^{kin} are modeled by modulating the corresponding maximum enzymatic reaction rates v_j^{max} using a scalar \mathbf{z} to denote the maximum allowable departure from their reference values $v_j^{max,ref}$. A value for \mathbf{z} of zero denotes a knock-out, a value less one implies a down-regulation whereas a

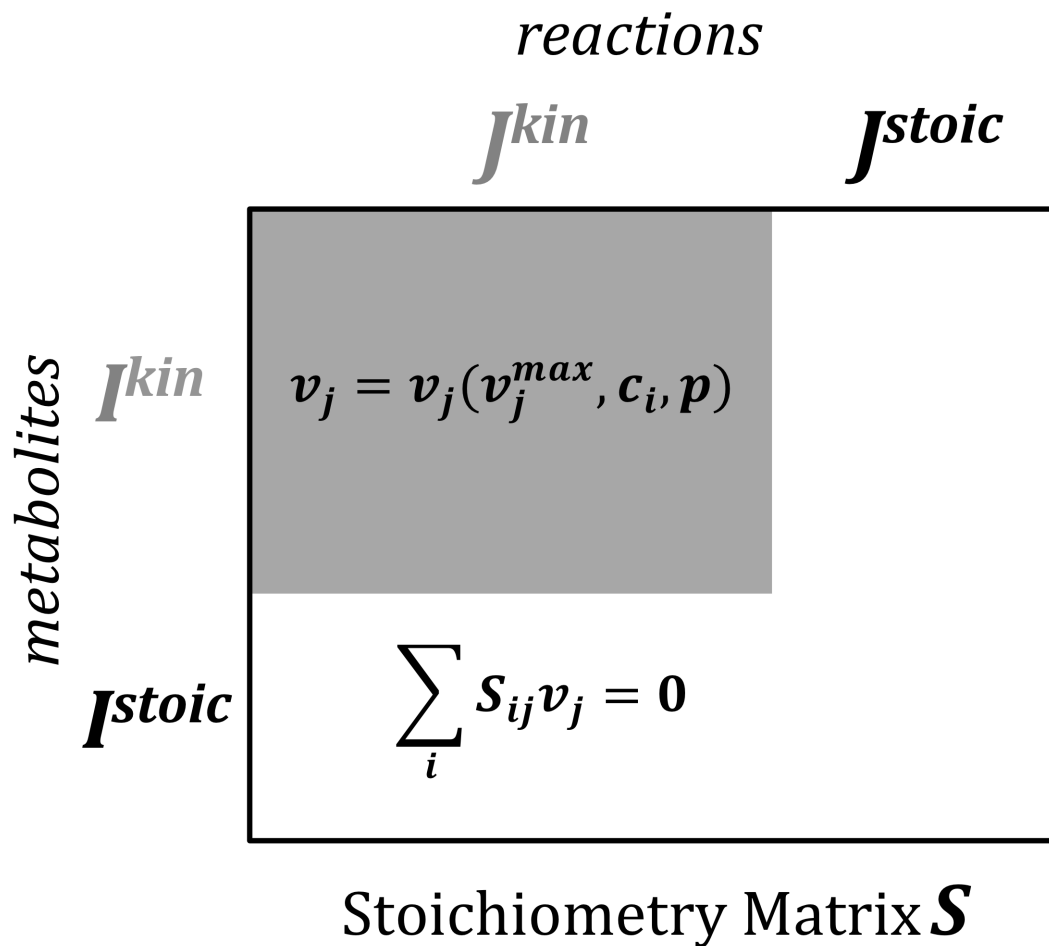


Figure 1. Schematic diagram showing the partition of reactions into the ones with kinetic information \mathbf{J}^{kin} and those linked by only stoichiometry \mathbf{J}^{stoic} . Note that some metabolites participate in only the stoichiometric part of the model (i.e., \mathbf{J}^{stoic}) whereas others participate in both (i.e., \mathbf{J}^{kin}). The flux of reaction j (v_j) in the kinetic part of the model (shaded region) is determined by the kinetic formalism, v_j^{max} , metabolite concentrations c_i and other kinetic parameters p while the ones in \mathbf{J}^{stoic} by the mass conservation laws (non-shaded region). doi:10.1371/journal.pcbi.1003487.g001

A. Characterization of reference phenotype

$$\begin{array}{l}
 \text{Maximize/Minimize} \quad \text{Flux of each reaction in the network } v_j \\
 \text{subject to} \quad \text{Network stoichiometry} \\
 \quad \text{Uptake conditions} \\
 \quad v_j = v_j^{ss} \text{ for all } j \in J^{kin} \\
 \quad v_{biomass} = v_{biomass}^{max,ref}
 \end{array}
 \Rightarrow [v_j^{WT,L}, v_j^{WT,U}]$$

B. Characterization of overproducing phenotype

$$\begin{array}{l}
 \text{Maximize/Minimize} \quad \text{Flux of each reaction in the network } v_j \\
 \text{subject to} \quad \text{Network stoichiometry} \\
 \quad \text{Uptake conditions} \\
 \quad v_j = v_j(v_j^{max}, c_i, p) \text{ for all } j \in J^{kin} \\
 \quad \text{Bounds on metabolite concentrations} \\
 \quad 0 \leq v_j^{max} \leq z \cdot v_j^{max,ref} \text{ for all } j \in J^{kin} \\
 \quad v_{biomass}^{min} \leq v_{biomass}, z \cdot v_{prod}^{target} \leq v_{prod}
 \end{array}
 \Rightarrow [v_j^{OS,L}, v_j^{OS,U}]$$

Figure 2. Optimization structure for the characterization of the reference and the phenotype consistent with overproduction of target chemical. (A) The reference phenotype characterization identifies the minimum $v_j^{WT,L}$ and maximum $v_j^{WT,U}$ flux limits of all reactions consistent with the steady-state reference flux v_j^{ss} of the reactions in J^{kin} and the maximum biomass production $v_{biomass}^{max,ref}$. (B) The overproducing phenotype is identified by calculating the minimum $v_j^{OS,L}$ and maximum $v_j^{OS,U}$ flux limits of all reactions consistent with the kinetic expressions for J^{kin} , the minimum target production of desired chemical v_{prod}^{target} and biomass $v_{biomass}^{min}$. The metabolite concentrations and enzyme activities are allowed to vary within a pre-specified range of their reference values. doi:10.1371/journal.pcbi.1003487.g002

value greater than one an up-regulation. The metabolite concentration ranges involved in the kinetic expressions are selected by selecting a percent allowed departure from the values obtained by solving the system of ODEs (e.g., $\pm 50\%$) and/or experimentally derived measurements. Nonlinear kinetic formalisms give rise to nonconvex nonlinear optimization problems (NLP) for identifying the flux ranges of the overproducing strain. These problems are solved to optimality using the global optimization solver BARON [80] accessed through GAMS. It is important to note that the kinetic expressions significantly restrict the range of allowable flux values consistent with experimental data and/or an overproduction target compared to flux ranges constrained only by stoichiometry. We quantify the average extent of this contraction by defining the Average Bound Contraction (ABC) factor:

$$ABC = \frac{\sum_{j \in J} \left[1 - \frac{(v_j^{U,kin} - v_j^{L,kin})}{(v_j^{U,stoic} - v_j^{L,stoic})} \right]}{N}$$

Here $v_j^{L,stoic}$ and $v_j^{U,stoic}$ denote the respective lower and upper bounds for the fluxes consistent with the overproduction target

when only stoichiometry is used and $v_j^{L,kin}$ and $v_j^{U,kin}$ represent the same bounds upon the incorporation of kinetic information. Interestingly, kinetic information for reactions in J^{kin} propagates into J^{stoic} leading to bound contraction even for reactions with stoichiometry-only description. Tighter description of allowable flux ranges for the reference and overproducing strains allow for sharper elucidation of reactions whose flux must change to meet the imposed target (i.e., MUST sets).

Identification of MUST sets

Similar to the OptForce procedure [9], by contrasting the flux space of the wild-type network with that of the overproducing strain, the sets of reactions that must be up-regulated ($MUST^U$), down-regulated ($MUST^L$), or be knocked out ($MUST^X$) are identified (see supplementary text of Ranganathan *et al* [9] for details). This procedure could be extended to identify higher order MUST sets (e.g., MUST Doubles, Triples etc.) where instead of comparing the flux ranges for individual reactions, we contrast the sum and/or difference of two or more fluxes (depending on the order) between the reference and the desired phenotypes. For example, this procedure can elucidate $MUST^{UU}$, $MUST^{LL}$ and $MUST^{UL}$ sets (see supplementary text of Ranganathan *et al* [9] for details).

Identification of FORCE sets

FORCE set is the minimum set of reactions (and by extension genetic) manipulations selected from within the MUST sets whose direct manipulation (i.e., updating of lower or upper bounds) ensures production of the desired chemical beyond a desired target even under the worst-case scenario where fluxes are re-apportioned to drain flux away from the target product.

The worst-case scenario is mathematically described by extending the bilevel optimization problem used for original OptForce [9], as shown in Figure 3 (also see Supplementary Text S1 for the detailed procedure). The outer problem maximizes the flux towards the desired chemical consistent with reaction kinetics and stoichiometry. Binary variables $y_j^{L,kin}$ and $y_j^{U,kin}$ associated with the $MUST^L$ and $MUST^U$ sets of reactions in J^{kin} respectively, are used to control the effect of engineering modifications to the corresponding v_j^{max} . If $y_j^{L,kin} = 1$ then v_j^{max} for that reaction can be down-regulated to a value between 0 and its wild-type $v_j^{max,ref}$. If $y_j^{U,kin} = 1$ then v_j^{max} for that reaction can be up-regulated to a value between $v_j^{max,ref}$ and $z v_j^{max,ref}$. Otherwise, v_j^{max} is kept unaffected at the reference $v_j^{max,ref}$ value. As in the original OptForce procedure, the inner problem simulates the worst-case scenario by minimizing product formation but only for the reactions in J^{stoic} . A separate set of binary variables $y_j^{L,stoic}$ and $y_j^{U,stoic}$ identify interventions in J^{stoic} required to guarantee a non-zero yield of the target chemical consistent with the flux distribution in J^{kin} . It is important to note that the metabolic fluxes in J^{kin} remain unaffected by the worst-case simulation of the reactions in J^{stoic} in the inner problem.

This bilevel formulation is converted into a single level mixed integer nonlinear problem (MINLP) using the conditions of strong duality. We construct the dual of the inner problem (called the primal), and add all the dual constraints, along with those of the primal, to the outer optimization problem. Since all the nonlinear kinetic expressions are present in the outer problem, the inner (primal) problem is linear in the reaction fluxes v_j . The fitness function of the inner problem is imposed setting the objectives of the primal and the dual equal to each other. It is to be noted here that the dualization of the flux variable for each reaction in J^{kin} introduces a bilinear term to the single-level formulation. This can be avoided by imposing the Karush-Kuhn-Tucker (KKT) conditions of complementary slackness between the primal constraints and their dual variables. This leads to the introduction of a binary variable for each constraint in the primal problem which is generally more tractable than the original bilinear constraints (see Supplementary Text S1 for detailed formulations). The above described sequence of equivalent problem re-formulations yields a single-level nonlinear MINLP (k-OptForce). The single-level k-OptForce optimization is successively solved using the global solver BARON [80] for an increasing number of interventions (by increasing κ) until the target yield is met.

Due to the nonconvex nature of the kinetic expressions and the large number of binary variables the resulting MINLP equivalent representation of the bilevel optimization problem may become computationally intractable. For these cases, we exploit the natural hierarchy of the model by first selecting interventions within J^{kin} and subsequently within J^{stoic} (see Figure 4 and Supplementary Text S1 for the algorithmic details). In the first step, a nonconvex optimization problem is solved to identify the minimum number of

$$\begin{array}{ll}
 \text{Maximize} & v_{prod} \\
 \text{(over } y_j^L, y_j^U, c_i) & \\
 \text{subject to} & r_j = v_j(v_j^{max}, c_i, p) \text{ for all } j \in J^{kin} \\
 & \text{Bounds on metabolite concentrations} \\
 & \text{If } y_j^{L,kin} = 1 \rightarrow v_j^{max} \in [0, v_j^{max,ref}] \text{ for all } j \in MUST^L \\
 & \text{If } y_j^{U,kin} = 1 \rightarrow v_j^{max} \in [v_j^{max,ref}, z \cdot v_j^{max,ref}] \text{ for all } j \in MUST^U \\
 \\
 \text{Minimize} & v_{prod} \\
 \text{(over } v_j \text{ for all } j \in J^{stoic}) & \\
 \text{subject to} & \text{Network stoichiometry} \\
 & \text{Uptake conditions} \\
 & \text{If } y_j^{L,stoic} = 1 \rightarrow v_j \in [LB_j, v_j^{OS,U}] \text{ for all } j \in MUST^L \\
 & \text{If } y_j^{U,stoic} = 1 \rightarrow v_j \in [v_j^{OS,L}, UB_j] \text{ for all } j \in MUST^U \\
 & v_j = r_j \text{ for all } j \in J^{kin} \\
 \\
 \sum_{j \in \{J^{kin} \cup J^{stoic}\}} y_j^U + y_j^L = \kappa
 \end{array}$$

Figure 3. Single-step bilevel formulation for the identification of FORCE sets. The outer problem maximizes the flux towards the desired chemical while the inner problem simulates the worst-case scenario by minimizing the product flux. Binary variables $y_j^{U,kin}$ and $y_j^{L,kin}$ propagate the effect of engineering modifications in J^{kin} while $y_j^{U,stoic}$ and $y_j^{L,stoic}$ do the same in J^{stoic} . The fluxes in J^{kin} remain unchanged in the inner problem. doi:10.1371/journal.pcbi.1003487.g003

manipulations in J^{kin} that are consistent with the overproduction target (Figure 4A). Keeping the fluxes in J^{kin} fixed at their optimized values the flux ranges of the overproducing strain are then re-calculated and the FORCE set for reactions in J^{stoic} are then identified (see Step 2) as in regular OptForce [9] (Figure 4B). By solving two separate problems the computational burden is significantly reduced at the expense of potentially missing synergistic interventions that share reactions between J^{kin} and J^{stoic} . It is to be noted that while the illustrated formulation only targets changes in v_j^{max} for reactions in J^{kin} the same analysis could be applied for the modulation of other kinetic parameters (e.g., K_m , K_i etc.) in the model.

Once the FORCE set of interventions are identified (along with alternative manipulation strategies), it is important to manually curate the results to understand the underlying metabolic reason behind each intervention. This is necessary since k-OptForce makes use of not just stoichiometry which imposes straightforward connections between reactants and products but also kinetics that introduce complex nonlinear couplings often between distal reactions through metabolite pools. While it is not possible to put forth an invariant set of rules for all applications, the following checks can be useful in general: (1) check whether any metabolite participating in affected reactions is hitting lower or upper concentration bounds, (2) identify if a common metabolite is

A. Identification of interventions in J^{kin}

$$\begin{array}{ll}
 \text{Minimize} & \text{Number of interventions within } J^{kin} \\
 \text{(over } y_j^{L,kin}, y_j^{U,kin}, c_i) & \\
 \text{subject to} & \text{Network stoichiometry} \\
 & \text{Uptake conditions} \\
 & v_j = v_j(v_j^{max}, c_i, p) \text{ for all } j \in J^{kin} \\
 & \text{Bounds on metabolite concentrations} \\
 & \text{If } y_j^{L,kin} = 1 \rightarrow v_j^{max} \in [0, v_j^{max,ref}] \text{ for all } j \in MUST^L \\
 & \text{If } y_j^{U,kin} = 1 \rightarrow v_j^{max} \in [v_j^{max,ref}, z \cdot v_j^{max,ref}] \text{ for all } j \in MUST^U \\
 & v_{prod} \geq v_{prod}^{target}
 \end{array}$$

B. Identification of interventions in J^{stoic}

$$\begin{array}{ll}
 \text{Maximize} & v_{prod} \\
 \text{(over } y_j^{L,stoic}, y_j^{U,stoic}) & \\
 \text{subject to} & \begin{array}{ll}
 \text{Minimize} & v_{prod} \\
 \text{(over } v_j \text{ for all } j \in J^{stoic}) & \\
 \text{subject to} & \text{Network stoichiometry} \\
 & \text{Uptake conditions} \\
 & \text{If } y_j^{L,stoic} = 1 \rightarrow v_j \in [LB_j, v_j^{OS,U}] \text{ for all } j \in MUST^L \\
 & \text{If } y_j^{U,stoic} = 1 \rightarrow v_j \in [v_j^{OS,L}, UB_j] \text{ for all } j \in MUST^U \\
 & v_j = v_j^{step1} \text{ for all } j \in J^{kin}
 \end{array} \\
 & \sum_{j \in J^{stoic}} y_j^{U,stoic} + y_j^{L,stoic} = \kappa
 \end{array}$$

Figure 4. Two-step formulation for the identification of the FORCE sets. (A) The first step of the formulation identifies the minimum number of interventions ($y_j^{U,kin}$ and $y_j^{L,kin}$) in J^{kin} required to meet the desired levels of overproduction of target chemical. (B) The second step identifies the additional interventions ($y_j^{U,stoic}$ and $y_j^{L,stoic}$) in J^{stoic} that would guarantee the desired yield in the worst case scenario. doi:10.1371/journal.pcbi.1003487.g004

fixing the branching ratio of fluxes in two pathways, (3) resolve whether a metabolite is limiting the flux of a reaction through substrate-level inhibition, (4) confirm if the flux in a pathway has been restricted because the v^{max} of one of the reactions has hit its upper bound, and, (5) analyze all alternate intervention strategies to identify common metabolites and/or enzymes that are being targeted.

Results

We contrasted k-OptForce with the original OptForce [9] predictions for the overproduction of L-serine and TAL in *E. coli* and *S. cerevisiae*, respectively. The goal was to assess how the addition of kinetic information into stoichiometric models changes the list of identified interventions and more importantly what are the underlying reasons for the new interventions.

Overproduction of L-serine in *E. coli*

L-Serine is a non-essential amino acid and a precursor for other amino acids such as cysteine, tryptophan and glycine. It also participates in the biosynthesis of purines and pyrimidines, and serves as an intermediate for phospholipids, sphingolipids and folate synthesis for several microorganisms [81,82]. The synthesis of L-serine in microorganisms such as *Escherichia coli* and *Corynebacterium* consists of a three-step pathway branching out of the glycolytic intermediate 3-phosphoglycerate (3pg). 3pg is converted to 3-phosphohydroxypyruvate (3php) by phosphoglycerate dehydrogenase (PGCD, EC 1.1.1.95), and phosphoserine transaminase (PSERT, EC 2.6.1.52) catalyzes the conversion of 3-phosphohydroxypyruvate (3php) to L-phosphoserine (pser-L) using L-glutamate as the amino acid donor. In the last step, phosphoserine phosphatase (PSP, EC# 3.1.3.3) catalyzes the final conversion to L-serine (see Figure 5). We used the genome-scale *iAF1260* model of *E. coli* [83] as the stoichiometric model for our simulations. The kinetic rate expressions for reactions of central metabolism were extracted from Chassagnole *et al* [76]. This kinetic model, which has been used before in variety of studies [5,36,39,84], consists of 25 metabolites and 25 reactions from glycolysis and pentose phosphate pathway (see Supplementary Material S1). All simulations were carried out in aerobic minimal medium with glucose as the sole carbon source.

The reference strain (i.e., wild-type *E. coli*) flux ranges were identified by finding the maximum flux variability in the entire network while keeping the fluxes in \mathbf{J}^{kin} fixed at the steady-state values obtained by solving the system of ODEs for the kinetic model (see Supplementary Figure S1). The L-serine overproducing network flux ranges were calculated for a target of 90% maximum theoretical yield (i.e., 180 mol/100 mol glucose uptake). The minimum biomass production was kept at 10% of its maximum achievable. The maximum enzyme activity v^{max} of reactions in \mathbf{J}^{kin} was allowed to vary from zero to two-fold up-regulation of its reference activity (i.e., $z = 2$). Also, the concentration of metabolites in \mathbf{I}^{kin} was allowed to vary within a two-fold range from their steady-state concentrations in the reference strain.

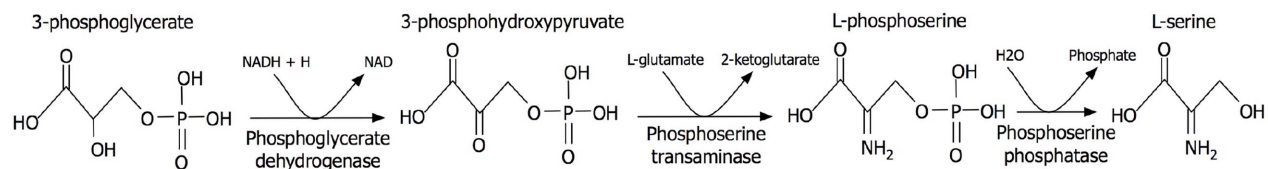


Figure 5. Pathway for L-serine production in *E. coli*.
doi:10.1371/journal.pcbi.1003487.g005

Figure 6 illustrates the reduction in flux ranges in the overproducing phenotype after the introduction of the kinetic constraints when compared with a stoichiometry-only description. The average bound contraction (ABC) was 52% for the fluxes in \mathbf{J}^{kin} . For example, the flux of glucose 6-phosphate dehydrogenase (G6PDH) in oxidative pentose phosphate (PP) pathway consistent with the imposed L-serine overproduction ranged from 0 to 136 mmol gDW⁻¹hr⁻¹ when constrained by only stoichiometry and from 0 to 62 mmol gDW⁻¹hr⁻¹ when imposing also kinetic information. This range reduction is due to the limitations of the maximum enzyme activity as well as metabolic concentrations of glucose-6-phosphate (g6p) regulating G6PDH flux. The restrictions implied by kinetics also propagate throughout the stoichiometric part of the network leading to an average ten percent range contraction for reactions in \mathbf{J}^{stoic} . For example, the fluxes ranges in 2-oxoglutarate dehydrogenase (AKGDH) and succinate dehydrogenase (SUCD) in the TCA cycle decreased 7.4% and 6%, respectively as a direct consequence of the flux range reduction for pyruvate dehydrogenase (PDH) in \mathbf{J}^{kin} which supplies acetyl-CoA (accoa) to TCA cycle. As a result of the tighter flux ranges in the overproducing network (and characterization of base strain) many more reactions are identified that must depart from their original ranges (i.e., MUST sets) compared to regular OptForce both in \mathbf{MUST}^U (38 vs. 3) and \mathbf{MUST}^L (293 vs. 108) sets. For example, up-regulation of glucose-6-phosphate isomerase (PGI) in upper glycolysis supplies more flux towards 3pg and L-serine production. The flux range for PGI in the overproducing phenotype (−36–100 mmol gDW⁻¹hr⁻¹) was wide enough to overlap with its reference flux value (35 mmol gDW⁻¹hr⁻¹) suggesting PGI up-regulation is not necessary for L-serine overproduction. In contrast, using k-OptForce we find that the flux range of PGI in the overproducing phenotype is restricted to 38–98 mmol gDW⁻¹hr⁻¹ which does not include the reference value of 35 mmol gDW⁻¹hr⁻¹. This implies that it is impossible to produce L-serine at 90% theoretical yield without directly (or indirectly) increasing the flux through PGI which becomes a member of \mathbf{MUST}^U . We also observe a significant increase in the number of reactions in \mathbf{MUST}^L . This is because the kinetic expressions in \mathbf{J}^{kin} fix the branching ratios for fluxes emanating from metabolites in \mathbf{I}^{kin} . As a result, many reactions in \mathbf{J}^{stoic} involving metabolites participating in reactions from \mathbf{J}^{kin} appear as down-regulations. For example, k-OptForce identifies that the acetyl-CoA carboxylase (ACCOAC) flux, which branches away from pyr and accoa towards membrane lipid metabolism must decrease (i.e., \mathbf{MUST}^L) as it goes from (9–527) mmol gDW⁻¹hr⁻¹ in the reference strain to (2.1–2.3) mmol gDW⁻¹hr⁻¹ in the overproducing strain.

Figure 7 summarizes the FORCE set of reactions as predicted by the original and k-OptForce. As expected, the first intervention suggested by both procedures is an at least 20-fold up-regulation in the activity of one of the three fluxes that directly lead to the synthesis of L-serine (i.e., PGCD, PSERT and PSP). However, the remaining interventions follow completely different strategies. k-OptForce emphasizes the need to remove substrate-level inhibition

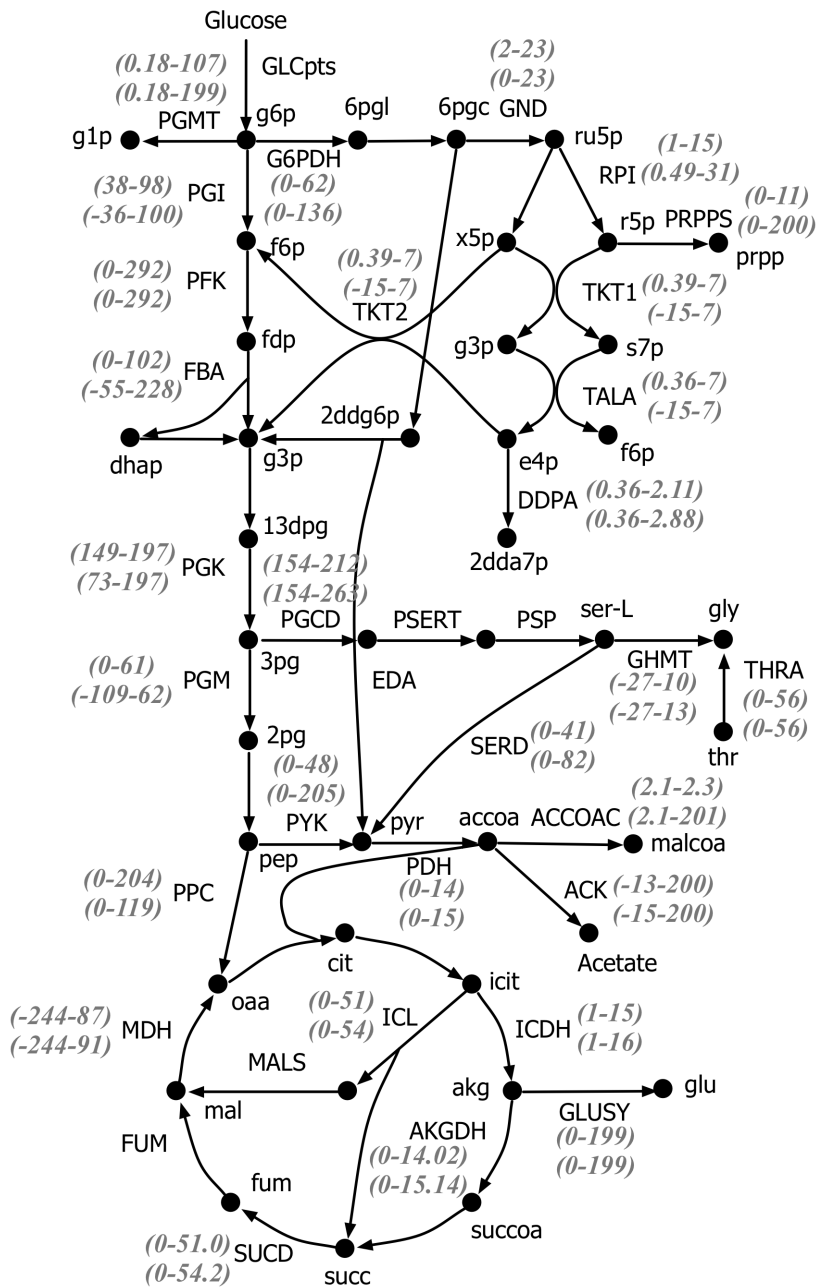


Figure 6. Reduction in the flux ranges of the desired phenotype for L-serine overproduction in *E. coli* after the incorporation of additional kinetic constraints. Values on top indicate reduced flux ranges (in mmol gDW⁻¹ hr⁻¹), while values in the bottom indicate the flux ranges (in mmol gDW⁻¹ hr⁻¹) without any kinetic information, for 100 mmol gDW⁻¹ hr⁻¹ of glucose uptake. doi:10.1371/journal.pcbi.1003487.g006

by making relatively small flux changes on a number of reactions to maintain concentrations within the imposed bounds (i.e., two-fold changes from wild-type measurements). Figure 7b illustrates that it is necessary to up-regulate upper glycolysis and down-regulate lower glycolysis to divert flux towards L-serine. The upper glycolytic pathway is tightly regulated by both product metabolites and nadh [85]. The kinetic expressions in Chassagnole *et al* [76] encode inhibition of PGI and phosphofruktokinase (PFK) by 6-Phospho-D-gluconate (6pgc) and phosphoenolpyruvate (pep). The high activity of the PP pathway and lower glycolysis in the wild-type requires elevated intracellular levels of 6pgc (0.8 mM) and pep (2.86 mM) to supply the fluxes through phosphogluconate

dehydrogenase (GND) (63 mmol gDW⁻¹ hr⁻¹) and pyruvate dehydrogenase (PDH) (93 mmol gDW⁻¹ hr⁻¹) reactions. The high concentrations of 6pgc and pep both prevent the up-regulation of upper glycolysis and the down-regulation of lower glycolysis. To alleviate this, k-OptForce suggests removal of PDH coupled with an at least three-fold down-regulation in the activity of either transaldolase (TALA) or transketolase (TKT1) reactions in PP pathway. Removal of PDH also allows the concentration of pep to be reduced from 2.8 to 1.43 mM thus alleviating its inhibitory effect on PFK. Likewise, down-regulation of the PP pathway activity reduces the flux in GND lowering the concentration of 6pgc (from 0.8 to 0.4 mM). Removal of substrate

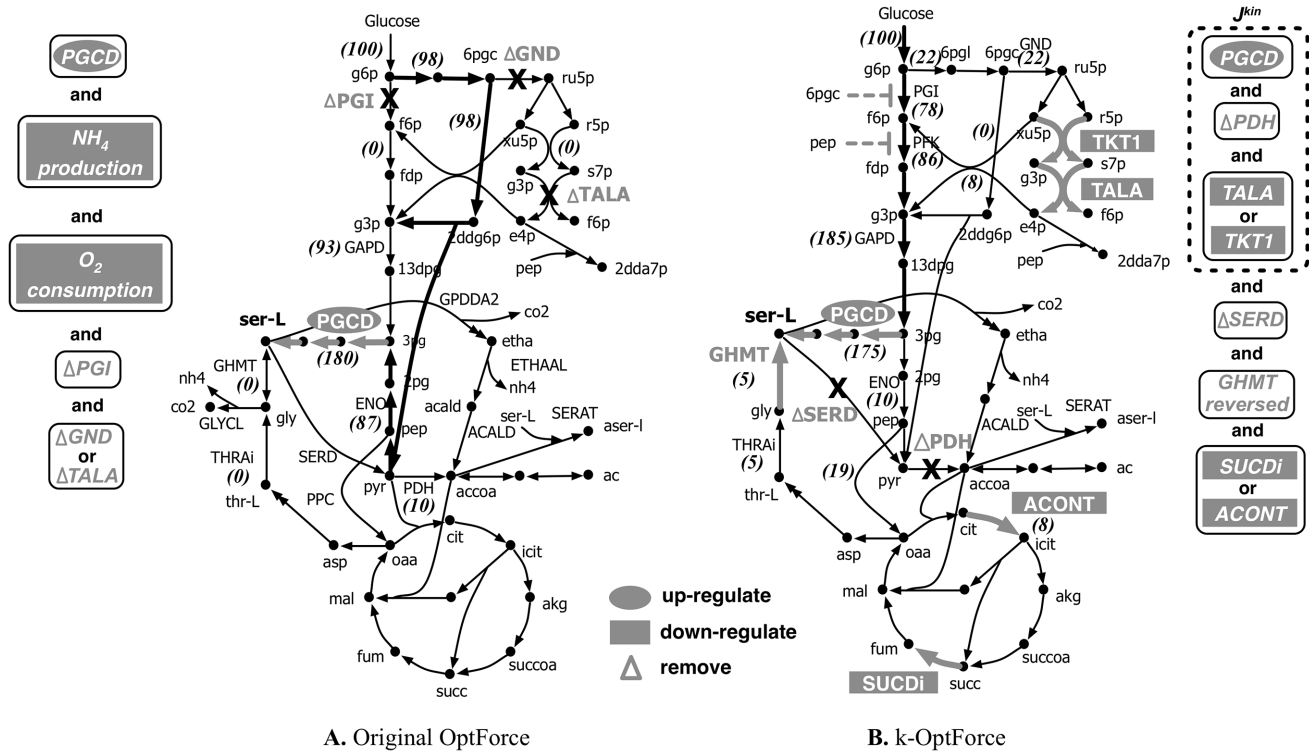


Figure 7. Comparison of intervention strategies predicted by A. regular OptForce and B. k-OptForce for overproduction of L-serine in *E. coli*. The values in brackets indicate the metabolic flux in $\text{mmol GDW}^{-1}\text{hr}^{-1}$ per $100 \text{ mmol GDW}^{-1}\text{hr}^{-1}$ glucose uptake. doi:10.1371/journal.pcbi.1003487.g007

inhibition on PFK and PGI by pep and 6pgc leads to increased flux towards L-serine. The original OptForce procedure cannot identify such interventions, as substrate inhibition is not captured through stoichiometric modeling.

The inhibitory effect of 6pgc and pep cannot be completely removed due to their prescribed lower limits in concentration (0.4 mM and 1.43 mM respectively). Moreover, the upper limits on concentration of metabolites involved in upper glycolysis put an upper bound on the amount of flux that can be carried by upper glycolysis. Therefore, additional interventions are needed to meet the L-serine target yield by modulating pathways away from glycolysis. k-OptForce suggests the reversal of glycine hydroxymethyltransferase (GHMT) thus converting glycine to L-serine (see Figure 7b). In contrast, the original OptForce predicts that the entire amount of flux required for L-serine can be supplied through the up-regulation of the serine synthase pathway as no inhibitory effect or concentration bound is considered. It is to be noted here that the forward activity of GHMT is essential *in vivo* [86,87]. If, however, the lower limits on the concentration of 6pgc and pep are reduced to 0.35 mM and 1.3 mM respectively, their inhibitory effect on upper glycolysis is alleviated sufficiently to route all the flux required for L-serine production through the serine synthase pathway. The upper glycolytic flux of PGI increases from 78 to 80 $\text{mmol gDW}^{-1}\text{hr}^{-1}$ and the PP flux is down-regulated further (from 22 to 20 $\text{mmol gDW}^{-1}\text{hr}^{-1}$) to provide the extra flux for L-serine (results not shown here). As a result, k-OptForce suggests down-regulation of GHMT by at least 3 folds from its reference flux instead of its reversal. All other interventions remain un-altered.

The remaining interventions suggested by k-OptForce aim at preventing the drain of metabolic flux from L-serine. Removal of L-serine deaminase (SERD_L) prevents the conversion of L-serine

to pyruvate. This is followed by an at least six-fold down-regulation (from 60 to 9 $\text{mmol gDW}^{-1}\text{hr}^{-1}$) of either citrate synthase (CS) or succinate dehydrogenase (SUCD) reactions to reduce the TCA cycle activity which arrests ATP production in the network. This prevents the conversion of L-serine to acetaldehyde whose activity requires five units of ATP per unit of flux. The original OptForce achieves the same goal by simply down-regulating the transport of oxygen and up-regulating the transport of ammonium into the cell. These interventions were not chosen by k-OptForce as they lead to flux values that are inconsistent with the kinetic expressions in J^{kin} .

Consistent with k-OptForce predictions, metabolic engineering studies on *C. glutamicum* have revealed that overexpression of the *serA/B* and *c* encoding for the three enzymes in the L-serine production pathway have a positive, though small, effect on L-serine production [88,89]. Removal of *sdaA* encoding for the SERD_L reaction, coupled with up-regulation of the L-serine pathway have been reported to lead to higher L-serine yields [89] consistent with k-OptForce predictions. Other studies have shown that down-regulation of GHMT reaction through the removal of *glyR* regulator further improves L-serine production [90]. In a recent study, overexpression of *pgk* was shown to divert more flux towards L-serine in *C. glutamicum* [91]. This could be viewed as an alternative strategy to the one suggested by k-OptForce involving alleviation of the substrate level inhibition of upper glycolysis through down-regulation of PP and lower glycolytic flux. It must be emphasized that the k-OptForce results depend heavily on the accuracy of the rate expressions of the kinetic model. For example, it has been found in both *E. coli* and *C. glutamicum*, that the activity of PGCD is feedback inhibited by L-serine [81,92]. Alleviating this feedback regulation significantly improves production of L-serine [91]. However, k-OptForce cannot capture this regulation since

the adopted kinetic model does not include this inhibitory effect. Accordingly, k-OptForce predictions must be carefully scrutinized to identify the driving forces for the identified interventions (e.g., substrate inhibition removal, ATP drain, cofactor sequestering, concentration increase, etc.) and the reason for the omission of seemingly straightforward interventions (e.g., concentration bound violation, inadequate v^{max} , lethal deletion, cofactor imbalance, etc.).

In addition to suggesting intervention strategies consistent with the kinetic constraints in the network, k-OptForce also pinpoints which ones and how original OptForce interventions violate network kinetics. For example, the original OptForce framework suggested the reversal of lower glycolytic reactions that converge to 3pg. This is accomplished by removing PGI and either GND, TKT1 or TALA in PP pathway to reroute the metabolic flux toward 3pg and pyruvate through E-D pathway by using 2-dehydro-3-deoxy-phosphogluconate aldolase (EDA). Reversal of enolase (ENO) and phosphoglycerate mutase (PGM) in lower glycolysis converts pyruvate to 3pg. k-OptForce finds that this redirection is not feasible. Reversible reactions PGM and ENO rely on the relative concentrations of its reactants and products to inform their directionality. Their reversal, to the extent suggested by the original OptForce procedure requires the respective metabolic levels of 2-phosphoglycerate (2pg) and pep to increase to 1.162 and 6.05 mM respectively, beyond the imposed upper limits of (0.856 and 4.726 mM respectively). Therefore, k-OptForce provides both a check on stoichiometry-only derived interventions and more importantly quantifies the impact of flux redirections on metabolite concentrations and required enzyme levels.

Sensitivity of the k-OptForce results to perturbations in kinetic parameters

Previous reports [93] on the sensitivity analysis of the *E. coli* kinetic model by Chassagnole *et al* [76] showed that simulation results are only sensitive to the values of nine (out of 25) enzyme activities in the model. In light of this analysis, we perturbed the enzyme activities of two sensitive ones (v_{PDH}^{max} and v_{G6PDH}^{max}) and two rather insensitive (v_{PGCD}^{max} and v_{TKT1}^{max}) by $\pm 20\%$ from their reference levels and repeated the k-OptForce calculations. Results showed that apart from v_{G6PDH}^{max} up-perturbation, the remaining interventions (both up-perturbation and down-perturbation) were identical to the original results. Increasing the value of v_{G6PDH}^{max} , which is one of the most highly sensitive parameters in the model [93], increased the PP flux for the reference phenotype by 11% (from 63 mmol gDW⁻¹hr⁻¹ to 70 mmol gDW⁻¹hr⁻¹), as the glycolytic pathway was inhibited by increased 6pgc concentration. As a result, down-regulation of GND was suggested as an additional intervention to reduce the increased PP activity and route more flux from glycolysis towards L-serine. In all other cases metabolite concentrations and fluxes in the kinetic model were minimally affected by perturbations in enzyme activity.

Production of triacetic acid lactone (TAL) in *S. cerevisiae*

4-Hydroxy-6-methyl-2-pyrone, commonly known as TAL, is a precursor for the production of phloroglucinol [94], which is an intermediate for many products such as 1,3,5-triamino-2,4,6-trinitrobenzene (TATB) and resorcinol [95]. Synthesis of TAL [96] in both *E. coli* and *S. cerevisiae* has been explored [97] [98]. Because neither *E. coli* nor *S. cerevisiae* can natively synthesize TAL, production routes for TAL rely on the heterologous expression of non-native enzymes such as 2-pyrone synthase (2-PS) (found in *Gerbera hybrida*) [96] or a genetically modified 6-methylsalicylic acid synthase (6-MSAS) [96,97] and *fasB* [96,98] with their ketoreductase domains deactivated. These efforts have led to TAL

yields in *S. cerevisiae* of only up to 6% of the theoretical maximum (with a titer of 1.8 g/l) in glucose medium. Figure 8 shows the targeted pathway for TAL synthesis in *S. cerevisiae*. We used the iAZ900 model [99] of *S. cerevisiae* as the stoichiometric network of metabolism. The kinetic expressions for reactions in central metabolism were imported from the kinetic model of central metabolism of *S. cerevisiae* described by van Eunen *et al* [79]. The model consists of twelve metabolites and twelve reactions, for the glycolytic pathway and the conversion of pyruvate to ethanol. Since the kinetic model did not include drains for amino acids from the central metabolic pathway metabolites (g6p, f6p, g3p, 3pg, pep), we added drains (similar to the method used in Chassagnole *et al* [76]) using MFA information of *S. cerevisiae* central metabolism from Gombert *et al* [100] to ensure biomass production (see Supplementary Material S1).

As in the first example, we allowed for up to two-fold changes in the metabolite concentrations and v^{max} from their wild-type values. Contrary to the previous example, here the kinetic expressions do not restrict further the flux ranges as the ABC metric (see Methods) for all J^{kin} and J^{stoic} fluxes is zero. As a result, we find no difference in the *MUST^U* (19 reactions) and *MUST^L* (61 reactions) sets for kinetic and original OptForce. This is due to the relatively few fluxes with kinetic expressions and the already fairly tight flux ranges achieved by stoichiometry-alone. For example, the PGI flux varies within the narrow range between 91 and 97 mmol gDW⁻¹hr⁻¹ for a 90% yield for TAL even when no kinetic expressions are used. This is because the imposed high production target for TAL largely fixes the flow in glycolysis. As a consequence of negligible contraction in flux ranges due to the kinetic constraints, no difference in predicted *MUST* sets by regular and k-OptForce is observed.

Figure 9 compares the FORCE sets and the respective guaranteed yield of TAL as suggested by original and k-OptForce. In general, both procedures suggest strategies that increase the availability of precursors accoa and malonyl-CoA (malcoa), up-regulating glycolysis, down-regulating PPP, and reducing nadph. However, while the original OptForce suggests that at least four interventions are required to achieve a 35% yield for TAL, k-OptForce suggests that a yield of 90% is achievable by only two interventions. Not surprisingly, both procedures suggest the up-regulation of the ACCOAC (by at least nine-fold of its reference flux) to increase the availability of the direct TAL precursor malonyl-CoA. The glycolytic flux is also up-regulated to divert flux from the PP pathway towards TAL. k-OptForce identifies that the kinetic expressions work in concert with the overproduction goal (given the imposed concentration ranges) without the need for any direct enzymatic interventions for reactions in J^{kin} . Figure 10 illustrates the required changes in metabolite concentrations in the overproducing network as predicted by k-OptForce. Elevated concentrations of metabolites in glycolysis lead to an increase the flux towards TAL. For example, the concentration of g6p in the upper glycolysis is increased by 8% (from 1.41 to 1.55 mM) leading to more flux through PGI (from 78 to 96 mmol gDW⁻¹hr⁻¹). This eliminates all flux through G6PDH and PP (from 19 to 0 mmol gDW⁻¹hr⁻¹) to maintain the steady-state metabolite balance of g6p. Without the benefit of any kinetics, the original OptForce suggests the removal of G6PDH reaction as a requirement for down-regulating the PP pathway. k-OptForce also requires that the concentrations of glyceraldehyde 3-phosphate (g3p) and 3pg be elevated by 6% and 31%, respectively from their reference states to up-regulate the lower glycolysis flux from 168 to 192 mmol gDW⁻¹hr⁻¹. This re-direction in glycolysis prevents the loss of metabolic flux towards glycerol synthesis. Instead, the original OptForce procedure suggests the

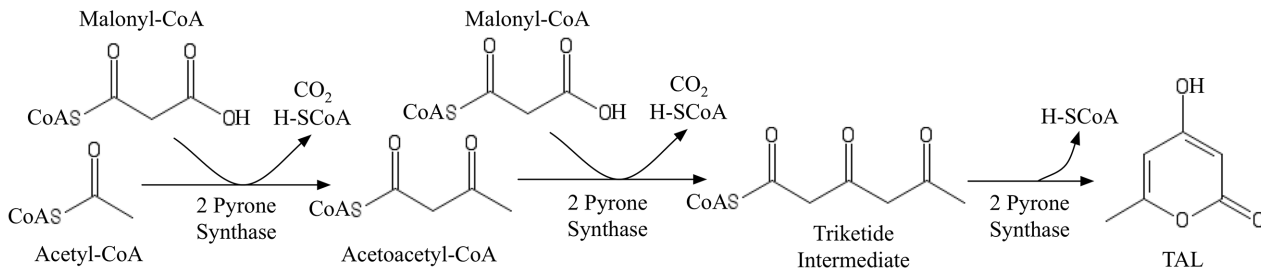


Figure 8. Pathway for TAL production in *S. cerevisiae*.
doi:10.1371/journal.pcbi.1003487.g008

removal of glycerol-3-phosphatase (G3PT) in the glycerol synthesis pathway to serve the same purpose and channel the flux towards lower glycolysis and TAL.

The fatty acid synthase is in direct competition to TAL production. It uses the same precursors as TAL (i.e., accoa and malcoa) to form medium chain fatty acids. Activity of fatty acid synthase requires the cofactor nadph for the reductive steps in the pathway. Not surprisingly, both kinetic and original OptForce identify strategies to lower the availability of nadph. k-OptForce achieves this by suggesting a 20 fold up-regulation (from 3.5 to 69 mmol gDW⁻¹ hr⁻¹) in the flux of aldehyde dehydrogenase (ALDD) which converts acetaldehyde towards acetate. The major routes of acetaldehyde production in *S. cerevisiae* are either from direct decarboxylation of pyruvate through pyruvate decarboxylase (PYRDC) or through the alternate threonine synthesis pathway followed by the cleavage of threonine by threonine aldolase (THRD) to acetaldehyde and glycine (see Figure 9B). The threonine synthesis pathway is favored in TAL overproduction as it consumes one unit of nadph for every unit of flux. We note that the fluxes in PYRDC and alcohol dehydrogenase (ALCD) are fixed by the kinetic constraints in J^{kin} . Therefore, up-regulation of ALDD causes most

of pyruvate to be routed through the threonine production pathway (to maintain the steady-state conservation of acetaldehyde) resulting in a decrease in nadph levels. The original OptForce does not arrive at this intervention as the kinetic control on the fluxes of PYRDC and ALCD is not captured. It instead suggests the removal of cytosolic isocitrate dehydrogenase (ICDH) to reduce the nadph production, and thus arrests fatty acid synthesis.

Unlike the previous example where k-OptForce required more interventions for the same overproduction target than the original OptForce, here the reverse trend is observed. The predicted yield for TAL by the original OptForce is only 35% of its theoretical maximum after four interventions whereas k-OptForce reaches 90% of theoretical maximum with only two manipulations. This is because the incorporation of kinetic information pushes metabolic flux in the direction that is needed for overproduction and away from the “worst-case” behavior.

Effect of metabolite concentration ranges on identified interventions

The steady-state balances of metabolites in a kinetic model [76,79] (with as many metabolites as rate equations) form a

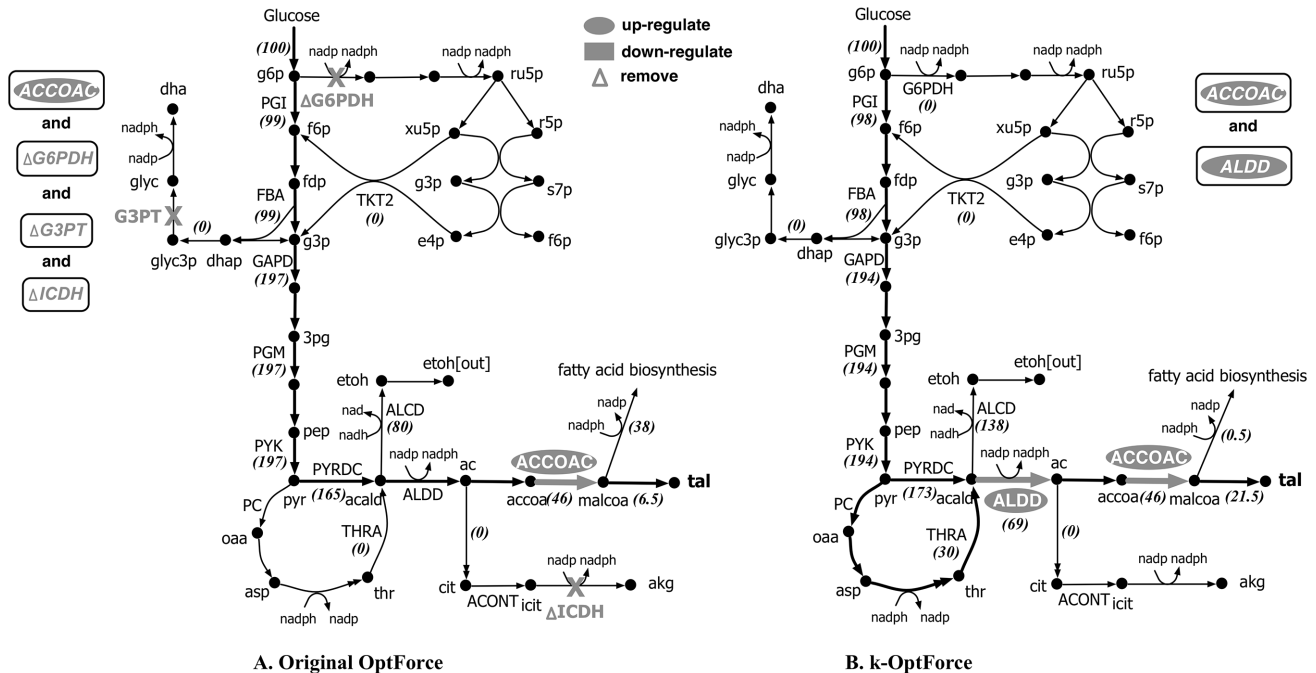


Figure 9. Comparison of intervention strategies predicted by A. regular OptForce and B. k-OptForce for overproduction of TAL in *S. cerevisiae*. The values in brackets indicate the metabolic flux in mmol gDW⁻¹ hr⁻¹ per 100 mmol gDW⁻¹ hr⁻¹ glucose uptake.
doi:10.1371/journal.pcbi.1003487.g009

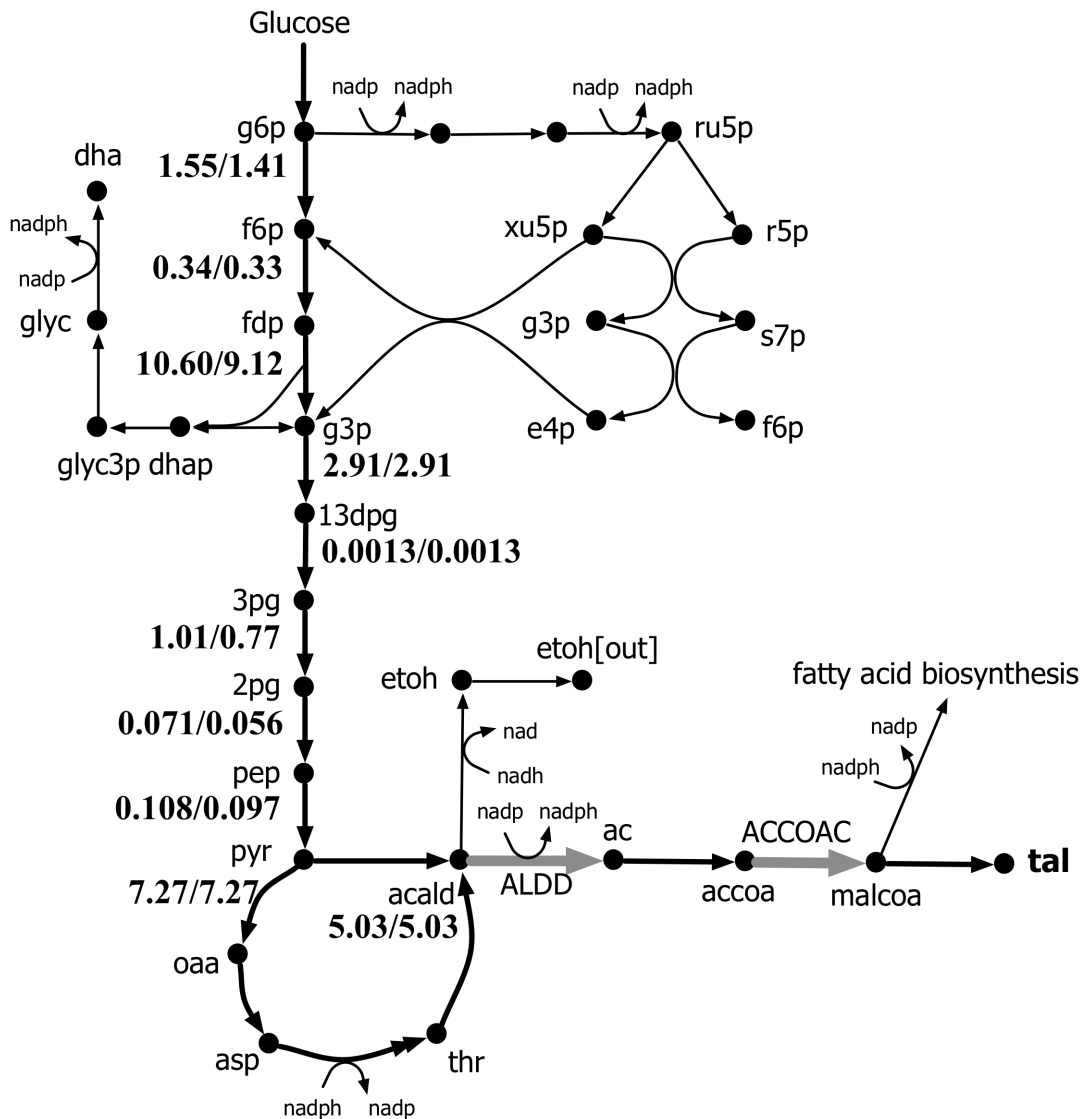


Figure 10. Comparison of the concentrations (in mM) of metabolites in I^{kin} for the engineered and reference strain of *S. cerevisiae* for overproduction of TAL. The first value indicates the overproducing strain concentration, while the second value refers to the reference strain concentration.

doi:10.1371/journal.pcbi.1003487.g010

square-system of equations with zero degrees of freedom. Assuming that there are no multiple steady-states due to the nonlinearity of the kinetic expressions, steady-state fluxes or concentrations cannot change unless accompanied by alterations in kinetic parameters (e.g. enzyme activities v^{max}). However, when a kinetic model is integrated with a stoichiometric genome-scale model, reactions in J^{stoic} that involve metabolites present in I^{kin} add in effect additional degrees of freedom to the square-system of equations thus decoupling metabolite concentrations from enzyme activities. As a result metabolite concentrations can change in such a way that fluxes are altered without requiring any enzymatic interventions as observed in the TAL overproduction example. In the absence of kinetic expressions for all reactions associated with the metabolites in I^{kin} a number of degrees of freedom remain for the metabolite concentrations. To avoid drastic concentration changes in response to the overproduction goal we explored penalizing deviations of metabolite concentrations from their reference steady-state values using a weight penalty factor ε . This

posture in essence imposes a homeostasis term in the optimization objective function. The outer objective function of the bilevel formulation for identifying FORCE sets is thus modified as follows:

$$\text{Maximize } (1-\varepsilon) \left(\frac{1}{v_{prod}^{max}} v_{prod} \right) - \varepsilon \left(\frac{1}{M^{kin}} \sum_{i \in I^{kin}} \left| \frac{c_i - c_i^{ss,ref}}{c_i^{ss,ref}} \right| \right)$$

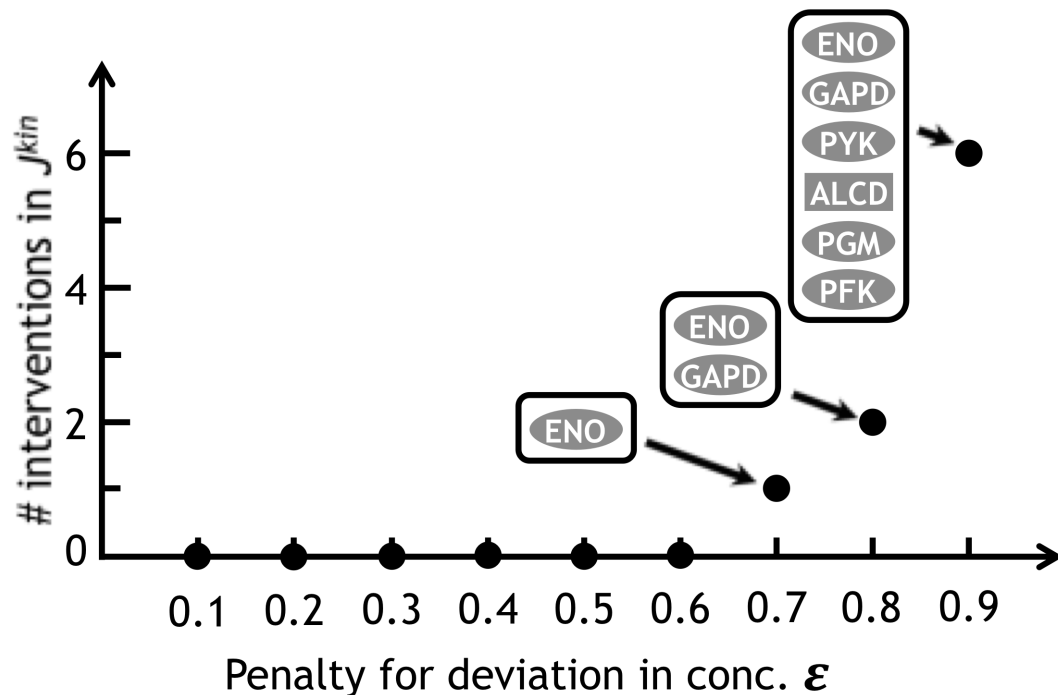
The first term in the objective function is the flux of the desired chemical v_{prod} , scaled by its theoretical maximum flux in the network v_{prod}^{max} . The second term is the average fractional departure of the metabolites in I^{kin} (c_i) from their reference values ($c_i^{ss,ref}$). M^{kin} represents the total number of metabolites in I^{kin} . When using the two-step procedure for identifying MUST sets (see Methods and Supplementary Material S1), the objective function for the first step is modified as follows:

$$\text{Minimize } (1-\varepsilon) \left(\frac{1}{MUST^{kin}} \sum_{j \in J^{kin}} y_j^U + y_j^L \right) + \varepsilon \left(\frac{1}{M^{kin}} \sum_{i \in I^{kin}} \left| \frac{c_i - c_i^{ss,ref}}{c_i^{ss,ref}} \right| \right)$$

In this objective function, the first term is the sum of interventions in J^{kin} , scaled by the total number of reactions in J^{kin} identified in MUST sets ($MUST^{kin}$). No other changes are made in the formulation of the algorithms. We tested this modified formulation on TAL production in *S. cerevisiae*.

Figure 11 describes the effect of penalizing concentration departures on enzymatic interventions in J^{kin} for the overproduction

of TAL in *S. cerevisiae*. We varied ε from 0.1 (low penalty) to 0.9 (very high penalty) on the identified interventions. Up to a ε value of 0.6, the penalty is not high enough to require direct interventions instead of concentration changes. For a ε value of 0.7 k-OptForce identified up-regulation of enolase (ENO) while maintaining the average deviation in concentration to 0.1053. Note that without the use of the penalty term the concentrations for 3pg and 2pg have to be elevated by 31% and 27% respectively, from their reference levels to redirect more flux through the lower glycolytic reactions of glyceraldehyde-3-phosphate dehydrogenase (PGM) and ENO. For $\varepsilon=0.8$ the metabolite concentrations remain even closer to their reference levels (average deviation is 0.0754) thereby requiring the up-regulation of glyceraldehyde-3-phosphate dehydrogenase (GAPD) in addition to ENO. For ε equal to or greater than 0.9 at least 6 additional enzymatic interventions in J^{kin} are necessary to increase glycolytic flux while concentrations remain very close to their



| $\frac{1}{M} \sum_{i \in I^{kin}} \left \frac{c_i - c_i^{ss,ref}}{c_i^{ss,ref}} \right $ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|------|------|------|------|------|------|------|------|-----|
| C_{2pg} | 27%↑ | 27%↑ | 27%↑ | 27%↑ | 27%↑ | 27%↑ | 14%↓ | 14%↓ | 0% |
| C_{3pg} | 31%↑ | 31%↑ | 31%↑ | 31%↑ | 31%↑ | 31%↑ | 0% | 0% | 0% |
| C_{g3p} | 6%↑ | 6%↑ | 6%↑ | 6%↑ | 6%↑ | 6%↑ | 6%↑ | 2%↓ | 3%↓ |
| C_{fdp} | 16%↑ | 16%↑ | 16%↑ | 16%↑ | 16%↑ | 16%↑ | 16%↑ | 0% | 0% |
| C_{pep} | 12%↑ | 12%↑ | 12%↑ | 12%↑ | 12%↑ | 12%↑ | 21%↑ | 21%↑ | 0% |

% deviation in concentration from their reference values with increase in ε

Figure 11. Predicted interventions in J^{kin} for TAL overproduction in *S. cerevisiae* as a function of penalty factor (ε). The graph above shows the variation in the minimum number of interventions required in J^{kin} to overproduce TAL with increase in ε . The corresponding table below shows of the average deviation (row 1) and the individual deviation in concentration of metabolites in J^{kin} for different values of ε . doi:10.1371/journal.pcbi.1003487.g011

reference values. By increasing the value of the penalty ϵ enables drawing trade-offs between allowable concentration changes and minimality of needed interventions for an overproduction goal while also providing a prioritization strategy for implementing engineering interventions.

Expression of a heterologous enzyme in *S. cerevisiae* to improve TAL yield

The conversion from cytosolic pyruvate to acetyl-CoA (precursor for TAL) in yeast follows a long and tightly regulated path involving the intermediate production of acetaldehyde and acetate [101]. We sought to computationally explore the use of a direct route from pyruvate to acetyl-CoA by adding a heterologous cytosolic PDH from *E. coli* in *S. cerevisiae* that directly converts pyruvate to acetyl-CoA. Note that *S. cerevisiae* has a pyruvate dehydrogenase activity in mitochondria but not in cytosol. The PDH complex in *E. coli* uses nad as the cofactor, however, an nadp-dependent PDH enzyme (constructed by site-directed mutagenesis in the $\beta\alpha\beta$ fold of the nad-binding domain of dihydrolipoamide dehydrogenase) has also been expressed in *E. coli* with identical kinetic properties [102]. The maximum theoretical yield of TAL using the nadp-dependent PDH enzyme increased by 40%. By bypassing the multi-step conversion of pyruvate to acetyl-CoA, two ATP equivalents of energy are conserved. No such maximum yield improvements are found for the nad-dependent PDH due to nad imbalance in the cytosol.

The kinetic expression for PDH was extracted from the kinetic model of *E. coli* proposed in Chassagnole *et al.* [76]. The interventions predicted by k-OptForce for maximizing TAL production are shown in Figure 12. Upon addition of the heterologous PDH, the entire amount of flux towards TAL production is routed through PDH. This eliminates the ACS activity that previously drained ATP. Pyruvate decarboxylase (PYRDC) is now down-regulated, but its activity is not reduced to zero. The entire flux of PYRDC goes towards ethanol production to regenerate NAD and maintain the cofactor balance in cytosol [101]. Instead of up-regulating ALDD, k-OptForce identifies an alternative intervention to lower nadph availability by up-regulating either aspartate kinase (ASPK), threonine synthase (THRS) or cystathionine synthase (METB1) in the hydroxybutyrate production pathway.

Discussion

k-OptForce integrates kinetic relations (whenever available) with stoichiometry based models to identify genetic perturbations that are consistent with enzyme expressions and metabolite concentrations. The resulting optimization problems pose significant computational challenges due to the bilevel nature of the formulation and the nonconvex terms in the objective function and constraints. We introduced tractable solution workflows for recasting the problems as equivalent single-level mixed-integer nonlinear optimization problems (MINLP) solved using the global optimization solver BARON to optimality. A hierarchical decomposition approach is also introduced for first identifying interventions within the kinetic part of the model followed by the interventions in the stoichiometry-only part of the model. As with other computational algorithms that make use of kinetics the results can be dependent upon the kinetic model structure and parameterization.

Computational results show that the introduction of kinetic expressions in strain design can significantly affect the identified interventions in sometimes non-intuitive ways. In some cases additional modifications are needed to substitute interventions that

cause enzyme saturation or concentration bound violations. The mechanism of action of these modifications is often subtle by alleviating substrate inhibition or draining away cofactors from competing pathways. In other cases, kinetic expressions shape flux distributions so as to favor the overproduction of the desired product requiring fewer direct interventions. Uncertainties in both the accuracy of the kinetic models and allowable concentration ranges imply that predicted interventions need to be carefully scrutinized to pinpoint the reasons for their inclusion. An important finding in this study was that concentration ranges have a very significant effect on the identified interventions. By penalizing departures of concentrations from the reference strain values substantial re-arrangements in the predicted interventions are observed. Each one of these changes can be analyzed and the underlying reason for its inclusion can be identified.

A key contribution of kinetic descriptions is that they can attribute performance bottlenecks to specific concentration bounds and/or enzymatic parameter ranges bottlenecks revealing avenues for model improvement and strain optimization. The case study for L-serine overproduction in *E. coli* provides an example of how k-OptForce can be used to both identify interventions and trace the reason(s) for the exclusion of others. k-OptForce revealed that inhibition of upper glycolysis by pep and 6pgc must be alleviated to route more flux towards L-serine. This is achieved through removal of PDH and down-regulation of TALA or TKT1 respectively. Flux analysis on single gene mutant strains of *E. coli* show that deletion of either *tala* or *tktA* increases the flux through PGI [103] corroborating k-OptForce predictions. However, MFA data for the *lpdA* mutant encoding the PDH enzyme in *E. coli* [104] showed that the upper glycolysis is down-regulated and that the flux through PP pathway is up-regulated, contrary to k-OptForce predictions. A possible reason for this discrepancy could be due to insufficiencies in the kinetic expressions used to describe the reactions in J^{kin} . Alternatively, since *lpdA* also encodes for the activity of ICDHy and the glycine cleavage system (GLYCL), its removal could have a combined effect on down-regulating the flux in upper glycolysis, which is not captured by the kinetic model. k-OptForce, however, correctly predicts that PDH removal down-regulates lower glycolysis which is observed in the *lpdA* mutant strain [104]. Down-regulation of lower glycolysis is necessary to prevent the flux towards L-serine from draining away towards pyruvate. k-OptForce also sets an upper limit on the activity of the L-serine synthase pathway that the original OptForce procedure failed to pinpoint. In addition, k-OptForce prevents rearrangement of fluxes that would violate kinetic constraints and metabolite concentration limits. The original OptForce suggested reversal of lower glycolysis by rerouting metabolic flux through ED pathway. However, such re-distribution results in the upper and lower glycolysis to operate in opposite directions which cannot be achieved as the same regulator, (i.e., *cra*), determines the directionality of both upper (i.e., PFK) and lower glycolysis (i.e., PYK) [105,106] and represses the ED pathway upon reversal of glycolysis.

k-OptForce may require fewer direct interventions if the kinetic expressions shape fluxes so as to favor the desired overproduction product as observed for the production of TAL where up-regulation of ALDD was suggested to redirect flux from pyruvate to acetyl-CoA. This is consistent with an experimental study for isoprenoid overproduction in *S. cerevisiae* [107] which demonstrated that overexpression of *ald6* (which encodes for the ALDD enzyme) increases flux towards acetyl-CoA. However, a fraction of the flux from pyruvate to acetaldehyde was routed through threonine degradation without the requirement of any additional interventions. This direct intervention-free flux redistribution may

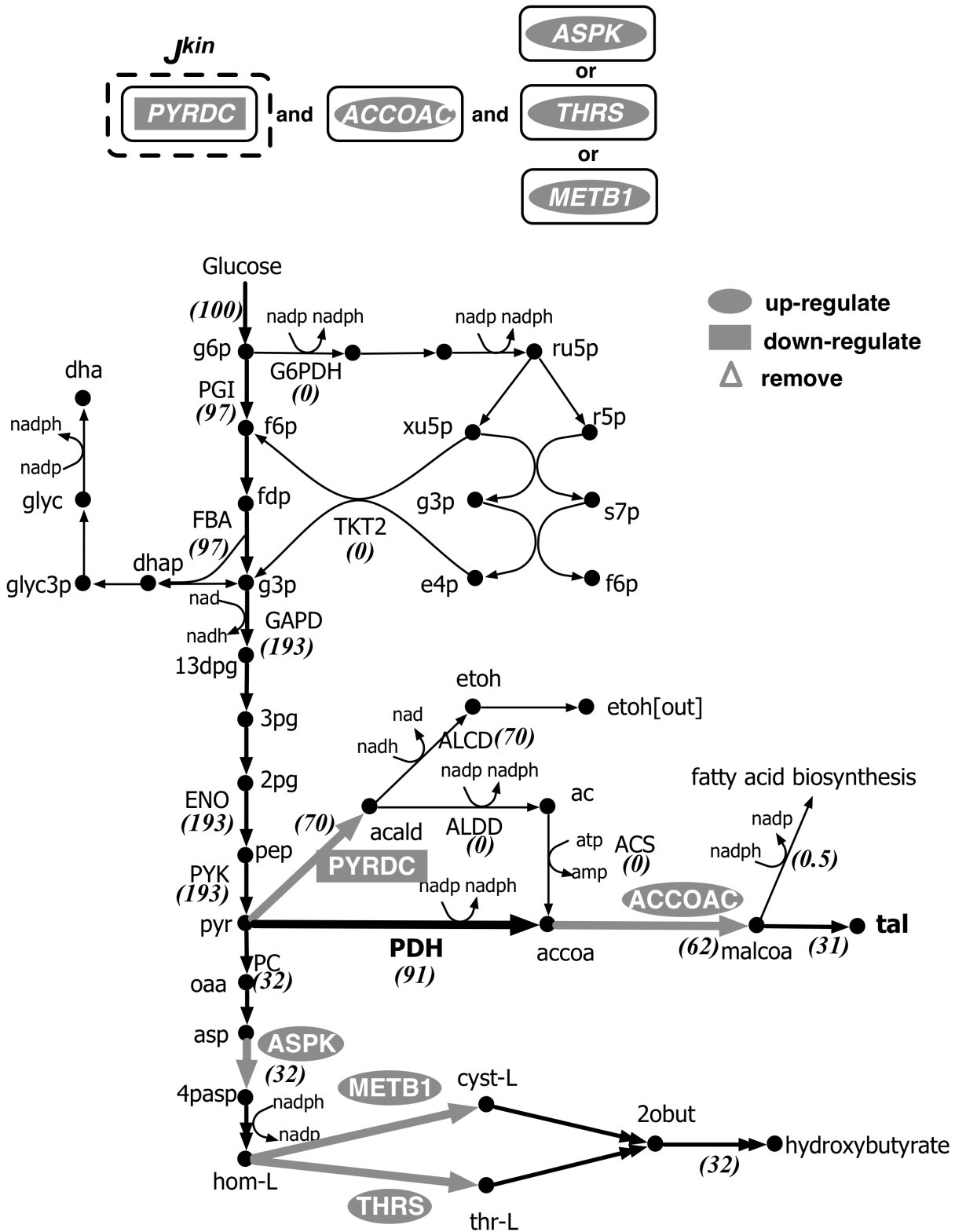


Figure 12. Metabolic interventions predicted by k-OptForce for overproduction of TAL in *S. cerevisiae* on heterologous expression of nadph-dependent PDH complex from *E. coli*. The values in brackets indicate the metabolic flux in mmol gDW⁻¹hr⁻¹ per 100 mmol gDW⁻¹ hr⁻¹ glucose uptake.
doi:10.1371/journal.pcbi.1003487.g012

be an artifact of the kinetic model and may require direct manipulations to engineer. Metabolome studies on single-gene mutant analysis in *E. coli* [103] revealed that, on average, internal metabolite concentrations were minimally altered from their reference concentrations as a result of the genetic perturbations. Changes in metabolic fluxes were largely the result of changes in enzyme activities. In response to this we postulated the use of a penalty term for violating homeostasis of metabolite levels. Alternatively, one could employ the method described in Smallbone *et al* [73] to formulate approximate lin-log expressions for all reactions associated with metabolites in J^{kin} that do not have a kinetic expression (i.e. not part of J^{kin}). This would restore the square-system of equations in J^{kin} and recouple all metabolite concentrations with enzyme activities.

The k-OptForce procedure is versatile enough to incorporate additional omics information, whenever available, to further improve prediction fidelity. For example, MFA data for reactions can be included as additional constraints to further tighten flux ranges. k-OptForce can also capture other types of metabolic regulation and select from a wider palette of direct interventions (e.g., enzymatic changes and transcriptional control) such as the dynamic hybrid model of *E. coli* metabolism by Lee *et al* [108] that integrates signaling and transcriptional regulation with FBA. Temporal consideration can also be addressed by deploying

k-OptForce within the dFBA framework [65] to explore the variation of metabolic interventions as a function of time alluding to RNAi type of interventions. We expect that k-OptForce predictions will help improve the breadth and accuracy of kinetic modeling descriptions by providing the quantitative means to assess model accuracy ultimately leading to improved fidelity of metabolic descriptions.

Supporting Information

Text S1 Optimization formulation and solution procedure of k-OptForce, kinetic models of central metabolism for *E. coli* and *S. cerevisiae*, supplementary figures for steady-state flux distribution and metabolite concentrations for the kinetic models, and flux ranges for the reference and overproducing phenotypes for L-serine and TAL production.

(PDF)

Author Contributions

Conceived and designed the experiments: AC ARZ CDM. Performed the experiments: AC. Analyzed the data: AC ARZ CDM. Contributed reagents/materials/analysis tools: AC ARZ CDM. Wrote the paper: AC ARZ CDM.

References

- Curran KA, Alper HS (2012) Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metab Eng* 14: 289–297.
- Hong KK, Nielsen J (2012) Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cell Mol Life Sci* 69: 2671–2690.
- Lee JW, Na D, Park JM, Lee J, Choi S, et al. (2012) Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat Chem Biol* 8: 536–546.
- Burgard AP, Pharkya P, Maranas CD (2003) OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng* 84: 647–657.
- Kim J, Reed JL, Maravelias CT (2011) Large-scale bi-level strain design approaches and mixed-integer programming solution techniques. *PLoS One* 6: e24162.
- Maia P, Vilaca P, Rocha I, Pont M, Tomb JF, et al. (2012) An integrated computational environment for elementary modes analysis of biochemical networks. *Int J Data Min Bioinform* 6: 382–395.
- Pharkya P, Burgard AP, Maranas CD (2004) OptStrain: a computational framework for redesign of microbial production systems. *Genome Res* 14: 2367–2376.
- Pharkya P, Maranas CD (2006) An optimization framework for identifying reaction activation/inhibition or elimination candidates for overproduction in microbial systems. *Metab Eng* 8: 1–13.
- Ranganathan S, Suthers PF, Maranas CD (2010) OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. *PLoS Comput Biol* 6: e1000744.
- Xu P, Ranganathan S, Fowler ZL, Maranas CD, Koffas MA (2011) Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. *Metab Eng* 13: 578–587.
- Yang L, Cluett WR, Mahadevan R (2011) EMILiO: a fast algorithm for genome-scale strain design. *Metab Eng* 13: 272–281.
- Gombert AK, Nielsen J (2000) Mathematical modelling of metabolism. *Curr Opin Biotechnol* 11: 180–186.
- Kadir TA, Mannan AA, Kierzek AM, McFadden J, Shimizu K (2010) Modeling and simulation of the main metabolism in *Escherichia coli* and its several single-gene knockout mutants with experimental verification. *Microb Cell Fact* 9: 88.
- Rizzi M, Baltes M, Theobald U, Reuss M (1997) In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol Bioeng* 55: 592–608.
- Heinrich R, Rapoport TA (1974) A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. *Eur J Biochem* 42: 89–95.
- van Riel NA (2006) Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. *Brief Bioinform* 7: 364–374.
- Heijnen JJ (2005) Approximative kinetic formats used in metabolic network modeling. *Biotechnol Bioeng* 91: 534–545.
- Wu L, Wang W, van Winden WA, van Gulik WM, Heijnen JJ (2004) A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics. *Eur J Biochem* 271: 3348–3359.
- del Rosario RC, Mendoza E, Voit EO (2008) Challenges in lin-log modelling of glycolysis in *Lactococcus lactis*. *IET Syst Biol* 2: 136–149.
- Hatzimanikatis V, Emmerling M, Sauer U, Bailey JE (1998) Application of mathematical tools for metabolic design of microbial ethanol production. *Biotechnol Bioeng* 58: 154–161.
- Wang FSKCLVEO (2006) Kinetic modeling of using S-systems and lin-log approaches. *Biochemical Engineering* 33: 238–247.
- Jamshidi N, Palsson BO (2008) Formulating genome-scale kinetic models in the post-genome era. *Mol Syst Biol* 4: 171.
- Pozo C, Marin-Sanguino A, Alves R, Guillen-Gosalbez G, Jimenez L, et al. (2011) Steady-state global optimization of metabolic non-linear dynamic models through recasting into power-law canonical models. *BMC Syst Biol* 5: 137.
- Sorribas A, Hernandez-Bermejo B, Vilaprinyo E, Alves R (2007) Cooperativity and saturation in biochemical networks: a saturable formalism using Taylor series approximations. *Biotechnol Bioeng* 97: 1259–1277.
- Liebermeister W, Klipp E (2006) Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor Biol Med Model* 3: 41.
- Varner J, Ramkrishna D (1999) Metabolic engineering from a cybernetic perspective. 1. Theoretical preliminaries. *Biotechnol Prog* 15: 407–425.
- Varner J, Ramkrishna D (1999) Metabolic engineering from a cybernetic perspective: aspartate family of amino acids. *Metab Eng* 1: 88–116.
- Varner J, Ramkrishna D (1999) Metabolic engineering from a cybernetic perspective. 2. Qualitative investigation of nodal architectures and their response to genetic perturbation. *Biotechnol Prog* 15: 426–438.
- Chou IC, Voit EO (2009) Recent developments in parameter estimation and structure identification of biochemical and genomic systems. *Math Biosci* 219: 57–83.
- Costa RS, Machado D, Rocha I, Ferreira EC (2010) Hybrid dynamic modeling of *Escherichia coli* central metabolic network combining Michaelis-Menten and approximate kinetic equations. *Biosystems* 100: 150–157.
- Famili I, Mahadevan R, Palsson BO (2005) k-Cone analysis: determining all candidate values for kinetic parameters on a network scale. *Biophys J* 88: 1616–1625.
- Miskovic L, Hatzimanikatis V (2011) Modeling of uncertainties in biochemical reactions. *Biotechnol Bioeng* 108: 413–423.
- Steuer R, Gross T, Selbig J, Blasius B (2006) Structural kinetic modeling of metabolic networks. *Proc Natl Acad Sci U S A* 103: 11868–11873.
- Tran LM, Rizk ML, Liao JC (2008) Ensemble modeling of metabolic networks. *Biophys J* 95: 5606–5617.
- Angermayr SA, Hellingwerf KJ (2013) On the use of Metabolic Control Analysis in the Optimization of Cyanobacterial Biosolar-Cell Factories. *J Phys Chem B* 117: 11169–11175.

36. Chassignole C, Fell DA, Rais B, Kudla B, Mazat JP (2001) Control of the threonine-synthesis pathway in *Escherichia coli*: a theoretical and experimental approach. *Biochem J* 356: 433–444.
37. Contador CA, Rizk ML, Asenjo JA, Liao JC (2009) Ensemble modeling for strain development of L-lysine-producing *Escherichia coli*. *Metab Eng* 11: 221–233.
38. Hoefnagel MH, Starrenburg MJ, Martens DE, Hugenholtz J, Kleerebezem M, et al. (2002) Metabolic engineering of lactic acid bacteria, the combined approach: kinetic modelling, metabolic control and experimental analysis. *Microbiology* 148: 1003–1013.
39. Nikolaev EV (2010) The elucidation of metabolic pathways and their improvements using stable optimization of large-scale kinetic models of cellular systems. *Metab Eng* 12: 26–38.
40. Rizk ML, Liao JC (2009) Ensemble modeling for aromatic production in *Escherichia coli*. *PLoS One* 4: e6903.
41. Song HS, Ramkrishna D (2012) Prediction of dynamic behavior of mutant strains from limited wild-type data. *Metab Eng* 14: 69–80.
42. Vital-Lopez FG, Armaou A, Nikolaev EV, Maranas CD (2006) A computational procedure for optimal engineering interventions using kinetic models of metabolism. *Biotechnol Prog* 22: 1507–1517.
43. Zomorodi AR, Lafontaine Rivera JG, Liao JC, Maranas CD (2013) Optimization-driven identification of genetic perturbations accelerates the convergence of model parameters in ensemble modeling of metabolic networks. *Biotechnol J* 8(9):1090–104.
44. Durot M, Bourguignon PY, Schachter V (2009) Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiol Rev* 33: 164–190.
45. Price ND, Reed JL, Palsson BO (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat Rev Microbiol* 2: 886–897.
46. Saha R, Suthers PF, Maranas CD (2011) *Zea mays* iRS1563: a comprehensive genome-scale metabolic reconstruction of maize metabolism. *PLoS One* 6: e21784.
47. Varma A, Palsson BO (1993) Metabolic capabilities of *Escherichia coli*: I. synthesis of biosynthetic precursors and cofactors. *J Theor Biol* 165: 477–502.
48. Mahadevan R, Schilling CH (2003) The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab Eng* 5: 264–276.
49. Edwards JS, Ibarra RU, Palsson BO (2001) In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat Biotechnol* 19: 125–130.
50. Park JM, Kim TY, Lee SY (2009) Constraints-based genome-scale metabolic simulation for systems metabolic engineering. *Biotechnol Adv* 27: 979–988.
51. Zomorodi AR, Suthers PF, Ranganathan S, Maranas CD (2012) Mathematical optimization applications in metabolic networks. *Metab Eng* 14: 672–686.
52. Cotten C, Reed JL (2013) Constraint-based strain design using continuous modifications (CosMos) of flux bounds finds new strategies for metabolic engineering. *Biotechnol J* 8: 595–604.
53. Ibarra RU, Edwards JS, Palsson BO (2002) *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* 420: 186–189.
54. Segre D, Vitkup D, Church GM (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci U S A* 99: 15112–15117.
55. Kim J, Reed JL (2010) OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Syst Biol* 4: 53.
56. Shlomi T, Berkman O, Ruppin E (2005) Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc Natl Acad Sci U S A* 102: 7695–7700.
57. Feist AM, Zielinski DC, Orth JD, Schellenberger J, Herrgard MJ, et al. (2010) Model-driven evaluation of the production potential for growth-coupled products of *Escherichia coli*. *Metab Eng* 12: 173–186.
58. Tepper N, Shlomi T (2010) Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways. *Bioinformatics* 26: 536–543.
59. Bro C, Regenb B, Forster J, Nielsen J (2006) In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metab Eng* 8: 102–111.
60. Asadollahi MA, Maury J, Patil KR, Schalk M, Clark A, et al. (2009) Enhancing sesquiterpene production in *Saccharomyces cerevisiae* through in silico driven metabolic engineering. *Metab Eng* 11: 328–334.
61. Alper H, Jin YS, Moxley JF, Stephanopoulos G (2005) Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab Eng* 7: 155–164.
62. Park JH, Lee KH, Kim TY, Lee SY (2007) Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proc Natl Acad Sci U S A* 104: 7797–7802.
63. Lee SJ, Lee DY, Kim TY, Kim BH, Lee J, et al. (2005) Metabolic engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and in silico gene knockout simulation. *Appl Environ Microbiol* 71: 7880–7887.
64. Choi HS, Lee SY, Kim TY, Woo HM (2010) In silico identification of gene amplification targets for improvement of lycopene production. *Appl Environ Microbiol* 76: 3097–3105.
65. Mahadevan R, Edwards JS, Doyle EJ, 3rd (2002) Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys J* 83: 1331–1340.
66. Covert MW, Xiao N, Chen TJ, Karr JR (2008) Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics* 24: 2044–2050.
67. Meadows AL, Karnik R, Lam H, Forestell S, Snedecor B (2010) Application of dynamic flux balance analysis to an industrial *Escherichia coli* fermentation. *Metab Eng* 12: 150–160.
68. Feng X, Xu Y, Chen Y, Tang YJ (2012) Integrating flux balance analysis into kinetic models to decipher the dynamic metabolism of *Shewanella oneidensis* MR-1. *PLoS Comput Biol* 8: e1002376.
69. Zhuang K, Izallalen M, Mouser P, Richter H, Risso C, et al. (2011) Genome-scale dynamic modeling of the competition between *Rhodospirillum rubrum* and *Geobacter* in anoxic subsurface environments. *ISME J* 5: 305–316.
70. Salimi F, Zhuang K, Mahadevan R (2010) Genome-scale metabolic modeling of a clostridial co-culture for consolidated bioprocessing. *Biotechnol J* 5: 726–738.
71. Jamshidi N, Palsson BO (2010) Mass action stoichiometric simulation models: incorporating kinetics and regulation into stoichiometric models. *Biophys J* 98: 175–185.
72. Smallbone K, Simeonidis E, Broomhead DS, Kell DB (2007) Something from nothing: bridging the gap between constraint-based and kinetic modelling. *FEBS J* 274: 5576–5585.
73. Smallbone K, Simeonidis E, Swainston N, Mendes P (2010) Towards a genome-scale kinetic model of cellular metabolism. *BMC Syst Biol* 4: 6.
74. Fleming RM, Thiele I, Provan G, Nasheuer HP (2010) Integrated stoichiometric, thermodynamic and kinetic modelling of steady state metabolism. *J Theor Biol* 264: 683–692.
75. Cotten C, Reed JL (2013) Mechanistic analysis of multi-omics datasets to generate kinetic parameters for constraint-based metabolic models. *BMC Bioinformatics* 14: 32.
76. Chassignole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M (2002) Dynamic modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnol Bioeng* 79: 53–73.
77. Peskov K, Mogilevskaia E, Demin O (2012) Kinetic modelling of central carbon metabolism in *Escherichia coli*. *FEBS J* 279: 3374–3385.
78. Hynne F, Dano S, Sorensen PG (2001) Full-scale model of glycolysis in *Saccharomyces cerevisiae*. *Biophys Chem* 94: 121–163.
79. van Eunen K, Kiewiet JA, Westerhoff HV, Bakker BM (2012) Testing biochemistry revisited: how in vivo metabolism can be understood from in vitro enzyme kinetics. *PLoS Comput Biol* 8: e1002483.
80. Sahinidis NV (1996) BARON: A general purpose global optimization software package. *Journal of Global Optimization* 8: 201–205.
81. Peters-Wendisch P, Netzer R, Eggeling L, Sahn H (2002) 3-Phosphoglycerate dehydrogenase from *Corynebacterium glutamicum*: the C-terminal domain is not essential for activity but is required for inhibition by L-serine. *Appl Microbiol Biotechnol* 60: 437–441.
82. Klomp LW, de Koning TJ, Malingre HE, van Beurden EA, Brink M, et al. (2000) Molecular characterization of 3-phosphoglycerate dehydrogenase deficiency—a neurometabolic disorder associated with reduced L-serine biosynthesis. *Am J Hum Genet* 67: 1389–1399.
83. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, 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: 121.
84. Visser D, Schmid JW, Mauch K, Reuss M, Heijnen JJ (2004) Optimal redesign of primary metabolism in *Escherichia coli* using linlog kinetics. *Metab Eng* 6: 378–390.
85. Kacser H, Burns JA (1995) The control of flux. *Biochem Soc Trans* 23: 341–366.
86. Patrick WM, Quandt EM, Swartzlander DB, Matsumura I (2007) Multicopy suppression underpins metabolic evolvability. *Mol Biol Evol* 24: 2716–2722.
87. Joyce AR, Reed JL, White A, Edwards R, Osterman A, et al. (2006) Experimental and computational assessment of conditionally essential genes in *Escherichia coli*. *J Bacteriol* 188: 8259–8271.
88. Park JH, Lee SY (2008) Towards systems metabolic engineering of microorganisms for amino acid production. *Curr Opin Biotechnol* 19: 454–460.
89. Peters-Wendisch P, Stolz M, Eterich H, Kennerknecht N, Sahn H, et al. (2005) Metabolic engineering of *Corynebacterium glutamicum* for L-serine production. *Appl Environ Microbiol* 71: 7139–7144.
90. Stolz M, Peters-Wendisch P, Eterich H, Gerharz T, Faurie R, et al. (2007) Reduced folate supply as a key to enhanced L-serine production by *Corynebacterium glutamicum*. *Appl Environ Microbiol* 73: 750–755.
91. Lai S, Zhang Y, Liu S, Liang Y, Shang X, et al. (2012) Metabolic engineering and flux analysis of *Corynebacterium glutamicum* for L-serine production. *Sci China Life Sci* 55: 283–290.
92. McKittrick JC, Pizer LI (1980) Regulation of phosphoglycerate dehydrogenase levels and effect on serine synthesis in *Escherichia coli* K-12. *J Bacteriol* 141: 235–245.
93. Tohsato Y, Ikuta K, Shionoya A, Mazaki Y, Ito M (2013) Parameter optimization and sensitivity analysis for large kinetic models using a real-coded genetic algorithm. *Gene* 518: 84–90.

94. Achkar J, Xian M, Zhao H, Frost JW (2005) Biosynthesis of phloroglucinol. *J Am Chem Soc* 127: 5332–5333.
95. Hansen CA, Frost JW (2002) Deoxygenation of polyhydroxybenzenes: an alternative strategy for the benzene-free synthesis of aromatic chemicals. *J Am Chem Soc* 124: 5926–5927.
96. Xie D, Shao Z, Achkar J, Zha W, Frost JW, et al. (2006) Microbial synthesis of triacetic acid lactone. *Biotechnol Bioeng* 93: 727–736.
97. Richardson MT, Pohl NL, Kealey JT, Khosla C (1999) Tolerance and specificity of recombinant 6-methylsalicylic acid synthase. *Metab Eng* 1: 180–187.
98. Zha W, Shao Z, Frost JW, Zhao H (2004) Rational pathway engineering of type I fatty acid synthase allows the biosynthesis of triacetic acid lactone from D-glucose in vivo. *J Am Chem Soc* 126: 4534–4535.
99. Zomorodi AR, Maranas CD (2010) Improving the iMM904 *S. cerevisiae* metabolic model using essentiality and synthetic lethality data. *BMC Syst Biol* 4: 178.
100. Gombert AK, Moreira dos Santos M, Christensen B, Nielsen J (2001) Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J Bacteriol* 183: 1441–1451.
101. Pronk JT, Yde Steensma H, Van Dijken JP (1996) Pyruvate metabolism in *Saccharomyces cerevisiae*. *Yeast* 12: 1607–1633.
102. Bocanegra JA, Scrutton NS, Perham RN (1993) Creation of an NADP-dependent pyruvate dehydrogenase multienzyme complex by protein engineering. *Biochemistry* 32: 2737–2740.
103. Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, et al. (2007) Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science* 316: 593–597.
104. Li M, Ho PY, Yao S, Shimizu K (2006) Effect of *lpdA* gene knockout on the metabolism in *Escherichia coli* based on enzyme activities, intracellular metabolite concentrations and metabolic flux analysis by ¹³C-labeling experiments. *J Biotechnol* 122: 254–266.
105. Ramseier TM, Bledig S, Michotey V, Feghali R, Saier MH, Jr. (1995) The global regulatory protein FruR modulates the direction of carbon flow in *Escherichia coli*. *Mol Microbiol* 16: 1157–1169.
106. Ramseier TM (1996) Cra and the control of carbon flux via metabolic pathways. *Res Microbiol* 147: 489–493.
107. Shiba Y, Paradise EM, Kirby J, Ro DK, Keasling JD (2007) Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab Eng* 9: 160–168.
108. Lee JM, Gianchandani EP, Eddy JA, Papin JA (2008) Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Comput Biol* 4: e1000086.