Metabolic Engineering 13 (2011) 60–75

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/ymben

Reducing the allowable kinetic space by constructing ensemble of dynamic models with the same steady-state flux

Yikun Tan [a], Jimmy G. Lafontaine Rivera [a,1], Carolina A. Contador [b,1], Juan A. Asenjo [b], James C. Liao [a,*]

[a] Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, 5531 Boelter Hall, Los Angeles, CA 90095-1592, USA
[b] Department of Chemical Engineering and Biotechnology, Institute for Cell Dynamics and Biotechnology: A Center for Systems Biology, University of Chile, Santiago, Plaza Ercilla 847, Santiago, Chile

ARTICLE INFO

Article history:
Received 26 June 2010
Received in revised form
3 November 2010
Accepted 5 November 2010
Available online 12 November 2010

Keywords:
Ensemble modeling
Dynamic model
Steady-state flux

ABSTRACT

Dynamic models of metabolism are instrumental for gaining insight and predicting possible outcomes of perturbations. Current approaches start from the selection of lumped enzyme kinetics and determine the parameters within a large parametric space. However, kinetic parameters are often unknown and obtaining these parameters requires detailed characterization of enzyme kinetics. In many cases, only steady-state fluxes are measured or estimated, but these data have not been utilized to construct dynamic models. Here, we extend the previously developed Ensemble Modeling methodology by allowing various kinetic rate expressions and employing a more efficient solution method for steady states. We show that anchoring the dynamic models to the same flux reduces the allowable parameter space significantly such that sampling of high dimensional kinetic parameters becomes meaningful. The methodology enables examination of the properties of the model's structure, including multiple steady states. Screening of models based on limited steady-state fluxes or metabolite profiles reduces the parameter space further and the remaining models become increasingly predictive. We use both succinate overproduction and central carbon metabolism in *Escherichia coli* as examples to demonstrate these results.

Published by Elsevier Inc.

1. Introduction

The construction of dynamic models for describing metabolic behavior is based on both network stoichiometry, which is reasonably well understood, and enzyme kinetics, which is often unavailable. Current approaches of constructing dynamic models start from the selection of lumped enzyme kinetics (e.g. Michaelis–Menten) and obtaining available parameters, then optimize the remaining parameters within a large parametric space in order to fit experimentally measured metabolite concentration profiles (Chassagnole et al., 2002; Usuda et al., 2010; Visser and Heijnen, 2003).

However, detailed kinetic parameters are rare and difficult to measure. On the other hand, steady-state flux distributions are relatively easy to characterize or estimate. Most importantly, the steady-state fluxes along with approximate metabolite pool sizes determine the time scale for the system. Heretofore, these data have not been used to construct dynamic models that capture a system's response upon perturbation. To take advantage of these data, the recently developed Ensemble Modeling (EM) approach (Tran et al., 2008) constructs a set of dynamic models that achieve a desired steady-state flux while allowing all possible dynamics within the framework of the chosen kinetic mechanisms and thermodynamic constraints. The steady state provides an anchor and sets the time scale for the dynamic models once the rough metabolite pool sizes are known. After a perturbation, such as enzyme expression tuning (Alper et al., 2005; Chao and Liao, 1993; Kojima et al., 1996; Lin et al., 2005a, 2005b) or alteration of enzyme properties (Zhang et al., 2008), these models will reach different fluxes as expected. The different fluxes provide a basis for screening models against experimental data commonly generated in metabolic engineering efforts (Atsumi et al., 2008b; Fischer et al., 2008; Nielsen et al., 2009; Ro et al., 2006). We have shown that the models would converge to a small set with only a few rounds of screening and become increasingly predictive (Contador et al., 2009; Dean et al., 2009; Rizk and Liao, 2009; Tran et al., 2008).

In this work, we demonstrate two benefits of anchoring these models to the same steady-state flux in the EM approach. First, by confining all the dynamic models to the same steady-state flux, the allowable range of possible kinetic parameters is significantly reduced, which allows meaningful sampling schemes to explore the dynamic behavior of the model. Second, by reducing the kinetic parameter space, further screening of models based on limited data (steady-state fluxes or transient metabolite profiles) becomes possible. Without such anchoring constraints, the parameter space is too large to be sampled effectively. Thus, EM differs from the traditional random sampling approaches in that it utilizes the

* Corresponding author. Fax: +1 310 206 4107.
E-mail address: liaoj@seas.ucla.edu (J.C. Liao).
[1] These authors are regarded as joint Second Authors.

1096-7176/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.ymben.2010.11.001

Abbreviations and nomenclature

Enzyme name

pts	phosphotransferase system
pgi	phosphoglucose isomerase
pfk	6-phosphofructokinase
fba	fructose bisphosphate aldolase
tpi	friose phosphate isomerase
gap	glyceraldehyde 3-phosphate dehydrogenase
pgk	3-phosphoglycerate kinase
gpm	phosphoglycerate mutase
eno	enolase
pyk	pyruvate kinase
pdh	pyruvate dehydrogenase
ppc	phosphoenolpyruvate carboxylase
g3pdh	glycerol 3-phosphate dehydrogenase
pgm	phosphoglucomutase
g1pat	glucose-1-phosphate adenylyltransferase
ser_synt	serine synthesis
zwf	glucose-6-phosphate-1-dehydrogenase
gnd	6-phosphogluconate dehydrogenase
rpe	ribulose phosphate 3-epimerase
rpi	ribose-5-phosphate isomerase
tkt1	transketolase 1
tkt2	transketolase 2
tal	transaldolase
aroG	2-dehydro-3-deoxyphosphoheptonate aldolase
rppk	ribose-phosphate pyrophosphokinase
synt1	chorismate synthesis
synt2	isoleucine, alanine, lysine, valine synthesis
trp_synt	tryptophan synthesis
met_trp_synt	methionine, tryptophan synthesis
pfl	pyruvate formate-lyase
glt	citrate synthase
acn	aconitase
icd	isocitrate dehydrogenase
sucAB	2-oxoglutarate dehydrogenase complex
sucCD	succinyl-CoA synthetase
sdh	succinate dehydrogenase
fum	fumarase
mdh	malate dehydrogenase
aceA	isocitrate lyase
aceB	malate synthase A
pyc	pyruvate carboxylase
poxB	pyruvate oxidase
pta	phosphate acetyltransferase
ackA	acetate kinase
pgl	6-phosphogluconolactonase
glk	glucokinase

Metabolite name

G6P	glucose-6-phosphate
F6P	fructose-6-phosphate
FBP	fructose-1,6-bisphosphate
DHAP	dihydroxyacetone phosphate
GAP	glyceraldehdye-3-phosphate
BPG	1,3-diphosphateglycerate
3PG	3-phosphoglycerate
2PG	2-phosphoglycerate
PEP	phosphoenolpyruvate
PYR	pyruvate
SER	serine
G3P	glyceral-3-phosphate
G1P	glucose-1-phosphate
ADPGlU	ADP-glucose
6PG	6-phosphogluconate
Ru5P	ribulose-5-phosphate
X5P	xylulose-5-phosphate
R5P	ribose-5-phosphate
S7P	sedoheptulose-7-phosphate
E4P	erythrose-4-phosphate
DAHP	3-deoxy-D-arabino-hepulosonate-7-phosphate
Intermediate 1	Intermediate 1
ACCOA	acetyl-CoA
OAA	oxaloacetate
Intermediate 2	Intermediate 2
Intermediate 3	Intermediate 3
CIT	citrate
ICIT	D-isocitrate
AKG	2-ketoglutarate
SUCCOA	succinyl-CoA
SUC	succinate
FUM	fumarate
MAL	malate
GLYX	glyoxylate
AC	acetate
ACTP	acetylphosphate
6PGL	6-phosphogluconolactone
6PGD	6-phosphogluconate
ATP	adenosine-triphosphate
ADP	adenosine-diphosphate
AMP	adenosine-monophosphate
NADH	diphosphopyridine nucleotide reduced
NAD	diphosphopyridine nucleotide
NADPH	nicotinamide adenine dinucleotide phosphate-reduced
NADP	nicotinamide adenine dinucleotide phosphate

available flux data and constrains the sampling space to a realistic space. These benefits are demonstrated first by using a simplified metabolic scheme and then two *Escherichia coli* models describing 1) succinate production and 2) central carbon metabolism.

In the previous EM work (Contador et al., 2009; Dean et al., 2009; Rizk and Liao, 2009; Tran et al., 2008), the elementary reaction rate law was used to model individual enzymatic reactions, since it is the most fundamental description of enzymatic mechanisms. The use of the elementary reaction kinetics enables the application of this approach to any metabolic network even when the enzymes are regulated at the protein level. In this work we further expand the concept of EM to construct models using lumped nonlinear kinetic rate laws, such as the Michaelis–Menten kinetics, Hill equation, and allosteric enzyme kinetics. This expansion makes the EM framework more flexible and applicable to a variety of situations.

Additionally, to bypass the heavy computational cost during the seeking of steady states using numerical integration method, we recast the system of material balance equations through a unique mathematical transformation (Savageau and Voit, 1987; Savageau, 1993) and solve it numerically by Newton's method. Such recasting technique is applicable when elementary enzyme kinetics are used and only steady state solutions are needed. The significant reduction of computational time makes the exploration of the allowable kinetic space more efficient, and can potentially open up a window of applying the EM approach for genome-scale model construction.

2. Methods

2.1. Dynamic models

The dynamic model is constructed by building a system of ordinary differential equations (ODEs) that consist of mass balances for intracellular metabolites that are involved in the metabolic system of interest, and the chosen kinetic rate expressions for individual enzymes are directly integrated in the mass balances as shown below:

$$\frac{dX}{dt} = SV \tag{1}$$

where X is the vector of intracellular metabolites, S is the stoichiometric matrix of the metabolic system, and V is the vector consisting of the collection of enzyme kinetic rate expressions as functions of X and kinetic parameters. If elementary reaction kinetics is used, the various enzyme complexes are included in X (Tran et al., 2008).

2.2. Enzyme kinetics

Similarly to previous EM work (Contador et al., 2009; Dean et al., 2009; Rizk and Liao, 2009; Tran et al., 2008), elementary reaction kinetics is used to describe individual enzyme reactions in the succinate production model. In addition, lumped enzyme kinetics such as Michaelis–Menten enzyme kinetics, Hill equation and allosteric enzyme kinetics (Cornish-Bowden, 1979; Hofmeyr and Cornish-Bowden, 1997) are used to model *E. coli* central metabolism. The detail formularization can be found in the Supplemental Material I and Table 1.

2.3. Thermodynamic constraints

For the lumped enzyme kinetics: reversible Michaelis–Menten and allosteric enzyme kinetics, the equilibrium constant k_{eq} can be calculated according the thermodynamic property of the specific enzyme as shown in the following equation:

$$k_{eq} = \exp\left(-\frac{\Delta G^o}{RT}\right) \tag{2}$$

where standard Gibbs free energies ΔG^o of enzymes involve in *E. coli* metabolism can be found in the literature (Henry et al., 2006).

For elementary reaction kinetics, the combined reversibilities $R_{i,j}$ of all the reaction steps catalyzed by the same enzyme are constrained by the Gibbs free energy ΔG_i of the overall reaction:

$$\sum_{j=1}^{n_i} \ln R_{i,j} = sign(V_{i,net})\frac{\Delta G_i}{RT} \tag{3}$$

where $sign(V_{i,net})$ denotes the direction of the net reaction catalyzed by enzyme i, positive for a forward proceeding reaction and negative for a reaction in the reverse direction. The Gibbs free energy is bounded by the standard Gibbs free energy by allowing metabolite levels to fluctuate in the following inequality:

$$\left(\frac{\Delta G_i}{RT}\right)_{lower\ bound} \le sign(V_{i,net}^{ref})\sum_j \ln R_{i,j}^{ref} \le \left(\frac{\Delta G_i}{RT}\right)_{upper\ bound} \tag{4}$$

where $(\Delta G_i/RT)_{lower\ bound}$ and $(\Delta G_i/RT)_{upper\ bound}$ reflect the Gibbs free energies at the lower and upper extremes, respectively. The derivation of this inequality and its relationship with elementary reaction rate constant has been previously described by Tran et al. (2008).

2.4. Constructing ensemble of models that all reach the same steady-state flux

Previously, elementary reaction enzyme kinetics was used in EM to construct an ensemble of models that can all reach the same steady state (Contador et al., 2009; Dean et al., 2009; Rizk and Liao, 2009; Tran et al., 2008). Through the normalization by steady-state metabolite concentrations, the knowledge of absolute steady-state concentrations becomes optional (Tran et al., 2008). By sampling the reversibilities of individual elementary reaction steps under thermodynamic constraint and enzyme complex fractions, the rate constants can be calculated. The detail derivation can be found in Supplemental Material II.

The concept of EM has been extended to construct such steady-state flux constrained dynamic models by using lumped enzyme

Table 1
The lumped kinetic rate expressions and their kinetic parameter sampling strategies.

Reaction	Enzyme kinetic type	Kinetic rate expression	Sampled parameters	Calculated parameters
$x_1 \rightleftharpoons x_2$	Reversible Michealis–Menten kinetic rate expression	$v = \frac{(v^f/k_m^1)(x_1 - (x_2/k_{eq}))}{1 + (x_1/k_m^1) + (x_2/k_m^2)}$	k_m^1, k_m^2	k_{eq},v^f
	Hill equation	$v = \frac{v_{\max}(x_1)^h}{k_m + x_1{}^h}$	k_m,h	$v_{\max}$
	Allosteric kinetic rate expression	$v = \frac{(v^f/k_m^1)(x_1 - (x_2/k_{eq}))((x_1/k_m^1) + (x_2/k_m^2))^{h-1}}{\prod_{j=1}^{j=m} \frac{1 + (M_j/k_{a,j})^h}{1 + \alpha_j (M_j/k_{a,j})^h} + ((x_1/k_m^1) + (x_2/k_m^2))^h}$	k_m^1,k_m^2,h,$k_{a,j}$,α_j M_j is the alloesteric regulator m is the number of regulators if $\alpha_j < 1$, M_j is an inhibitor if $\alpha_j > 1$, M_j is an activator if $\alpha_j = 1$, M_j has no effect on the rate	k_{eq},v^f
$x_1 + x_2 \rightleftharpoons x_3 + x_4$	Reversible Michealis–Menten kinetic rate expression	$v = \frac{(v^f/(k_i^1 k_m^2))(x_1 x_2 - ((x_3 x_4)/k_{eq}))}{1 + \frac{x_1}{k_i^1} + \frac{x_2}{k_i^2} + \frac{x_3}{k_i^3} + \frac{x_4}{k_i^4} + \frac{x_1 x_2}{k_i^1 k_m^2} + \frac{x_3 x_4}{k_m^3 k_i^4}}$	k_i^1,k_i^2,k_i^3,k_i^4, k_m^2,k_m^3	k_{eq},v^f
	Hill equation	$v = \frac{v_{\max}(x_1)^{h_1}(x_2)^{h_2}}{(k_m^1 + x_1^{h1})(k_m^2 + x_2^{h2})}$	k_m^1,k_m^2,h_1,h_2	v_{max}
	Allosteric kinetic rate expression	$term\ 1 = \frac{x_1}{k_i^1} + \frac{x_2}{k_i^2} + \frac{x_3}{k_i^3} + \frac{x_4}{k_i^4} + \frac{x_1 x_2}{k_i^1 k_m^2} + \frac{x_3 x_4}{k_m^3 k_i^4}$ $v = \frac{(v^f/(k_i^1 k_m^2))(x_1 x_2 - ((x_3 x_4)/k_{eq}))term\ 1^{h-1}}{\prod_{j=1}^{j=m} \frac{1 + (M_j/k_{a,j})^h}{1 + \alpha_j (M_j/k_{a,j})^h} + term\ 1^h}$	k_i^1,k_i^2,k_i^3,k_i^4, k_m^2,k_m^3,$k_{a,j}$,α_j,h	k_{eq},v^f

Box 1–Mathematical summary of the Ensemble Modeling approach.

Modeling objectives

Flux-oriented objective:
To match model-predicted flux with data upon enzyme tuning

$$\min\sum_i \left(V_{i,ss}\left(\frac{E_{tot}+\Delta E_{tot}}{E_{tot}}\right)-J_{i,ss}^{measured}\left(\frac{E_{tot}+\Delta E_{tot}}{E_{tot}}\right)\right)^2$$

Pros:
- Objective function directly reflects metabolic engineering needs
- Flux readily measureable
- Effects of kinetics on flux accountable

OR

Metabolite-oriented objective:
To match metabolite profiles with data $\min\sum_i \left(x_i(t)-x_i^{measured}(t)\right)^2$

Pro:
- Direct test of ODE

Con:
- Not directly applicable to metabolic engineering

by changing K and E_{tot}, such that

$$\frac{dX}{dt}=F(X,K,E_{tot}),\quad V_{ss}=G(X_{ss},\ K,\ E_{tot},\ \Delta E_{tot})$$

where ss represents steady state of the ODE, G represents the function of steady-state flux.

Ensemble modeling features

(1) Use an additional constraint for the reference strain $(\Delta E_{tot}=0)V_{i,ss}=G(X_{ss},K,E_{tot},\Delta E_{tot}=0)=J_{i,ss}^{measured}$

Pros:
- Anchor the ODE to the measured steady-state flux
- Reduce the parameter space significantly

(2) Normalize the ODE so that the metabolite concentrations are dimensionless, but the fluxes are dimensional.

(3) Elementary reactions can be used when detailed enzyme kinetics are unknown

Pros:
- Metabolite concentrations are not needed if not available
- Fluxes carry the correct dimension

(4) Parallel (ensemble) iteration

Pros:
- Most fundamental and flexible
- Can include regulation

Pro:
- High throughput screening of parameters

kinetics. These lumped enzyme kinetics include Michaelis–Menten kinetics, Hill equation and allosteric enzyme kinetics. The kinetic parameters are randomly assigned to satisfy both thermodynamic constraint and steady-state flux of individual enzymes. The detail formularization and sampling strategy can be found in Supplemental Material II.

A mathematical summary of the EM approach and comparison with the traditional approach is shown in Box 1. The novel contribution of the EM framework is to anchor the ODE to the measured steady-state flux, therefore reducing the parameter space significantly. By doing so, the sampling of kinetic parameters and further screening of parameter sets become possible.

2.5. Solving the system of algebraic equations

When elementary reaction kinetics is used, steady-state concentrations (normalized or absolute) can be obtained through the solution of the system of equations described by

$$\frac{dx_i}{dt}=SV=0 \tag{5}$$

$$\sum x_k=\text{constant} \tag{6}$$

Eq. (5) is required for each intracellular compound (metabolites and enzyme complexes) and Eq. (6) is required for each conserved species (e.g. enzymes and cofactors). Although the system of Eq. (5) already contains as many equations as unknowns, the equations of conserved metabolites are linearly dependent on each other. In order to make the system fully determined, it is important to add sets of Eq. (6) and remove one equation belonging to a member of each conserved species if desired. Such system of nonlinear equations (Eqs. (5) and (6)) is then re-written through reductive recasting in which algebraic manipulations are used (Savageau and Voit, 1987; Savageau, 1993). The goal of recasting is to reduce the original system to a standard canonical form in which each equation has only one positive term. The recast system can be readily solved using Newton's method and tends to have quadratic convergence rates. Such recast system is less sensitive to initial guesses than the original system due to the decomposition of nonlinear constraints into disjoint sets of convex constraint relationships (Irvine, 1988; Savageau and Voit, 1987; Savageau, 1993). As is often the case with Newton's Method, convergence is dependent on initial guess and therefore should be selected intelligently. To this effect, we simulate the system of ODEs (Matlab ode15s) for a short period of time (2000 time units) and use the state of the system at the final time point as the initial guess. In order to compare the benefits of using this numerical method instead of relying on ODE solution to obtain steady-state concentrations, we analyze 5000 models through both methods and record the time individual method takes.

2.6. Matlab module

A Matlab module has been developed to assist the construction of ensembles of dynamic models with any of the enzyme reaction

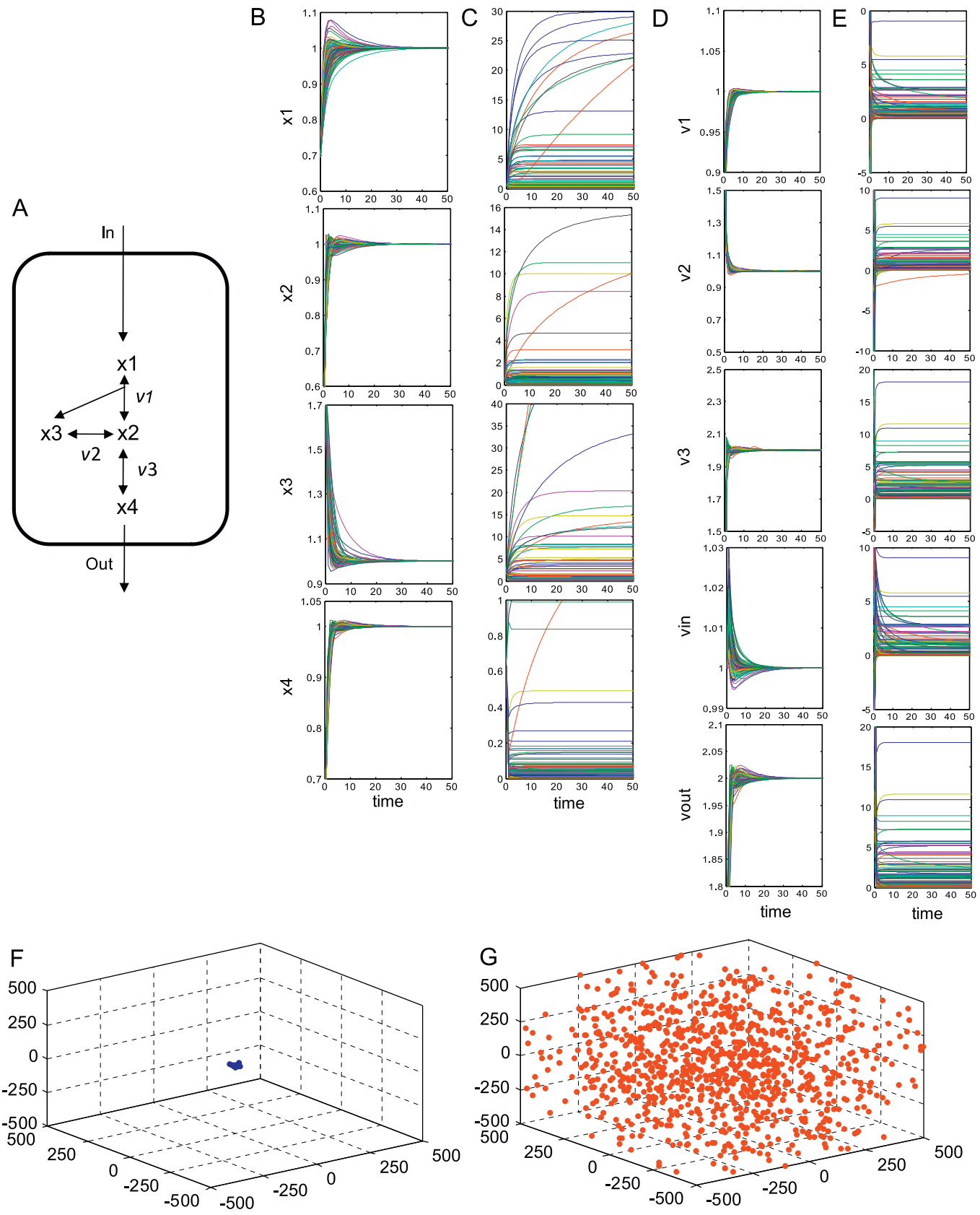

Fig. 1. (A) Toy network 1. (B) Metabolite profiles for 100 models that are constructed by anchoring to the predetermined steady-state flux. Each color represents one model with a different set of kinetic parameters. The initial condition is randomly selected and is altered from the predefined steady-state value. (C) Metabolite profiles for 100 control models that are constructed without the steady-state flux constraint. (D) Flux profiles for 100 models that are constructed with the steady-state flux constraint. (E) Metabolite profiles for 100 control models that are constructed without the steady-state flux constraint. The initial condition is randomly selected and is altered from the predefined steady-state value. (F) Kinetic parameter space spanned by the EM (1000 models are used for analysis). First three principal components are plotted here. (F) Kinetic parameter space spanned by the control models (1000 models are used for analysis) without the steady-state flux constraint. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
The EM algorithm is able to select meaningful models, where it is computationally infeasible for random sampling.

	Random sampling	EM
Small network		
Total models	10,000	10,000
# of models reach a steady state	1250	9940
# of models reach the desired steady state ($\pm$10% for random sampling)	4	9940
Large network		
Total models	10,000	10,000
# of models reach a steady state	0	9702
# of models reach the desired steady state	0	9702

kinetics mentioned above (http://www.seas.ucla.edu/~liaoj). The required inputs include the reaction network stoichiometry matrix, reference steady-state data (steady-state flux distribution, metabolite concentration if it is available), individual reaction's kinetic types (e.g. Elementary reaction kinetics, Michaelis–Menten kinetics) and enzyme regulations. The kinetic rate expressions of each enzyme can be chosen from different types of kinetics. The module will generate an ensemble of kinetic parameter sets. These parameter sets guarantee that the predefined steady-state flux can be achieved.

A separate Matlab module has been developed to recast the system of ODEs and solve it numerically by Newton's method when elementary reaction kinetics is used to describe individual enzymatic reactions (http://www.seas.ucla.edu/~liaoj).

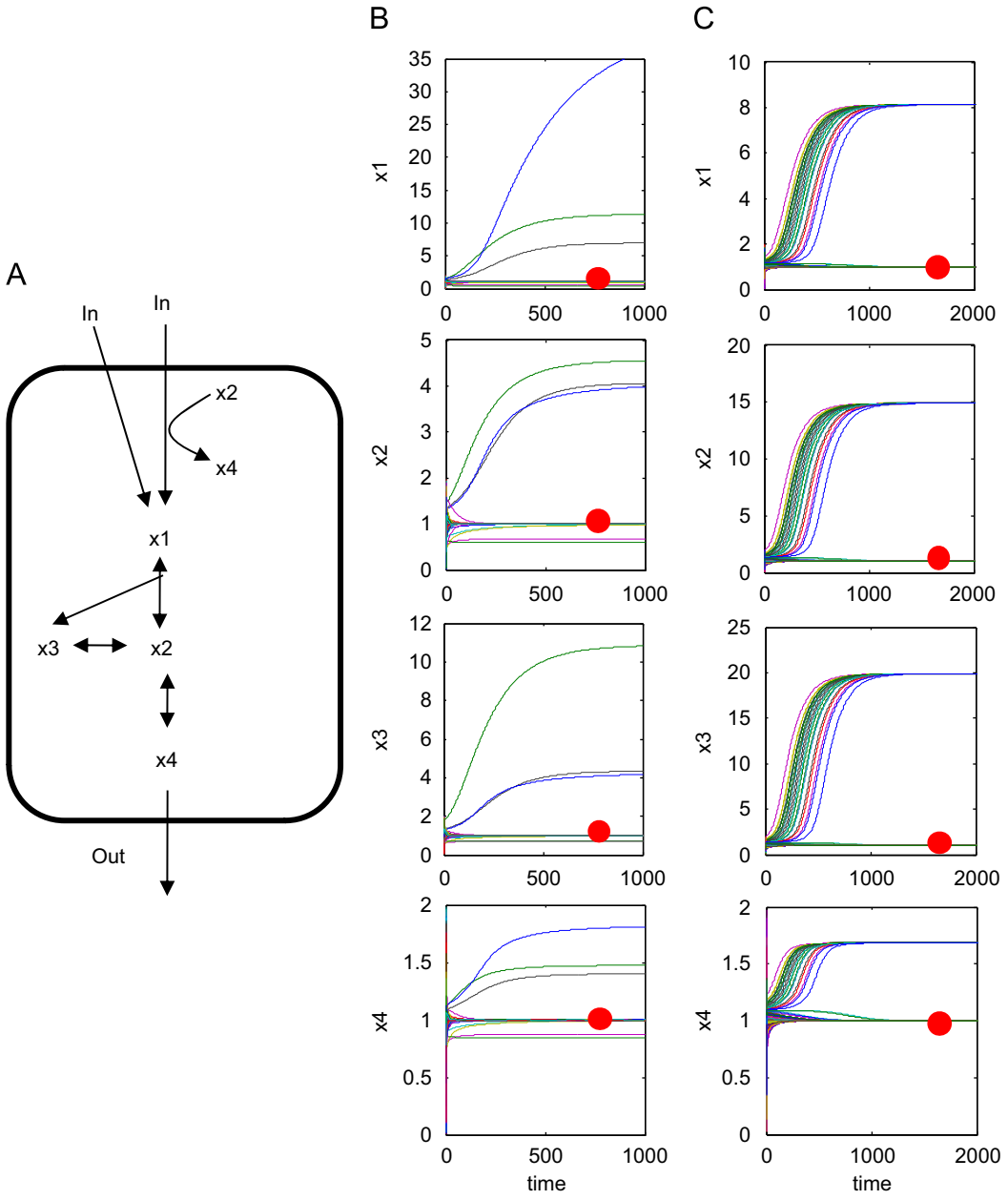

Fig. 2. (A) Toy network 2. (B) A small fraction of the ensemble (less than 10%) reaches a different steady state when the initial condition is randomly selected. The red dots represent the predefined steady state. (C) A specific model that arrives at a different steady state is isolated and various randomly chosen initial conditions are used. For some initial conditions, the model reaches the predetermined steady state as labeled with the red dots, while for other initial conditions, the model reaches the other steady state. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. Anchoring to steady-state flux reduces parameter space

The EM strategy allows models with different parameter sets (referred to as different models hereafter) to reach the same steady-state flux. To demonstrate this point, we construct toy network 1 (Fig. 1A), which is composed of four metabolites and five flux streams. Elementary reaction kinetics is used to construct an ensemble of models using the Matlab module developed in this work (see Section 2). By starting from an initial condition that deviates from the assigned steady-state values, all models in the ensemble constructed with the EM strategy converge to the assigned steady state in both metabolites and fluxes, but with different dynamics (Fig.1B, D). Although the steady states of these models are all stable in this example, EM does not guarantee stability. Models which have unstable steady states will be screened out immediately and have no impact further. This point will be demonstrated shortly.

To demonstrate the importance of the steady-state flux constraints, we construct a control ensemble whose kinetic parameters are randomly sampled without the steady-state flux constraint. As expected, all models in the control ensemble diverge to different steady states or no steady state at all, as shown in Fig. 1C, E.

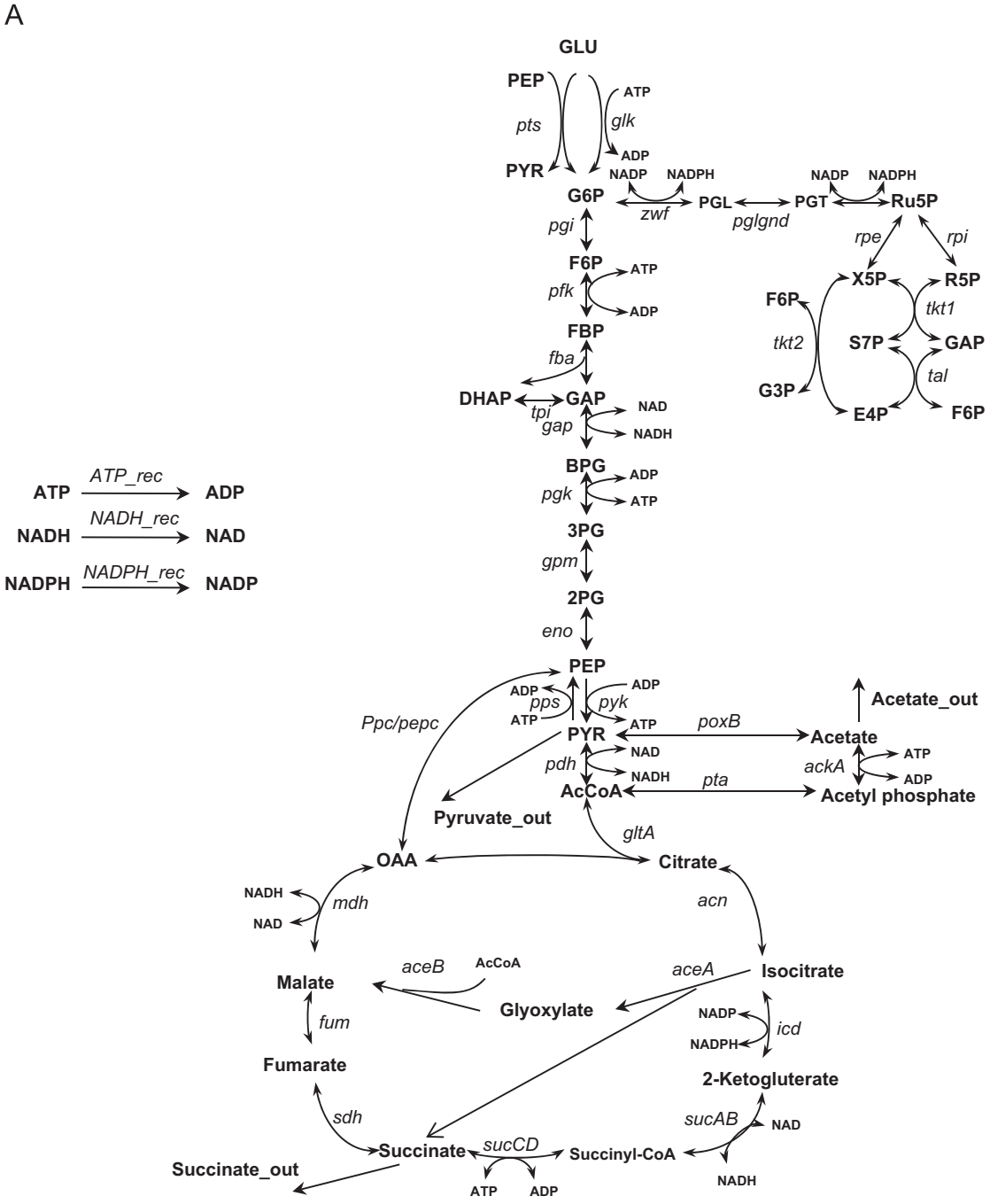

Fig. 3. (A) The *E. coli* succinate production model. PTS, glycolysis, pentose–phosphate pathway and TCA cycle are included in the model. (B) External fluxes (glucose uptake, acetate, succinate and pyruvate) of 100 models in the initial ensemble are plotted. A random initial condition is used here. After screening out a few models that possess multiple steady states, the rest of the models all reach the predetermined steady state. (C) External flux responses caused by *sdh* knockout. Each model behaves differently. (D) After comparing with experimental data, the models that behave qualitatively correct are retained for the next step screening.

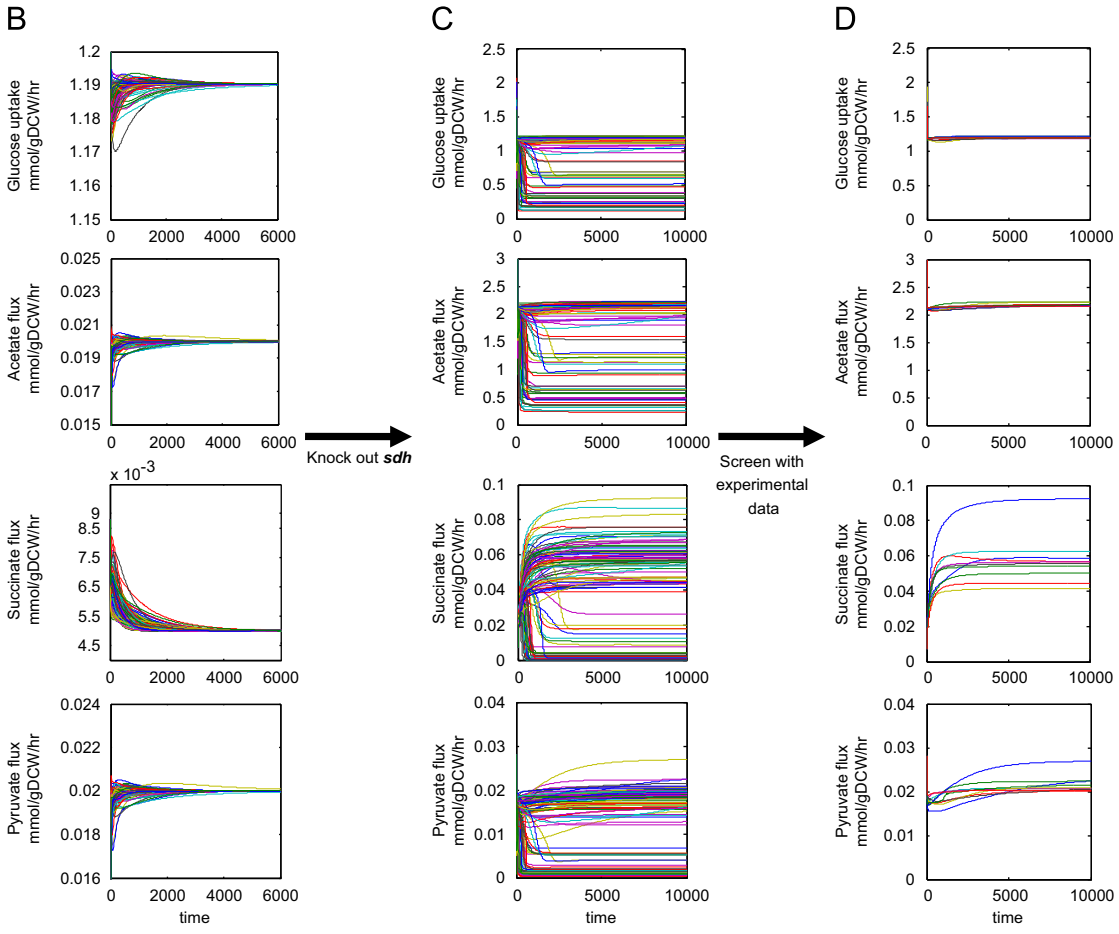

Fig. 3. (*Continued*)

The kinetic parameter spaces spanned by sampling with and without the steady-state constraint are compared using principal component analysis (PCA) in order to visualize high dimensional data. To do so, PCA is first performed for the kinetic parameter sets obtained using the EM approach. For illustration, the first three principal components are plotted (Fig. 1F). Then, the kinetic parameters that are sampled without the steady-state constraints are plotted using the same coordinate system (Fig. 1G). As shown in Fig. 1F, G, with the constraints of steady-state fluxes, the allowable kinetic parameter space is dramatically reduced. This result suggests that flux anchoring is highly beneficial for exploring the kinetic space using random sampling. This result is reasonable, since EM utilizes realistic data to constrain the system. However, the degree to which the parameter space is reduced is unexpected.

Additionally, we have shown that without anchoring, random sampling of parameters will yield models of which 87.5% do not even reach a steady state, and only 0.04% of the models eventually converge to a point in the vicinity (10%) of the desired steady state fluxes. These numbers become even smaller as the network becomes larger (Table 2). Thus by anchoring the model to the desired steady state, the sampling efficiency increased by several (4–8) orders of magnitude.

3.2. Multiple steady states

When the network topology becomes more complicated, as in toy network 2 (Fig. 2A), we observe that some sets of parameters in the ensemble exhibit multiple steady-state behavior (Fig. 2B). When the initial condition is altered from the desired steady state, a fraction of the models do not return to the desired steady state, but reach different steady states. To further illustrate this behavior, a specific model that reaches a different steady state is identified. As shown in Fig. 2C, the model with this specific parameter set converges to two different steady states depending on the initial condition, with one of the steady states being the predetermined steady state. This indicates that some of the models constructed with EM approach may possess an alternative steady state. Although multiple steady states in metabolic systems have been reported (Europa et al., 2000; Lu et al., 2007), the models that possess multiple steady states or unstable steady states are eliminated from the ensemble. On the other hand, multiple steady state behavior can be further studied using EM, if the problem at hand is to investigate this phenomenon specifically.

3.3. Screening of EM with enzyme tuning data: succinate production example

The size of the ensemble can be reduced by screening with experimental data, such as flux changes upon enzyme level tuning. Succinate production in *E. coli* is used as an example. The scope of the succinate production model includes the phosphotransferase system (PTS), glycolysis, pentose–phosphate pathway and the tricarboxylic acid cycle (TCA cycle). The detail structure is shown in Fig. 3A. The elementary reaction rate laws are used here. The steady-state flux distribution is calculated based on the measured external fluxes (Lin et al., 2005a) and the pseudo-steady state assumption (Table 3). The EM approach described above was used

Table 3
Reference steady-state flux and enzymes standard Gibbs free energies for succinate model.

Enzyme	Overall reaction	Net flux at reference state (mmol/gDCW/h)	Standard Gibbs free energy (kcal/mol)
pgi	G6P↔F6P	1.09	−2.5
pfk	F6P+ATP↔FBP+ADP	1.16	−4.5
fba	FBP↔DHAP+GAP	1.16	1.1
tpi	DHAP↔GAP	1.16	0.2
gap	GAP+NAD↔BPG+NADH	2.35	4.2
pgk	BPG+ADP↔3PG+ATP	2.35	4.7
gpm	3PG↔2PG	2.35	−2.2
eno	2PG↔PEP	2.35	−0.2
pyk	PEP+ADP↔PYR+ATP	1.4	−8.4
pps	PYR+ATP↔PEP+ADP	0.15	−3.6
pdh	PYR+NAD↔ACCOA+NADH	1.81	−5.1
pfl	PYR↔ACCOA	0.001	−2.5
glt	ACCOA+OAA↔CIT	0.17	−7.4
acn	CIT↔ICIT	0.17	0
icd	ICIT+NADP↔AKG+NADPH	0.13	−3.52
sucAB	AKG+NAD↔SUCCOA+NADH	0.13	−5.12
sucCD	SUCCOA+ADP↔SUC+ATP	0.13	0.8
sdh	SUC↔FUM	0.16	0.7
fum	FUM↔MAL	0.12	−1.3
mdh	MAL+NAD↔OAA+NADH	0.16	4.8
aceA	ICIT↔GLYX+SUC	0.04	−2.2
aceB	GLYX↔MAL	0.04	−10.3
ppc	PEP↔OAA	0.01	−11.74
pyc	PYR↔OAA	0.001	−12
poxB	PYR↔AC	0.5	−46.2
pta	ACCOA↔ACTP	1.65	−3.9
ackA	ACTP +ADP↔AC+ATP	1.65	−4.7
zwf	G6P+NADP↔6PGL+NADPH	0.1	−0.9
pgl	6PGL↔6PGC	0.1	−13.3
gnd	6PGC+NADP↔Ru5P+NADPH	0.1	−0.8
rpe	Ru5P↔X5P	0.07	0
rpi	Ru5P↔R5P	0.03	−0.7
tkt(1)	X5P+R5P↔S7P+G3P	0.03	0.9
tal	S7P+G3P↔E4P+F6P	0.03	−0.6
tkt(2)	X5P+E4P↔F6P+G3P	0.03	0.9
ATP_rec	ATP↔ADP	4.12	−0.1
NADH_rec	NADH↔NAD	4.45	−0.1
NADPH_rec	NADPH←NADP	0.33	−0.1
EI	PEP↔PYR+P1	1.09	−3
HPr	P1↔P2	1.09	−0.1
EIIA	P2↔P3	1.09	−0.1
EIICB	Glu+P3↔G6P	1.09	−9
glk	Glu+ATP↔G6P+ADP	0.1	−4.5

to construct 5000 models in the initial ensemble. As shown in Fig. 3B, the models in the initial ensemble reach the same steady states, after a small fraction of models that possess multiple steady states are filtered out. As expected, enzyme perturbations shift the steady flux to different extents (Fig. 3C), which differentiates each model in the ensemble.

A series of genetic manipulations have been reported in the literature to construct a succinate over-producing strain (Lin et al., 2005a, 2005b; Sanchez et al., 2005). These manipulations include 1) knockout of succinate dehydrogenase (*sdh*), 2) knockout of pyruvate oxidase (*poxB*), 3) knockout of acetate kinase (*ackA*) and phosphotransacetylase (*pta*), and 4) overexpression of the glyoxylate shunt enzymes isocitrate lyase (*aceA*) and malate synthase (*aceB*) by knocking out the *aceBAK* operon repressor (*iclR*). After the first manipulation, the fluxes of glucose uptake, acetate production and pyruvate accumulation remain unchanged while succinate flux becomes detectable. Both the second and the third steps result in dramatic drops in acetate production and glucose uptake. Meanwhile, pyruvate begins to accumulate and the succinate production rate further increases. After the fourth manipulation, succinate production flux increases significantly along with an increase in the glucose uptake rate when compared to the preceding strain. Pyruvate accumulation decreased and acetate production exhibits no change.

In order to screen the ensemble with these experimental data, we perturb each model in the ensemble according to the enzyme knockouts and overexpressions carried out experimentally (Lin et al., 2005a, 2005b; Sanchez et al., 2005). We then retain models whose predictions are consistent with the experimental observation (Fig. 3D). The models whose responses are in qualitative agreement with that of the experimental data are retained for the next step. We here consider any flux changes *in silico* that are less than 5% as non-change, otherwise increase or decrease, respectively. We expect that the first application of EM is qualitative prediction of enzyme targets to guide the experimentalist in metabolic engineering. Thus, the level of accuracy should be adjusted according to the need of the experimentalist. In our experience, only qualitative predictions of increase, decrease, or no change in the desired flux would be useful to the experimentalist. As such, the results are not sensitive to the overexpression level (whether 5-, 10- or 15-fold). As more data are used in the construction of EM, the accuracy of prediction will increase and more quantitative agreement will be desired. To avoid numerical singularity, a 99% decrease in enzyme level represents a genetic knockout . A 10-fold increase in enzyme level is used to reflect the experimental overexpression. Other levels of increase can be used if experimental evidence is available. As shown in Fig. 4A, 24 out of 5000 models exhibit correct changes in glucose uptake rate, acetate production rate, pyruvate accumulation rate and succinate production rate after four sequential genetic perturbations.

Meanwhile, the control ensemble of 5000 models is constructed with their kinetic parameters randomly sampled without any constraints. These unconstrained models are then screened against the same experimental data. As shown in Fig. 4B, none of the models can even describe the system's response caused by the first step of manipulation: knockout of succinate dehydrogenase (*sdh*). Fig. 4C, D show the contrast of kinetic space spanned by the ensemble with and without the steady-state constraint. Because of the reduction in the allowable kinetic space, EM enables model screening using available data. Such a process is impossible when the steady-state anchor is not in place.

3.4. Reduction in computation time through the solution of algebraic equations

Instead of using the numerical integration to compute new steady states after perturbations, we solve the system of algebraic equations directly to reduce computation time. To be noticed that such recasting technique is only applicable when elementary enzyme kinetics are used, and it is limited to steady-state solutions only. The solution approach is described in Section 2 and a Matlab module is developed to automate this process (http://www.seas.ucla.edu/~liaoj). For a direct comparison, we analyze 5000 models for a series of perturbations using both methods. In this example, solving the system of algebraic equations offers a 63% reduction in the computation time required. In addition, we are expecting more time saving when the network becomes larger and more complicated as shown in Fig. 4E.

3.5. The retained models can predict the outcome of additional perturbations

In addition to describing and understanding the systematic behavior of the succinate producing pathway, it is desirable to predict the results of further experiments with the assistance of the retained models after screening. Lin et al. (2005b) further engineered the strain by knocking out PTS and overexpressing

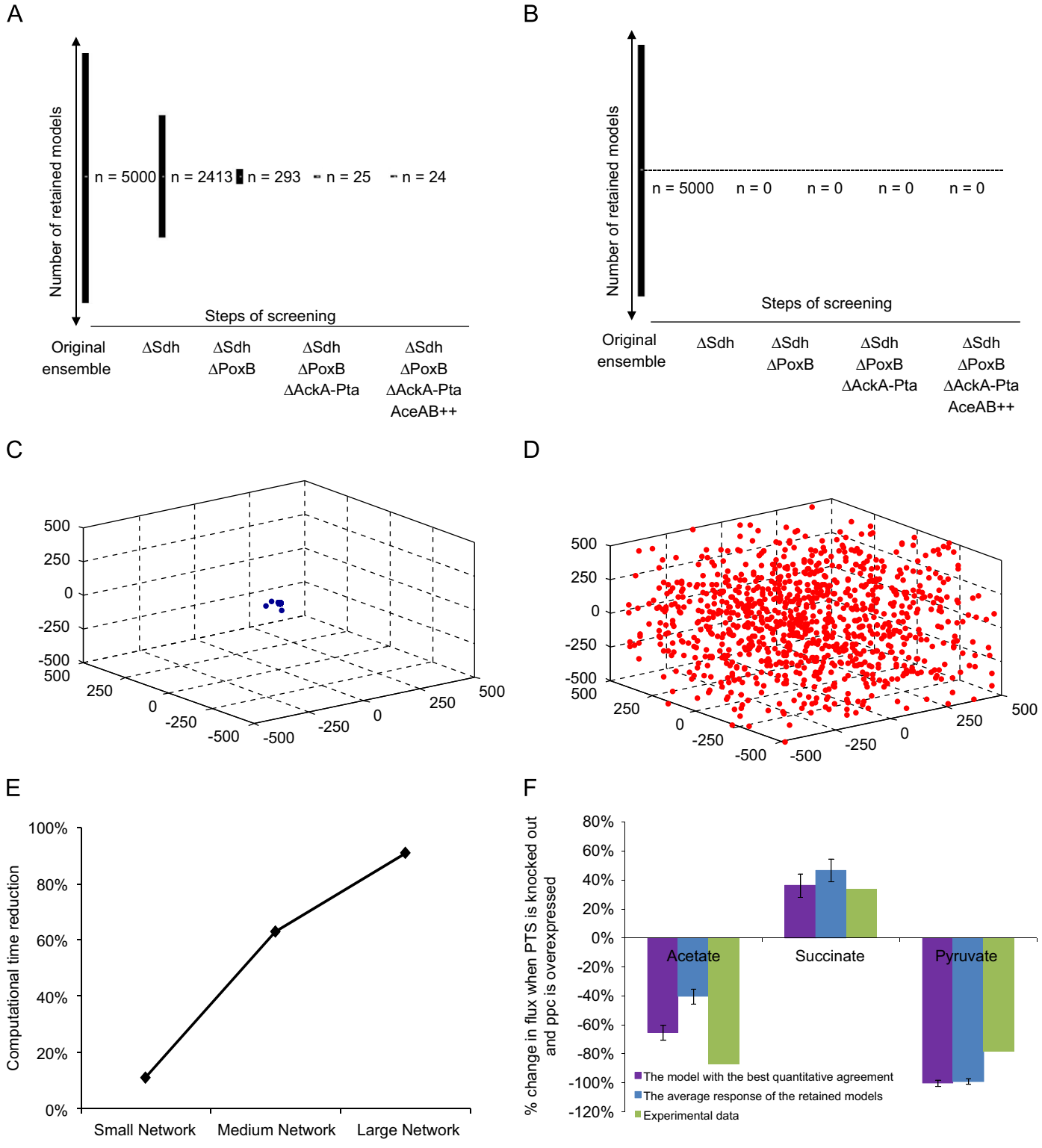

Fig. 4. (A) When the models are anchored to the predetermined steady state, the ensemble converges to a subset of 24 models after four screening steps. These 24 models are able to describe all four experimentally observed phenotypes. (B) Without the steady-state constraint, none of the models even survive the first step of screening. (C) Kinetic parameter space spanned by the steady-state flux constrained models (1000 models are used for analysis). The first three principal components are plotted here. (D) Kinetic parameter space spanned by the control models (1000 models are used for analysis) without the steady-state flux constraint. (E). Computation time reduction by using recasting method. The computational time reduction is defined as $1-(time_{recast}/time_{ODE})$. The small network is an artificial 4 metabolites × 5reactions network; the medium network is the succinate example network (35 metabolite × 42 reactions); the large network is an artificial 68 metabolites × 86 reactions network. (F). The 24 models retained can all predict system response upon further experimental perturbations. Comparison between the EM prediction (the average response of 24 models and the response of the model with the best quantitative agreement) and experimental data for the percent change in the succinate production rate, acetate accumulation rate and pyruvate accumulation rate when PTS is inactivated and *ppc* is overexpressed. The model with the best quantitative agreement with the screening experimental data has be identified by ranking retained 24 models' summed squared errors. The error bars in the model data represent the standard deviation.

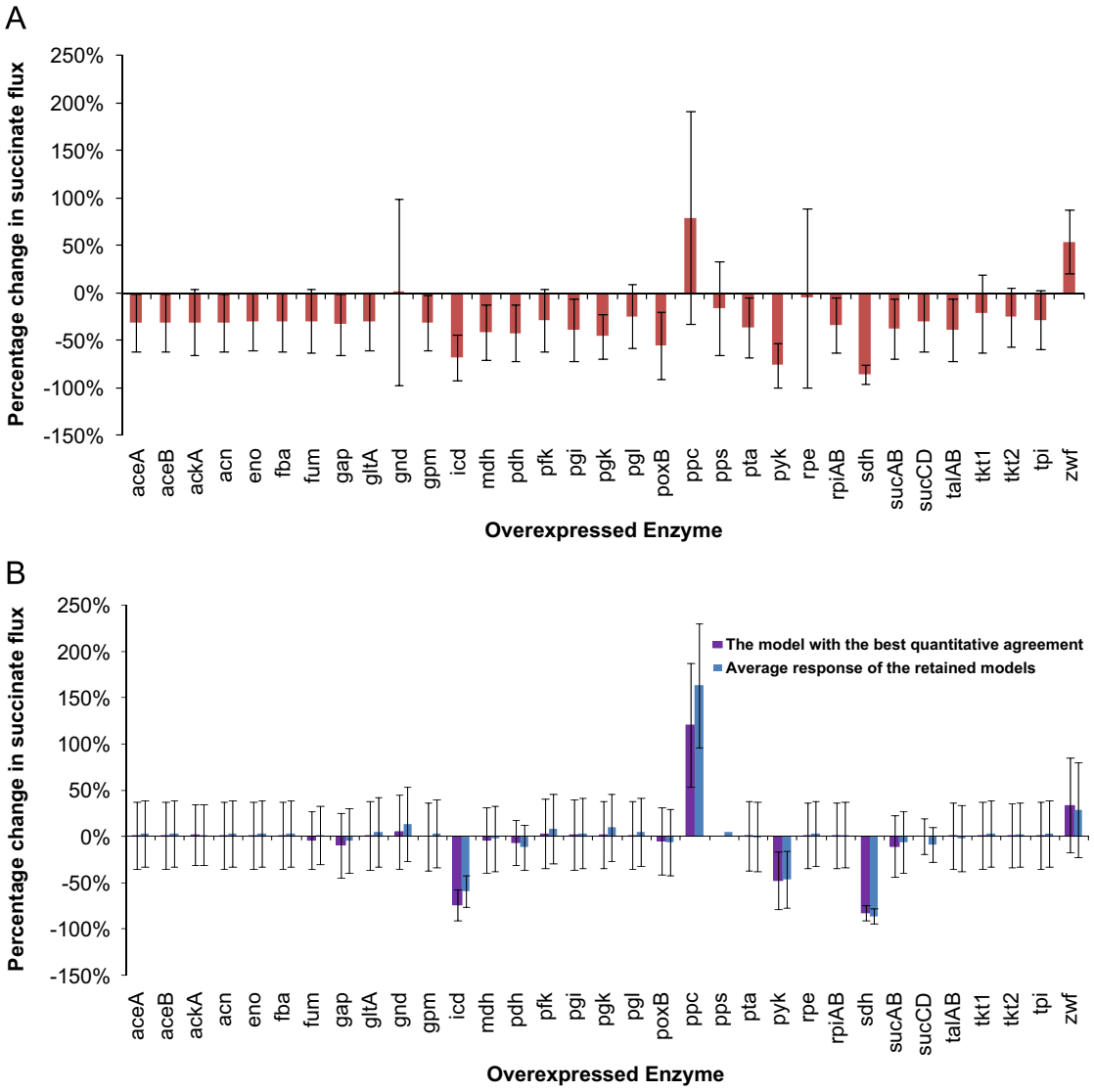

Fig. 5. The succinate production percentage change upon individual enzyme overexpressions (10-fold change) and the error bars are standard deviations. (A) Base on initial ensemble of 5000 models and (B) base on retained 24 models. The purple bars are the responses of the model with the best quantitative agreement with the screening data. The model with the best quantitative agreement with the screening experimental data has been identified by ranking retained 24 models' summed squared errors. The blue bars are the average responses of the retained 24 models.

phosphoenolpyruvate carboxylase (*ppc*). These manipulations were intended to increase phosphoenolpyruvate (PEP) pool for succinate production (Chao and Liao, 1993; Chao et al., 1993; Lin et al., 2005b; Millard et al., 1996; Patnaik et al., 1992; Rizk and Liao, 2009). The 24 models all predict that these additional modifications could increase the succinate flux while decreasing pyruvate and acetate fluxes. Indeed, the reported experimental data (Lin et al., 2005b) demonstrated a significantly increased succinate production rate, as shown in Fig. 4F.

3.6. The importance of model screening with enzyme tuning data

As a part of the initial ensemble of 5000 models, the retained 24 models are based on the same stoichiometry. After screening with the enzyme tuning data, the kinetics of the retained 24 models are refined and distinguish them from other 4976 models of the initial ensemble. To demonstrate the significance of enzyme tuning data utilization and importance of kinetics refining, we survey and compare all possible overexpression targets after four existing perturbations using both the initial 5000 models and the retained 24 models. Individual enzymes are overexpressed 10-fold respectively *in silico*, and the effects of individual enzyme overexpression on succinate production predicted by initial ensemble and the retained subset are shown in Fig. 5. According to the survey based on overall 5000 models as shown in Fig. 5A, the average effect of *ppc* overexpression is to benefit production. However the large error bar that lies across *x*-axis diminishes the reliability of candidacy. Instead, *zwf* stands out as a possible target, but it has never been shown to benefit succinate production. On the contrary, as shown in Fig. 5B, the retained 24 models altogether predict that *ppc* is the only overexpression target. This overexpression target has been confirmed by the experiment, therefore the false positive prediction rate is near zero. Such comparison shows that the screening of the ensemble with experimental data is necessary to get the correct prediction. This simply illustrates the mutualism between mathematical models and experiment data: only when provided and trimmed with enough experimental data, the model predictions can be reliable.

3.7. Expanding the EM strategy to use lumped enzyme kinetics

The concept of EM is to anchor the ODE models to the same steady-state flux with kinetic parameters sampled under thermodynamic constraints. Instead of solely using elementary reaction kinetics, we further expand the concept of EM to construct models using lumped kinetic rate laws, such as the Michaelis–Menten kinetics, Hill equation and allosteric enzyme kinetics. They can be implemented in the same fashion of elementary reaction kinetics such that the knowledge of absolute steady-state metabolite concentrations is optional, detailed derivations can be found in Supplemental Material II. Although these lumped kinetic expressions are empirical in nature, they contain less kinetic parameters. This expansion makes the EM approach more flexible and applicable to situations when lumped kinetic expressions are available.

3.8. Screening by dynamic data: central metabolism model

Metabolite dynamic profiles, although difficult to obtain (Buchholz et al., 2002), can also be used as the data for screening models in the EM approach. The dynamic model of *E. coli* central carbon metabolism is used as an example. Fig. 6A illustrates the modeling structure that is similar to the previous work (Chassagnole et al., 2002) including the PTS, glycolysis, pentose–phosphate pathway and storage metabolites. The dynamic model was constructed using different types of enzyme kinetics (Table 4). In general, enzymes with known mechanisms are assigned lumped kinetic expressions. Others are modeled using elementary reactions. The steady-state fluxes and metabolite concentrations of this system were determined previously (Chassagnole et al., 2002), and are shown in Table 5. The cofactors (NAD, NADH, NADPH, NADP, AMP, ADP, ATP) in this case are modeled as analytical functions of time as shown in Table 6. These are derived from the previous study (Chassagnole et al., 2002), and are not necessarily balanced throughout the time course after the perturbation is introduced. An ensemble of 10,000 models is constructed by anchoring to the predetermined steady-state data, as shown in Fig. 6B. Meanwhile, the control ensemble of 10,000 models is constructed with their kinetic parameters randomly sampled without any constraints. We then perturb individual models and compare the screening results of these two ensembles.

After two initial ensembles of models are constructed, we then perturb these models according to the experimental designs and compare the model predictions with the experimental data reported in the literature (Chassagnole et al., 2002). For the ensemble with the steady-state constraint, a glucose pulse is introduced to each model. Fig. 6C illustrates the metabolite responses. The root mean square of deviation between model response and experimental measurement is calculated as the fitting error for each model. The model that has the smallest fitting error of 2.13 is retained to describe the system response as shown in Fig. 6D. Fig. 7A shows the number of retained models as a function of the fitting errors. In particular, 84 out of 10,000 models behave very similarly to the experimental observation with fitting errors smaller than 3. In the case of the control ensemble without the steady-state constraint, none of the models has a fitting error smaller than 10 which is the range that could produce a reasonable match with the measured dynamic response (Fig. 7B). This indicates that the screening of kinetic parameters without anchoring to the steady state is impossible due to the large parameter space. The kinetic parameter spaces spanned by the two ensembles with and without flux anchoring are shown in Fig. 7C and D, respectively.

The retained model fits the observed trend for most metabolite trajectories. After injecting glucose, both model and experiment

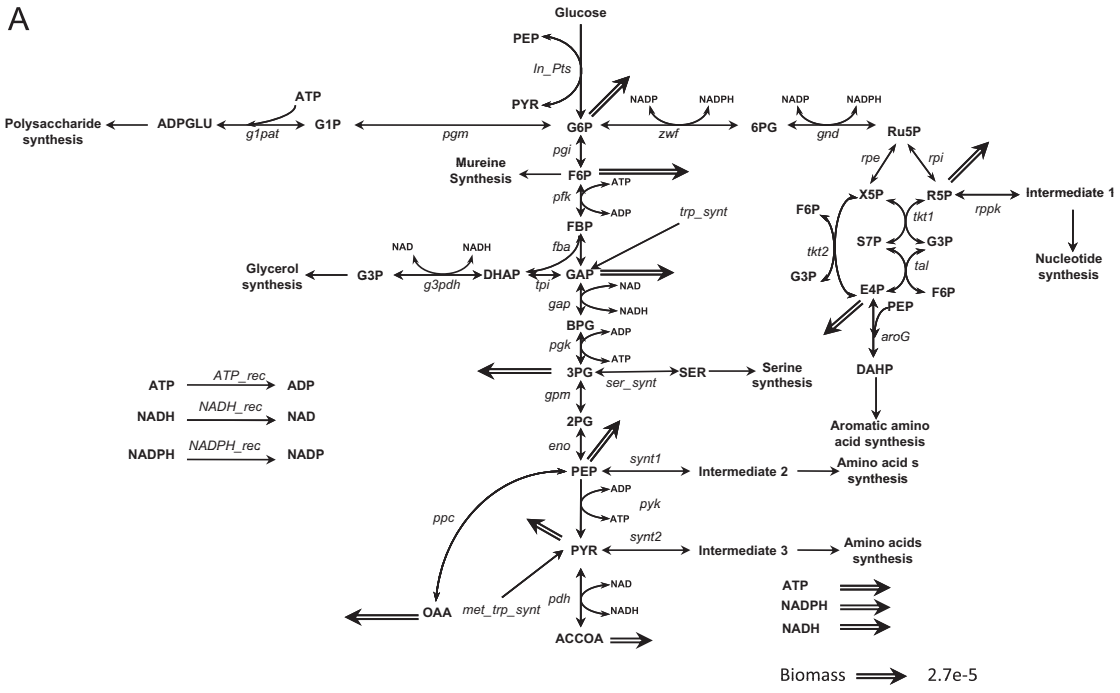

Fig. 6. (A) Structure of the *E. coli* central metabolism model. PTS, glycolysis, pentose–phosphate pathway and storage material are included in the model. The double arrows represent the required materials for biomass generation. (B) Metabolite profiles of the initial ensemble of 100 models. A random initial condition is used here. After screening out a few models that possess multiple steady states, the rest of the models all reach the predetermined steady state. (C) Metabolite dynamic profile caused by the glucose pulse. Each model behaves differently. The predetermined steady-state metabolite concentrations are used as the initial condition. (D) After comparing with the experimental data, the model with the smallest fitting error (2.13) is retained to describe system's response. The error bar is the standard deviation of 84 models that have fitting errors smaller than 3. The red dots represent the experimental measurements.

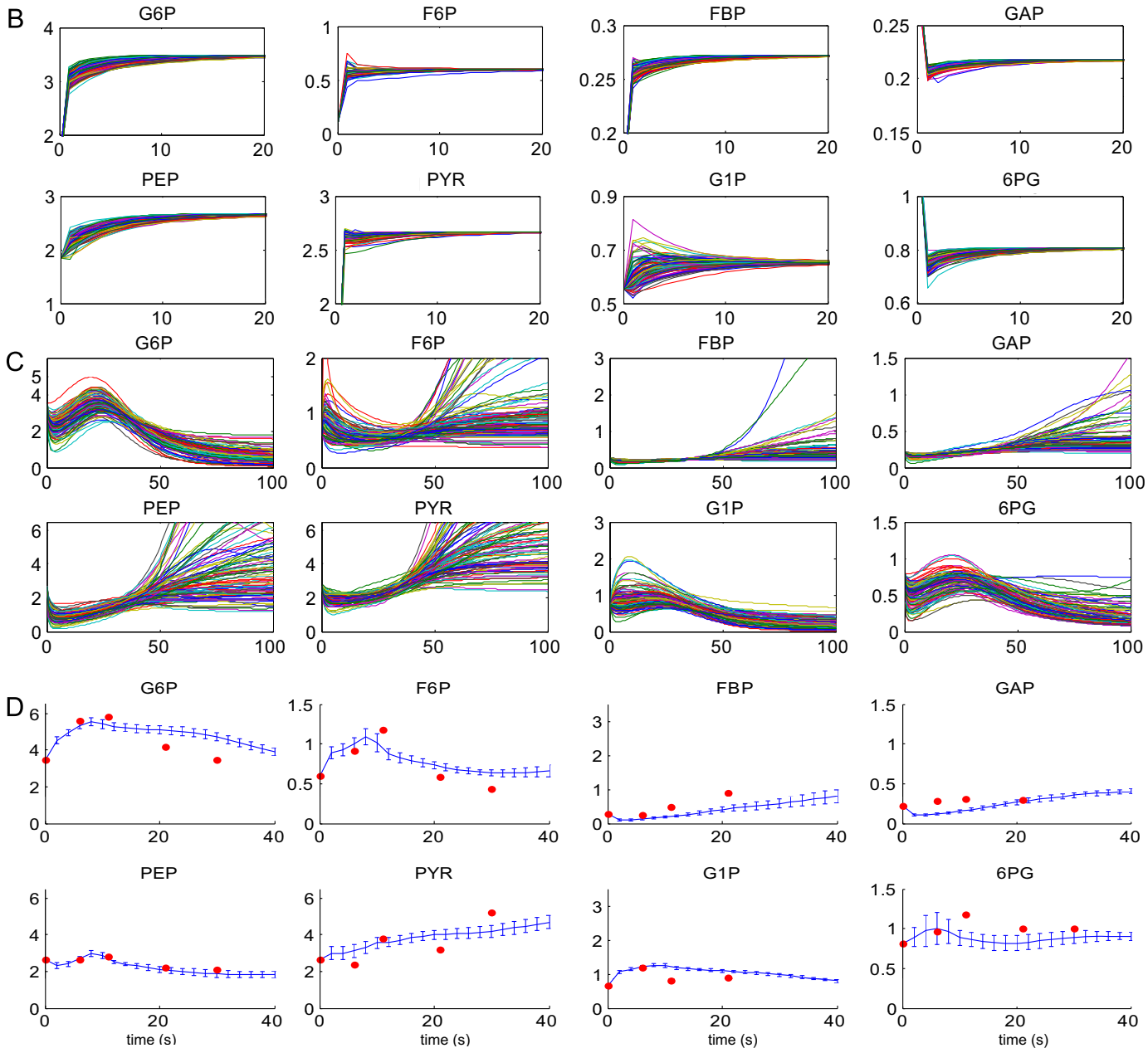

Fig. 6. (*Continued*)

measurements show a sharp increase in the glucose-6-phosphate (G6P) level (Chassagnole et al., 2002). Simultaneously, a sharp decline in the PEP concentration can be observed. This rapid response is expected, as the sudden abundance of glucose after stimulation causes a rapid decline in PEP, which is used in the conversion of glucose to G6P through PTS. The drop in G6P after 10 s suggests that PEP becomes limited for assisting glucose uptake through PTS, as PEP concentration is in a continuous decline after 10 s.

There are slight deviations between model responses and experimental observations, in particular, the model predicted profiles of fructose-1,6-bisphosphate (FBP) and glyceraldehdye-3-phosphate (GAP) do not follow the experimental data perfectly. These two metabolites are directly downstream of the heavily regulated enzyme 6-phosphofructokinase (Pfk) (Johannes and Hess, 1973), the complexity of its regulatory mechanisms may not be fully captured by the model structure used here.

4. Discussion

In this work, we have demonstrated that by anchoring the models to the same steady-state flux significantly reduces the kinetic parameter space. With this reduction in kinetic space, the screening of dynamic models with available experiment data becomes possible. Conversely, the screening procedure is fruitless without anchoring to the steady-state flux.

Here we use the steady-state flux instead of the steady-state metabolite concentrations as the anchor for various reasons: (1) Given a collection of steady-state metabolite concentrations without the fluxes, the unique steady-state fluxes are not guaranteed when kinetic parameters are randomly sampled. (2) The employment of steady-state fluxes sets the time scale for the dynamic models once the approximated metabolite pool sizes are known, whereas steady-state metabolite concentrations do not carry time

Table 4
Overall reactions, kinetic types, regulations and standard Gibbs free energies of enzymatic reactions involved in *E. coli* central carbon metabolic model.

Enzyme	Overall reaction	Kinetic types	Activation	Inhibition	Standard Gibbs free energy (kcal/mol)
pts	GLU+PEP↔G6P+PYR	c		G6P	−13.1
pgi	G6P↔F6P	c		6PG	−2.5
pfk	F6P+ATP↔FBP+ADP	c	AMP, ADP	PEP	−4.5
fba	FBP↔DHAP+GAP	e			1.1
tpi	DHAP↔GAP	a			0.2
gap	GAP+NAD↔BPG+NADH	a			4.2
pgk	BPG+ADP↔3PG+ATP	a			4.7
gpm	3PG↔2PG	a			−2.2
eno	2PG↔PEP	a			−0.2
pyk	PEP+ADP→PYR+ATP	c	FBP,AMP	ATP	−8.4
pdh	PYR+NAD↔ACCOA+NADH	b			−5.1
ppc	PEP↔OAA	c	FBP		−11.74
g3pdh	DHAP+NADH↔G3P+NAD	a			−2
pgm	G6P↔G1P	a			2.2
g1pat	G1P+ATP↔ADPGLU	c	FBP		2.8
ser_synt	3PG↔SER	a			−2
zwf	G6P+NADP↔6PG+NADPH	c		NADPH	−0.9
gnd	6PG+NADP↔Ru5P+NADPH	c		NADPH, ATP	−0.8
rpe	Ru5P↔X5P	a			0
rpi	Ru5P↔R5P	a			−0.7
tkt1	X5P+R5P↔S7P+G3P	a			0.9
tkt2	X5P+E4P↔F6P+G3P	a			0.9
tal	S7P+G3P↔E4P+F6P	a			−0.6
aroG	PEP+E4P↔DAHP	e			−17.9
rppk	R5P↔Intermediate 1	a			−3.5
synt1	PEP↔Intermediate 2	a			−3.5
synt2	PYR↔Intermediate 3	a			−3.5
trp_synt	in→GAP	d			−3.5
met_trp_synt	in→PYR	d			−3.5

a: reversible Michaelis–Menten, b: Hill equation, c: allosteric enzyme kinetics, d: constant flux, e: elementary reaction kinetics.

Table 5
The steady-state flux distribution and metabolite concentrations of central carbon metabolism.

Enzyme reaction	Flux (mmol/gDCW/s)	Metabolite	Metabolite concentration (mM)
pts	0.1	G6P	3.48
pgi	0.37	F6P	0.60
pfk	0.73	FBP	0.272
fba	0.73	DHAP	0.167
tpi	0.71	GAP	0.218
gap	1.62	BPG	0.008
pgk	1.62	3PG	2.13
gpm	1.53	2PG	0.399
eno	1.53	PEP	2.67
pyk	0.18	PYR	2.67
pdh	0.93	Ser	1
ppc	0.25	G3P	1
g3pdh	0.02	G1P	0.653
pgm	0.01	ADPGlU	1
g1pat	0.01	6PG	0.808
ser_synt	0.09	Ru5P	0.111
zwf	0.62	X5P	0.138
gnd	0.62	R5P	0.398
rpe	0.37	S7P	0.276
rpi	0.25	E4P	0.098
tkt1	0.20	DAHP	1
tkt2	0.17	RPPK	1
tal	0.20	ACCOA	1
aroG	0.03	OAA	1
rppk	0.05	Synt1	1
synt1	0.07	Synt2	1
synt2	0.27	ATP	4.27
trp_synt	0.01	ADP	0.595
met_trp_synt	0.02	AMP	0.955
		NADH	0.1
		NAD	1.47
		NADPH	0.062
		NADP	0.195

information. (3) Flux data are easily measureable or estimated, but metabolite concentrations are not.

Furthermore, we extend the EM approach to lumped enzyme kinetic rate expressions. In some cases, enzyme rate expressions are available but some of the kinetic parameters (such as v_{max}) are not. This extension allows EM to take advantage of well characterized lumped kinetic rate expressions while using elementary reaction kinetics for others. Known parameters can be fixed while others are sampled in EM (Dean et al., 2009). We also demonstrate the use of metabolite profiles in screening, in addition to the flux data used before. These results make the EM approach more versatile.

Existing dynamic modeling methods (Chassagnole et al., 2002; Usuda et al., 2010; Visser and Heijnen, 2003) heavily rely on the knowledge of enzyme kinetic form and parameters. Detailed kinetic parameters are typically obtained through *in vitro* enzyme assay experiments, as a result they are rare and difficult to measure. Additionally, the *in vitro* measurements do not necessarily reflect *in vivo* performance hence post-adjustment of these parameters and estimating other unknown parameters according to experimentally measured metabolite concentration profiles are required. Such process needs extra experiment data (metabolite profiles) which are difficult to measure and not commonly done, while the EM approach takes advantageous of flux measurements (production formation rates, not necessarily C13 labeled measurements) that are routinely generated through metabolic engineering efforts. Even with the metabolite profile data, the parameter estimation is very inefficient because of the large dimension and wide range of parameters. Most importantly, the fitting of metabolite profiles does not guarantee prediction of fluxes, which is the main goal of metabolic engineering.

In contrast to related works (Fell, 1992; Kahn and Westerhoff, 1991; Liao and Delgado, 1993; Link and Weuster-Botz, 2007; Schellenberger and Palsson, 2009; Wang et al., 2004), the EM approach generates nonlinear models that use parameters closely

Table 6
Analytical functions for cofactors.

$$C_{atp} = 4.27 - 4.163\frac{t}{0.35 + 1.25t + 0.05t^2}$$

$$C_{adp} = 0.582 + 1.73(2.731^{-0.15t})(0.12t + 0.000214t^3)$$

$$C_{amp} = 0.123 + 7.25\frac{t}{7.25 + 1.47t + 0.17t^2} + 1.073\frac{t}{1.29 + 8.05t}$$

$$C_{nadph} = 0.062 + 0.332(2.718^{-0.464t})(0.0166t^{1.58} + 0.000166t^{4.73} + 1.13 \times 10^{-10}t^{7.89} + 1.36 \times 10^{-13}t^{11} + 1.23 \times 10^{-16}t^{14.2})$$

$$C_{nadp} = 0.18 + 2 \times 10^{-6}t^3 + \frac{t^2}{5000} - \frac{t}{100}$$

$$C_{nadh} = 0.0934 + 0.0011(2.371^{-0.123t})(0.844t + 0.104t^3)$$

$$C_{nad} = 1.314 + 1.314(2.73^{(-0.0435t - 0.342)}) - \frac{(t + 7.871) \times 2.73^{(-0.0218t - 0.171)}}{8.481 + t}$$

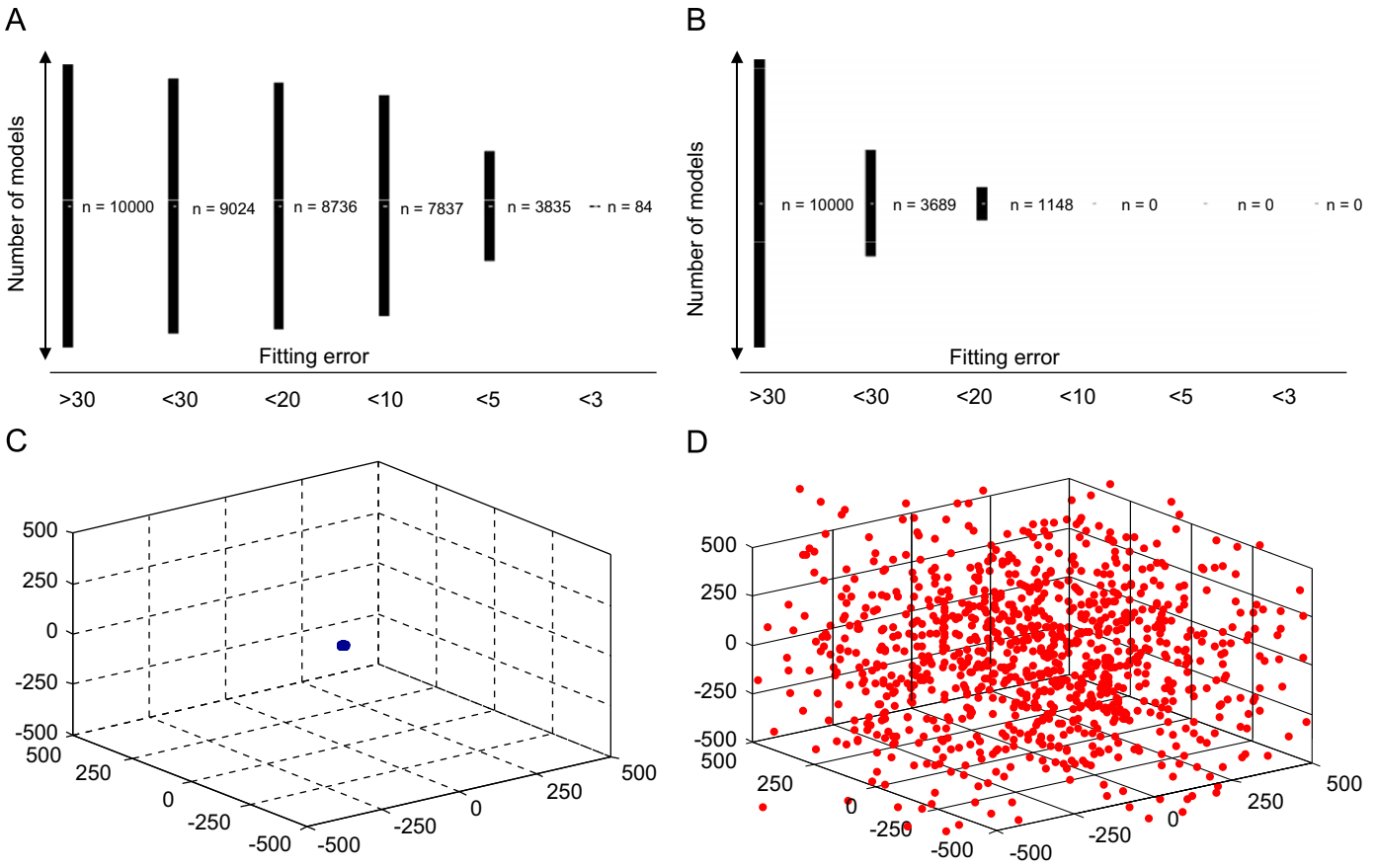

Fig. 7. (A) The distribution of models within certain fitting errors. When all the models are anchored to the same steady state, 84 out of 10,000 models behave very similarly to the experimental observation with fitting errors smaller than 3. (B) When the kinetic parameters are generated randomly without anchoring the models to the same steady state, none of the models has a fitting error smaller than 10, which is the range that could produce a reasonable match with the measured dynamic response. (C) Kinetic parameter space spanned by the steady-state flux constrained models (1000 models are used for analysis). The first three principal components are plotted here. (D) Kinetic parameter space spanned by the control models (1000 models are used for analysis) without the steady-state flux constraint.

related to enzyme molecular properties. As such, effects of genetic manipulations at the molecular level can be assessed using EM. This modeling approach is expected to aid metabolic engineering (Atsumi et al., 2008a, 2008b; Barkovich and Liao, 2001; Fischer et al., 2008; Ro et al., 2006) by interpreting data and guiding molecular manipulations in the future.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ymben.2010.11.001.

References

Alper, H., Miyaoku, K., Stephanopoulos, G., 2005. Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. Nat. Biotechnol. 23, 612–616.

Atsumi, S., Cann, A.F., Connor, M.R., Shen, C.R., Smith, K.M., Brynildsen, M.P., Chou, K.J., Hanai, T., Liao, J.C., 2008a. Metabolic engineering of *Escherichia coli* for 1-butanol production. Metab. Eng. 10, 305–311.

Atsumi, S., Hanai, T., Liao, J.C., 2008b. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. Nature 451, 86–89.

Barkovich, R., Liao, J.C., 2001. Metabolic engineering of isoprenoids. Metab. Eng. 3, 27–39.

Buchholz, A., Hurlebaus, J., Wandrey, C., Takors, R., 2002. Metabolomics: quantification of intracellular metabolite dynamics. Biomol. Eng. 19, 5–15.

Chao, Y.P., Liao, J.C., 1993. Alteration of growth yield by overexpression of phosphoenolpyruvate carboxylase and phosphoenolpyruvate carboxykinase in *Escherichia coli*. Appl. Environ. Microbiol. 59, 4261–4265.

Chao, Y.P., Patnaik, R., Roof, W.D., Young, R.F., Liao, J.C., 1993. Control of gluconeogenic growth by pps and pck in *Escherichia coli*. J. Bacteriol. 175, 6939–6944.
Chassagnole, C., Noisommit-Rizzi, N., Schmid, J.W., Mauch, K., Reuss, M., 2002. Dynamic modeling of the central carbon metabolism of *Escherichia coli*. Biotechnol. Bioeng. 79, 53–73.
Contador, C.A., Rizk, M.L., Asenjo, J.A., Liao, J.C., 2009. Ensemble modeling for strain development of L-lysine-producing *Escherichia coli*. Metab. Eng. 11, 221–233.
Cornish-Bowden, A., 1979. Fundamentals of Enzyme Kinetics. Butterworths & Co, London.
Dean, J.T., Rizk, M.L., Tan, Y., Dipple, K.M., Liao, J.C., 2009. Ensemble modeling of hepatic fatty acid metabolism with a synthetic glyoxylate shunt. Biophys. J. 98, 1385–1395.
Europa, A.F., Gambhir, A., Fu, P.C., Hu, W.S., 2000. Multiple steady states with distinct cellular metabolism in continuous culture of mammalian cells. Biotechnol. Bioeng. 67, 25–34.
Fell, D.A., 1992. Metabolic control analysis: a survey of its theoretical and experimental development. Biochem. J. 286 (Pt 2), 313–330.
Fischer, C.R., Klein-Marcuschamer, D., Stephanopoulos, G., 2008. Selection and optimization of microbial hosts for biofuels production. Metab. Eng. 10, 295–304.
Henry, C.S., Jankowski, M.D., Broadbelt, L.J., Hatzimanikatis, V., 2006. Genome-scale thermodynamic analysis of *Escherichia coli* metabolism. Biophys. J. 90, 1453–1461.
Hofmeyr, J.H., Cornish-Bowden, A., 1997. The reversible Hill equation: how to incorporate cooperative enzymes into metabolic models. Comput. Appl. Biosci. 13, 377–385.
Irvine, D.H., 1988. Efficient solution of nonlinear models expressed in S-system canonical form. Math. Comput. Modelling 11, 123–128.
Johannes, K.J., Hess, B., 1973. Allosteric kinetics of pyruvate kinase of *Saccharomyces carlsbergensis*. J. Mol. Biol. 76, 181–205.
Kahn, D., Westerhoff, H.V., 1991. Control theory of regulatory cascades. J. Theoret. Biol. 153, 255–285.
Kojima, H., Ogawa, Y., Kawamura, K., Sano, K., Method of producing L-lysine by fermentation. U.S. Patent vol. 6040160, Ajinomoto Co., Inc., 1996.
Liao, J.C., Delgado, J., 1993. Advances in metabolic control analysis. Biotechnol. Prog. 9, 221–233.
Lin, H., Bennett, G.N., San, K.Y., 2005a. Genetic reconstruction of the aerobic central metabolism in *Escherichia coli* for the absolute aerobic production of succinate. Biotechnol. Bioeng. 89, 148–156.
Lin, H., Bennett, G.N., San, K.Y., 2005b. Metabolic engineering of aerobic succinate production systems in *Escherichia coli* to improve process productivity and achieve the maximum theoretical succinate yield. Metab. Eng. 7, 116–127.
Link, H., Weuster-Botz, D., 2007. Steady-state analysis of metabolic pathways: comparing the double modulation method and the lin-log approach. Metab. Eng. 9, 433–441.
Lu, P., Rangan, A., Chan, S.Y., Appling, D.R., Hoffman, D.W., Marcotte, E.M., 2007. Global metabolic changes following loss of a feedback loop reveal dynamic steady states of the yeast metabolome. Metab. Eng. 9, 8–20.
Millard, C.S., Chao, Y.P., Liao, J.C., Donnelly, M.I., 1996. Enhanced production of succinic acid by overexpression of phosphoenolpyruvate carboxylase in *Escherichia coli*. Appl. Environ. Microbiol. 62, 1808–1810.
Nielsen, D.R., Leonard, E., Yoon, S.H., Tseng, H.C., Yuan, C., Prather, K.L., 2009. Engineering alternative butanol production platforms in heterologous bacteria. Metab. Eng. 11, 262–273.
Patnaik, R., Roof, W.D., Young, R.F., Liao, J.C., 1992. Stimulation of glucose catabolism in *Escherichia coli* by a potential futile cycle. J. Bacteriol. 174, 7527–7532.
Rizk, M.L., Liao, J.C., 2009. Ensemble modeling for aromatic production in *Escherichia coli*. PLoS One 4, e6903.
Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., Chang, M.C., Withers, S.T., Shiba, Y., Sarpong, R., Keasling, J.D., 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440, 940–943.
Sanchez, A.M., Bennett, G.N., San, K.Y., 2005. Novel pathway engineering design of the anaerobic central metabolic pathway in *Escherichia coli* to increase succinate yield and productivity. Metab. Eng. 7, 229–239.
Savageau, M.A., Voit, E.O., 1987. Recasting nonlinear differential equations as S-systems: a canonical nonlinear form. Math. Biosci. 87, 33.
Savageau, M.A., 1993. Finding multiple roots of nonlinear algebraic equations using S-System Methodology. Appl. Math. Comput. 55, 12.
Schellenberger, J., Palsson, B.O., 2009. Use of randomized sampling for analysis of metabolic networks. J. Biol. Chem. 284, 5457–5461.
Tran, L.M., Rizk, M.L., Liao, J.C., 2008. Ensemble modeling of metabolic networks. Biophys. J. 95, 5606–5617.
Usuda, Y., Nishio, Y., Iwatani, S., Van Dien, S.J., Imaizumi, A., Shimbo, K., Kageyama, N., Iwahata, D., Miyano, H., Matsui, K., 2010. Dynamic modeling of *Escherichia coli* metabolic and regulatory systems for amino-acid production. J. Biotechnol. 147, 17–30.
Visser, D., Heijnen, J.J., 2003. Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. Metab. Eng. 5, 164–176.
Wang, L., Birol, I., Hatzimanikatis, V., 2004. Metabolic control analysis under uncertainty: framework development and case studies. Biophys. J. 87, 3750–3763.
Zhang, K., Sawaya, M.R., Eisenberg, D.S., Liao, J.C., 2008. Expanding metabolism for biosynthesis of nonnatural alcohols. Proc. Natl. Acad. Sci. U.S.A. 105, 20653–20658.