Modeling the metabolic dynamics at the genome-scale by optimized yield analysis

Metabolic Engineering 75 (2023) 119–130

Contents lists available at ScienceDirect

Metabolic Engineering

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

Modeling the metabolic dynamics at the genome-scale by optimized yield analysis

Hao Luo [a], Peishun Li [a], Boyang Ji [a,b], Jens Nielsen [a,b,*]

[a] Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden
[b] BioInnovation Institute, Ole Måløes Vej 3, DK2200, Copenhagen N, Denmark

ARTICLE INFO

Keywords:
Genome-scale metabolic model
Cybernetic modeling
Flux-balance analysis
Elementary flux mode
Optimize yield
Yield space

ABSTRACT

The hybrid cybernetic model (HCM) approach is a dynamic modeling framework that integrates enzyme synthesis and activity regulation. It has been widely applied in bioreaction engineering, particularly in the simulation of microbial growth in different mixtures of carbon sources. In a HCM, the metabolic network is decomposed into elementary flux modes (EFMs), whereby the network can be reduced into a few pathways by yield analysis. However, applying the HCM approach on conventional genome-scale metabolic models (GEMs) is still a challenge due to the high computational demands. Here, we present a HCM strategy that introduced an optimized yield analysis algorithm (opt-yield-FBA) to simulate metabolic dynamics at the genome-scale without the need for EFMs calculation. The opt-yield-FBA is a flux-balance analysis (FBA) based method that can calculate optimal yield solutions and yield space for GEM. With the opt-yield-FBA algorithm, the HCM strategy can be applied to get the yield spaces and avoid the computational burden of EFMs, and it can therefore be applied for developing dynamic models for genome-scale metabolic networks. Here, we illustrate the strategy by applying the concept to simulate the dynamics of microbial communities.

1. Introduction

Mathematical modeling of cellular metabolism supports and accelerates the development of metabolic engineering and industrial bioprocesses (Aboulmouna et al., 2020; Nielsen and Keasling, 2016; Young, 2015). Among various types of mathematical models, genome-scale metabolic models (GEMs) and cybernetic models are widely used and have many successful applications underlying different assumptions and frameworks (Chen et al., 2022; Geng et al., 2021; Lu et al., 2019; Orth et al., 2011; Ramkrishna and Song, 2012).

A GEM is a collection of all the biochemical reactions occurring in the organism of interest, and these models systematically organize and process metabolic information. For example, GEMs can be used to calculate flux distributions at steady state or simulate phenotypes of different mutants and/or growth at different conditions (Bordbar et al., 2014). With the availability of genome sequences for many organisms, more than thousands of GEMs have been constructed. For analysis of these models, many standardized tools and analysis methods have been developed, such as flux balance analysis (FBA) (Gu et al., 2019; Henry et al., 2010; Lieven et al., 2020; Orth et al., 2010b; Amit Varma and Palsson, 1994).

Cybernetic modeling is an ordinary differential equation (ODE) modeling approach that incorporates control strategies for the optimal investment of enzymatic resources (Dhurjati et al., 1985; Straight and Ramkrishna, 1994). The hypothesis underlying cybernetic models is that microorganisms can regulate the synthesis and activity of enzymes as a survival mechanism for cells growing in different environments (Young and Ramkrishna, 2007). The approach has been applied to describe not only the dynamic behaviors of microorganisms in a mixture of carbon sources, but also to describe multi-species relationships (Martínez et al., 2020; Perrin et al., 2020; Song et al., 2009). To process more complex networks, some diverse methods of cybernetic modeling were developed such as first-generation lumped cybernetic models (Song and Ramkrishna, 2010, 2011) and second-generation hybrid cybernetic models (HCMs). The HCM is a widely used dynamic model in recent years, that uses elementary flux modes (EFMs) and yield analysis to reduce the number of pathways to be described with kinetic expressions (Geng et al., 2012; Kim et al., 2008; Kompala et al., 1986; Song and

Abbreviations: HCM, hybrid cybernetic model; GEM, genome-scale metabolic model; FBA, flux-balance analysis; EFM, elementary flux mode; EFV, elementary flux vector; opt-yield-FBA, optimized yield analysis algorithm; LP, linear programming.

* Corresponding author. Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden.
E-mail address: nielsenj@chalmers.se (J. Nielsen).

https://doi.org/10.1016/j.ymben.2022.12.001
Received 20 June 2022; Received in revised form 1 December 2022; Accepted 5 December 2022
Available online 9 December 2022

Table 1
The comparison of current methods to obtain yield space for HCM.

Methods	Correct yield space	Constrains bounded	Process complex GEMs
EFMs	×	×	×
EFVs	✓	✓	×
FBA modes	×	✓	✓
Opt-yield-FBA[a]	✓	✓	✓

[a] This study.

Ramkrishna, 2009). Using this framework, it is possible to describe metabolic dynamic behaviors well by only requiring the measurement of exchange fluxes and considering intracellular metabolites under pseudo-steady-state conditions (Kim et al., 2008; Song et al., 2009).

However, applying the HCM approach on conventional genome-scale metabolic models is still a challenge due to the high computational demand for calculating EFMs. EFMs are a set of all non-decomposable pathways containing a minimal set of reactions and these pathways can be considered as a catalog of all possible flux balance analysis (FBA) results (Schuster et al., 1999, 2000; Terzer and Stelling, 2008; Vilkhovoy et al., 2016). The number of EFMs is exponentially growing with the number of reactions in a network, and this is therefore a bottleneck for applying the HCM approach at the genome-scale directly (Vilkhovoy et al., 2016). According to the metabolic yield analysis (MYA) by Song and Ramkrishna, EFMs provide a master yield space and convex hull for pathways selection, and selected pathways' yield values are located on the boundaries of the yield space (Song and Ramkrishna, 2009). Therefore, to fill the gap between GEMs and HCMs without EFMs, the replacement methods should be able to obtain yield spaces from GEMs.

Some studies tried to replace the EFMs in the HCM approach by elementary flux vectors (EFVs) or FBA modes (Ahamed et al., 2021; Vilkhovoy et al., 2016). EFVs are alternatives to EFMs but can account for inhomogeneous constraints (Klamt et al., 2017, 2018; Müller and Regensburger, 2016). With inhomogeneous constraints, the EFVs can provide a theoretically correct yield space better than EFMs. Although EFVs are more reliable for calculating the yield space, the computational demand is still large (Ahamed et al., 2021). FBA modes cannot provide a complete yield space because they optimize the output rate of target metabolites, not yield (Klamt et al., 2018; Vilkhovoy et al., 2016). The comparison of current methods for optimizing yield or obtaining yield space for HCM is shown in Table 1.

To obtain the yield spaces from GEMs without EFMs, a yield optimization method is required. Typically, the yield is defined as a ratio of product fluxes and substrate uptake fluxes ($Y = r_p/r_s$), and represents a relative value of the amount of product or biomass formed per amount of substrate consumed (Klamt et al., 2018). Some studies tried to optimize yield by FBA, such as fixing the substrate uptake rate (r_s) as an experimentally measured value or one, then the maximized product fluxes or growth (r_p) represent maximized yield (Schuster et al., 2008; Teusink and Smid, 2006). The other mathematical frameworks for yield optimization have been developed by linear-fractional programming and higher-dimensional linear programming (Klamt et al., 2018).

In this study, we present an optimized yield analysis algorithm (opt-yield-FBA) that converts the objective function of yield from the nonlinear problem of r_p/r_s to a linear problem of $r_p - Y_{temp} \bullet r_s$, and the algorithm can thus compute, analyze optimal yield solutions and obtain yield spaces from GEMs. With the yield space obtain using opt-yield-FBA rather than EFMs, the HCM strategy can be applied to describe metabolic dynamics at the genome-scale. We illustrate the HCM strategy by metabolic networks of different sizes and use the generated models to simulate the dynamics of microbial communities. Finally, we also evaluate the robustness of the yield space and the challenges of HCM.

2. Methods

2.1. Metabolic flux analysis

Flux balance analysis (FBA) is a widely used method for studying GEMs. It is based on using an objective function to calculate flux distributions by linear programming (Orth et al., 2010b; Schuetz et al., 2012; Amit Varma and Palsson, 1994) combined with a mass balance constraints around the individual metabolites. GEM manipulation was performed by the cobrapy 0.17.0 (Ebrahim et al., 2013) on Python 3.8.

EFMs and EFVs calculations were performed by efmtool 4.7.1 (Terzer and Stelling, 2008) and CellNetAnalyzer 2019.2.842 (Klamt et al., 2007) on MATLAB 2018b.

2.2. Metabolic yield analysis and pathway selection

Metabolic yield analysis is used to reduce pathways based on the convex hull in a yield space. Pathways with flux distributions come from EFMs, EFVs or opt-yield-FBA, and yield values of products could be calculated by $Y = r_p/r_s$. All yield values construct a convex shape and provide the master yield space. For prior reduction, a subset of pathways that cover 99% yield space can be obtained by calculating subsets convex hull volumes (Song and Ramkrishna, 2009). The threshold of 99% is suggested by Song and Ramkrishna's study and the resulting subset is a minimal number of pathways that contribute 99% volumes of yield space (Song and Ramkrishna, 2009).

For active pathways selection with experimental data, the sum of squared weights or the distance is minimized, as described by the following equations. When the yield of experimental data in the master yield space,

$$\min_{\mathbf{h}} \frac{1}{2}\|\mathbf{h}\|_2^2$$

$$\text{Subject to : } \mathbf{Z}_y\mathbf{h} - \mathbf{y}_\mathbf{m} = \mathbf{0}$$

$$\mathbf{h} \geq 0$$

$$\|\mathbf{h}\|_\mathbf{1} = 1$$

When the yield of experimental data is outside the master yield space,

$$\min_{\mathbf{h}} \frac{1}{2}\left\|\mathbf{Z}_y\mathbf{h} - \mathbf{y}_\mathbf{m}\right\|_2^2$$

$$\text{Subject to : } \mathbf{h} \geq 0$$

$$\|\mathbf{h}\|_\mathbf{1} = 1$$

Here $\mathbf{h}$ (n_{paths}), is a vector of weight for pathways and all weights are positive ($\mathbf{h} \geq 0$) and the sum of weights is one ($\|\mathbf{h}\|_\mathbf{1} = 1$). $\mathbf{Z}_y$ (n_{mets} * n_{paths}) is a yield matrix of all pathways and products, and $\mathbf{y_m}$ (n_{mets}) is a product yield vector from experimental data. The "$\mathbf{Z}_y\mathbf{h} - \mathbf{y}_\mathbf{m} = 0$" is the distance between the master yield space and the experimental yield vectors. If the experimental yield is in the master yield space the distance is zero. When the experimental yield is outside, the distance should be minimized (Song and Ramkrishna, 2009).

The convex hull and its volume were calculated by SciPy 1.8.0 and the selection of active pathways was performed by CVXPY1.2.0 (Agrawal et al., 2018) on Python.

2.3. Cybernetic modeling

The ordinary differential equations of HCM describe metabolite mass balances and enzyme balances. The $\mathbf{x}$ is variables vector that contains extracellular metabolite concentrations ($\mathbf{m_{ex}}$), biomass concentration (c) and enzyme levels ($\mathbf{e}$), the $\mathbf{x}$ vector could be written like Equation (1)

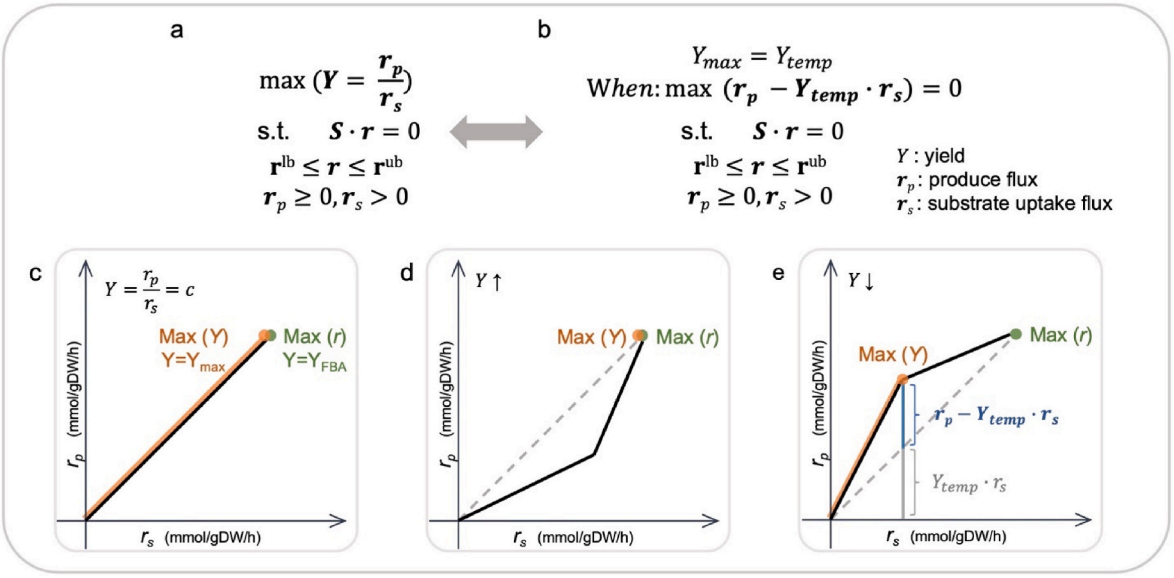

Fig. 1. Overview of the yield optimization algorithm. a) Formulas of the yield optimization problem. The objective function of yield is the ratio of r_p and r_s, where r_p represents the production rate and r_s represents the substrate uptake rate. The optimization is subject to a steady-state ($S \bullet r = 0$) with constraints ($r^{lb} \le r \le r^{ub}$) and the directions of r_p and r_s are both defined as positive. b) The approximate solution of the optimal yield. The optimal yield solution is generated by a series of FBA solutions iteratively. In each iteration, an assumed yield is defined as Y_{temp} and the optimized objective function ($r_p - Y_{temp} \bullet r_s$) is set as the differences between r_p and the product of Y_{temp} and r_s. The initial value of Y_{temp} is Y_{FBA} (yield of green dots), which is the yield value when rate optimized by FBA. The return of each iteration is a vector of flux distribution that updates Y_{temp} for the next iteration. The iterations will be terminated when the optimization return zero and the maximum yield is found. c-e) Potential relationships of the substrate uptake rate (r_s) and the production rate (r_p). The maximal r_p value from FBA is shown as a green dot and the maximal yield is shown with an orange dot or line. In c and d, the max(r) and max(y) overlap and $Y_{FBA} = Y_{max}$, and the iterations are terminated in the first iteration. In e, the max(r) and max(y) points do not overlap and the maximal Y requires multi optimal calculations. The objective function of $r_p - Y_{temp} \bullet r_s$ is shown by the blue line and text. c) Illustration of a case where the maximum yield is at the maximum flux and the yield is constant for all set of fluxes. d) Illustration of a case where the maximum yield is at the maximum flux but the yield is not constant for all set of fluxes. e) Illustration of a case where the maximum yield is at not the maximum flux.

$$\mathbf{x} = \begin{bmatrix} \mathbf{m_{ex}} \\ \mathbf{e} \\ c \end{bmatrix} \tag{1}$$

Here biomass concentration c can be included in the vector of extracellular metabolite concentrations ($\mathbf{m_{ex}}$). HCM focuses on extracellular metabolites and all intracellular metabolites under a pseudo-steady-state $d\mathbf{m_{in}}/dt = 0$. Mass balances for extracellular metabolite concentrations ($\mathbf{m_{ex}}$), enzyme levels ($\mathbf{e}$) and biomass concentration (c) are described by Equations (2)–(4).

$$\frac{d\mathbf{m_{ex}}}{dt} = \mathbf{S_{ex}} \bullet \mathbf{Z} \bullet \mathbf{diag}(\mathbf{v}) \bullet \mathbf{r} \bullet c \tag{2}$$

$$\frac{d\mathbf{e}}{dt} = \boldsymbol{\alpha} + \mathbf{diag}\left(\mathbf{u}\right) \bullet \mathbf{r_E} - \left(\mathbf{diag}\left(\boldsymbol{\beta}\right) + \mu\right) \bullet \mathbf{e} \tag{3}$$

$$\frac{dc}{dt} = \mu \bullet c \tag{4}$$

In Equation (2), $\mathbf{m_{ex}}$ (n_{mets}) is a vector of extracellular metabolite concentrations; $\mathbf{S_{ex}}$ (n_{mets} * n_{rxns}) is the stoichiometric matrix of the extracellular metabolites; $\mathbf{Z}$ is the (n_{rxns} * n_{paths}) stoichiometric matrix of pathways and extracellular metabolites, the pathways could come from EFMs or EFVs or opt yield FBA; $\mathbf{diag}(\mathbf{v})$ is a diagonalization $\mathbf{v}$ (n_{paths}) and $\mathbf{v}$ is the vector of the cybernetic control variables that regulate the enzyme activities; $\mathbf{r}$ (n_{paths}) is the vector of exchange fluxes from corresponding pathways, c is the biomass concentration per unit volume. In Equation (3), $\boldsymbol{\alpha}$ (n_{paths}), denotes the vector of the constitutive rates of enzyme synthesis; $\mathbf{r_E}$ represents the maximum enzyme synthesis rate which inductive synthesis of the enzyme occurs without resources limitation. The $\mathbf{diag}(\mathbf{u})$ is a diagonalization $\mathbf{u}$ (n_{paths}) and $\mathbf{u}$ is a vector of the cybernetic control variables that regulate the enzyme synthesis; $\boldsymbol{\beta}$ (n_{paths}) is a vector of is enzyme degradation rates, $\mu \bullet \mathbf{e}$ is the dilution rate by growth. In Equation (4), μ is the specific growth rate.

The specific pathway rate, enzyme synthesis rate and growth rate are described by Equation (5) (6) and (7).

$$r_i = k_{maxi} \bullet e_i^{rel} \bullet \frac{s_j}{(K_{ij} + s_j)} \tag{5}$$

$$r_{Ei} = k_{ei} \bullet \frac{s_j}{(K_{ij} + s_j)} \tag{6}$$

$$\mu = \mathbf{h} \bullet \mathbf{r} \tag{7}$$

Equations (5) and (6) are the Michaelis–Menten type of kinetic expression of the pathway rate and the enzyme synthesis rate. Here i is the index of the pathway, j is the index of the substrate, different pathways could consume different substrates; where k_{maxi} is the reaction rate constant (g/gDW/h); K_{ij}*and* k_{ei} are the Michaelis–Menten constant (g/L); s_j denotes the substrate concentration; $\mathbf{h}$ is the matrix to find the biomass-related pathways and sum their rates. The e_i^{rel} is the relative enzyme level, which is described by Equation (8):

$$e_i^{rel} \equiv \frac{e_i}{e_i^{max}} \tag{8}$$

$$e_i^{max} = \frac{\alpha_i + k_{ei}}{\beta_i + \mu_i^{max}} \tag{9}$$

Equation (9) is under the condition of maximized enzyme level and comes from enzyme balance in equation (3). When the left-hand side of Equation (3) is zeroed and u = 1, the maximum enzyme level can be established as equation (9).

The matching law and proportional law to specify the cybernetic control variables shown by Equations (10) and (11).

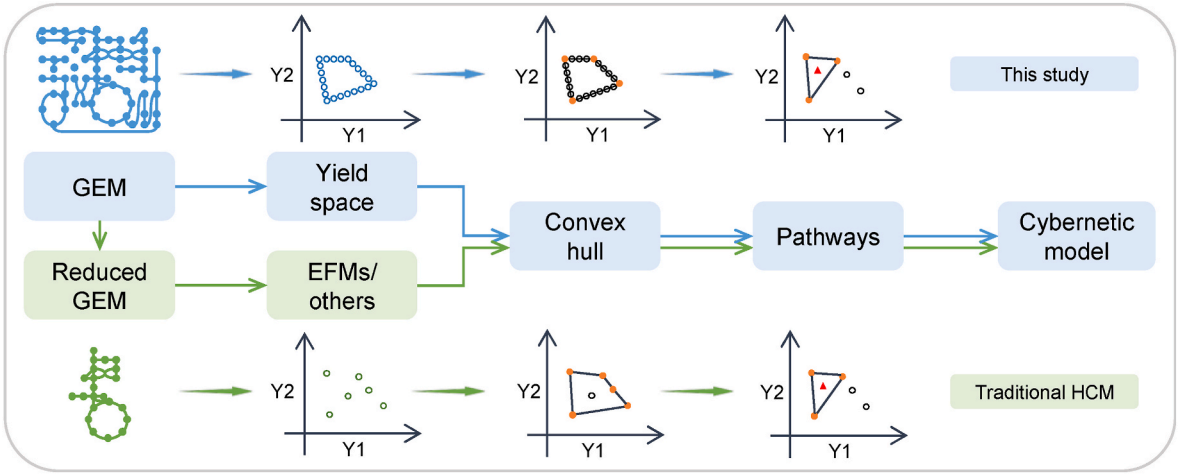

Fig. 2. Overview of the HCM strategy and comparison with the traditional HCM. The HCM strategy in this study (top part) and the traditional HCM strategy (bottom part). In our HCM strategy, a yield space is calculated by the opt-yield-FBA from a complete GEM. All yield values of pathways from opt-yield-FBA are located on the boundaries of the yield space. The traditional HCM strategy processes small or medium size models and calculates EFMs to obtain the yield space. The yield values of EFMs are located both on the boundaries of space and inside of the yield space. The pathways from opt-yield-FBA or EFMs provide the master yield space for pathways selection by MYA and convex hull. The active pathways can be further selected with experimental data. Based on the pathways selected, the HCM approach can be applied.

$$u_i = \frac{p_i}{\sum_j p_j} \tag{10}$$

$$v_i = \frac{p_i}{\max(p_j)} \tag{11}$$

The variables v_i and u_i are bounded between 0 and 1, and p is the return on investment defined by Equation (12)

$$p_i = f_{carboni} r_i \tag{12}$$

Here $f_{carboni}$ is the number of carbon atoms in the substrate through the i_{th} pathway.

Codes are available in GitHub: https://github.com/SysBioChalmers/GEM2CB_model

3. Results and discussions

3.1. Optimized yield analysis for GEMs and the opt-yield-FBA

Yield is a crucial requirement for industrial fermentation and commercial production, which describes the units of product synthesized per unit of substrate consumed. As shown in Fig. 1a, the yield is defined as a fractional form of $Y = r_p/r_s$, which is the ratio of the product output rate (r_p) and the substrate uptake rate (r_s). This optimization is a simplified linear-fractional programming (LFP) problem in mathematics, which is challenging to solve directly with current LP methods or tools. To be compatible with current GEM analysis tools and LP solvers, the LFP needs to be transformed into an LP. In Fig. 1b, we transform the objective function of yield optimization from a fraction into a linear formula. The optimization is subject to steady-states ($S \bullet r = 0$) and reversibility bounds ($r^{lb} \le r \le r^{ub}$). The directions of r_p and r_s are defined as being positive in the model and both reactions are irreversible because the product cannot be absorbed, and the substrate cannot be produced in this case. Based on the definition of yield, the denominator cannot be zero, and the additional constraints of $r_p \ge 0$ and $r_s > 0$ are added. We assume there is a temporary yield ($Y_{temp} \ge 0$) to approach the optimal yield (Y_{max}), and the objective function is set as $r_p - Y_{temp} \bullet r_s$. In fact, when the Y_{temp} is set as the minimum, i.e. at zero, $Y_{temp} \bullet r_s = 0$, and the objective function will be r_p, which means that the objective function is the same as maximizing the production rate and it can be solved by FBA. Thus, the initial value of Y_{temp} can be set as the yield when rate is optimized (Y_{FBA}). The objective function is set as the difference between the real production rate (r_p) and the production rate from the substrate rate (r_s) according to the temporary yield (Y_{temp}). In the first iteration, a vector of flux distribution that includes r_{p1} and r_{s1} is returned and can be used to generate a new $Y_{temp1} = r_{p1}/r_{s1}$ and the Y_{temp1} which will be integrated into the objective function for the next round of optimization. The iteration is applied to find a yield value higher than the current Y_{temp}. If the objective function $r_p - Y_{temp} \bullet r_s$ is reaching zero, it means that the current Y_{temp} is the maximized yield.

As shown in Fig. 1c–e, the potential relationships between substrate uptake rate (r_s ,x − axis) and the production rate (r_p, y-axis) are shown. If the model constraints are continuous and homogeneous, the r_p and r_s usually have a linear correlation (Fig. 1c), and the yield will be a constant, which can be represented by the slope of the lines (grey dashed lines). In practice, inhomogeneous constraints are usually present in a GEMs, especially when enzymatic or expression constraints are integrated (Sánchez et al., 2017; Yang et al., 2021). With inhomogeneous constraints, the flux distribution can be concave or convex, as in Fig. 1d and **e**. In Fig. 1d when there are limitations for growth requirements like maintaining basic growth or energy. As substrate uptake increases, the r_p increases rapidly when growth requirements are fulfilled, and more substrates are used for products. In this case, the maximum yield (orange point) is reached to yield of maximized production rate r_p (green point). In Fig. 1e when the model encounters more constraints with substrate uptake like oxygen limitation. The flux distributions of r_p and r_s could be described as convex and the state with maximized r_p (green point) will not correspond to the state with maximized yield (orange point). The objective function of differences between the current calculated yield with the maximum yield is shown in blue text in Fig. 1e. Similarly, we can also use this strategy to get the minimum yield Y_{min}. When the substrate uptake rate is close to zero ($r_s \to 0$), there is a problem of yield closing to infinite ($Y_{max} \to \infty$). That is against yield definitions and product cannot be produced without substrate and the substrate must bound larger than zero because the denominator of yield cannot be zero. Finally, the opt-yield-FBA could be considered as a derivative of the FBA method with an objective function of $r_p - Y_{temp} \bullet r_s$, which can be used to search maximal yield with a serial FBA solution.

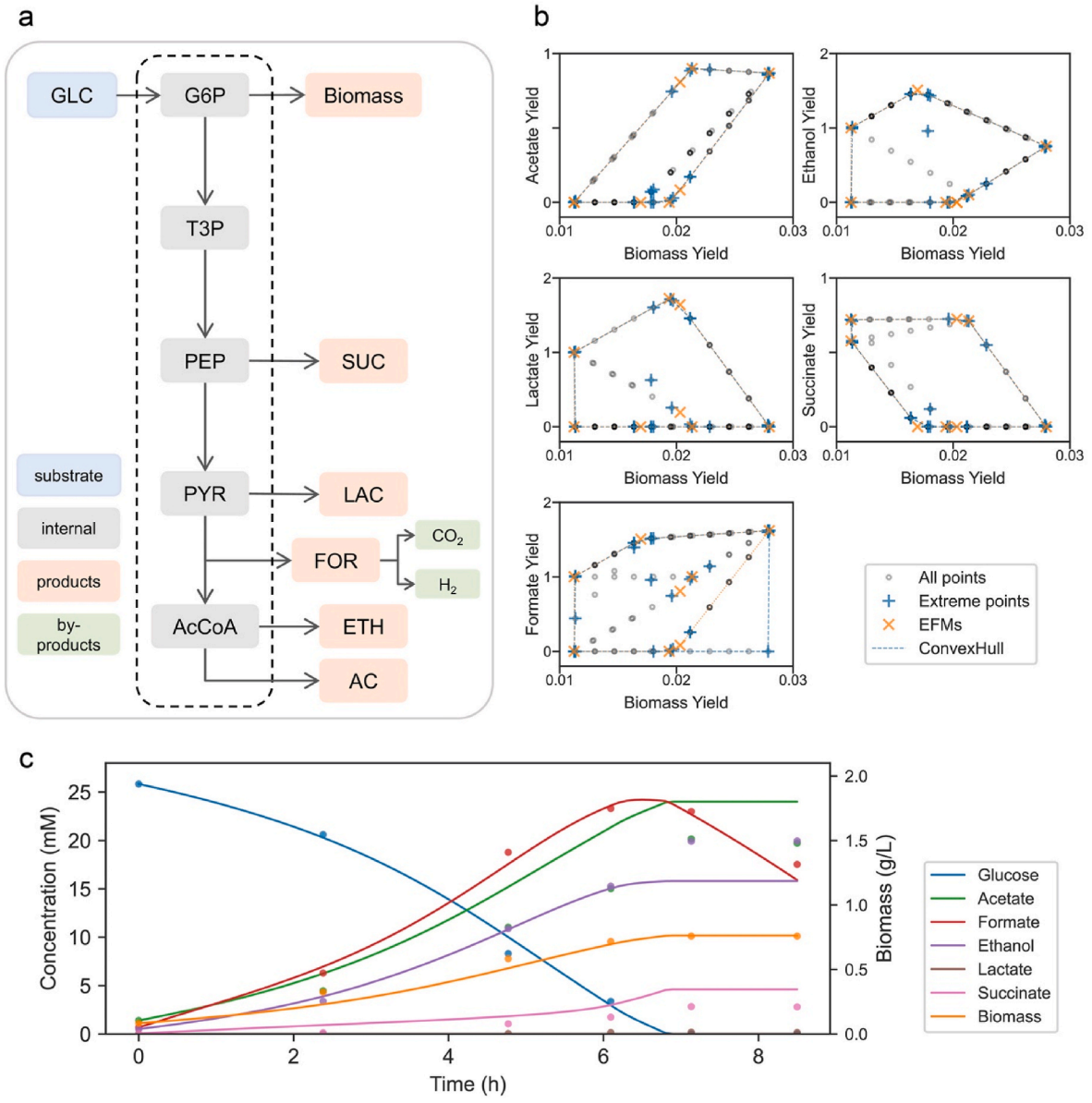

Fig. 3. Overview of a reduced *E. coli* network used for illustration, the yield space, and simulation results. a) Overview of the receded *E. coli* network, which contains 12 reactions and 14 metabolites. GLC (glucose), G6P (glucose 6-phosphate), T3P (fructose 1,6-biphosphate), PEP (phosphoenolpyruvate), PYR (pyruvate), AcCoA (acetyl coenzyme-A), SUC (succinate), FOR (formate), ACT (acetate), LAC (lactate), ETH (ethanol), B (biomass), CO_2 and H_2. b) The yield distributions are calculated by EFMs and opt-yield-FBA for acetate, ethanol, lactate, succinate, and formate. The grey dot indicates the pathway generated by opt-yield-FBA (the dark grey dots are overlapping yield values from different pathways), most of them located on the boundaries. The blue '+' marker indicates the extreme point of opt-yield FBA pathways, and the dashed line indicates the convex hull of the yield space. The orange 'x' marker indicates the pathway by EFMs. c) The simulation of biomass and metabolic dynamics under anaerobic conditions. The dot indicates experimental data, while the lines are obtained from the simulation.

3.2. The HCB strategy with opt-yield-FBA and application at the small scale

Our HCM strategy gets the yield space by opt-yield-FBA and avoids the calculation difficulties associated with identification of EFMs. To obtain the yield space, there are two types of methods, one is to calculate all feasible pathways in a network like EFMs and EFVs, the other method is to identify the boundaries of the yield space by yield optimization like opt-yield-FBA. Usually, the second method can save computing sources for complex models because of few pathways calculation and focus on the boundaries of the yield space. In the second method like opt-yield-FBA, the range of yield values can be identified by maximizing and minimizing a single target product, for pairs of target products, the yields space can also be calculated by sampling. For example, as shown in Fig. 2, after getting the maximum and minimum yield values of Y1 ($Y1_{max}$, $Y1_{min}$, X-axis), any values between $Y1_{min}$ and $Y1_{max}$ can be sampled as constraints to optimize the Y2 (Y-axis). After calculating the maximum and minimum yield values of Y2 in different sampled intervals from $Y1_{max}$ and $Y1_{min}$, the yield space of Y1 and Y2 can be obtained. Compared to EFMs or EFVs, opt-yield-FBA can obtain yield spaces straightforwardly. The opt-yield-FBA algorithm can be implemented by the FBA framework and is compatible with most current modeling tools or solvers since they process models by linear programming such as COBRApy (Ebrahim et al., 2013). With the replacement of EFMs by opt-yield-FBA, the strategy of HCM makes it possible to process complex GEMs directly and apply the HCMs approach to describe metabolic dynamics GEMs (Fig. 2).

We first illustrate the HCM strategy by using a reduced *E. coli*

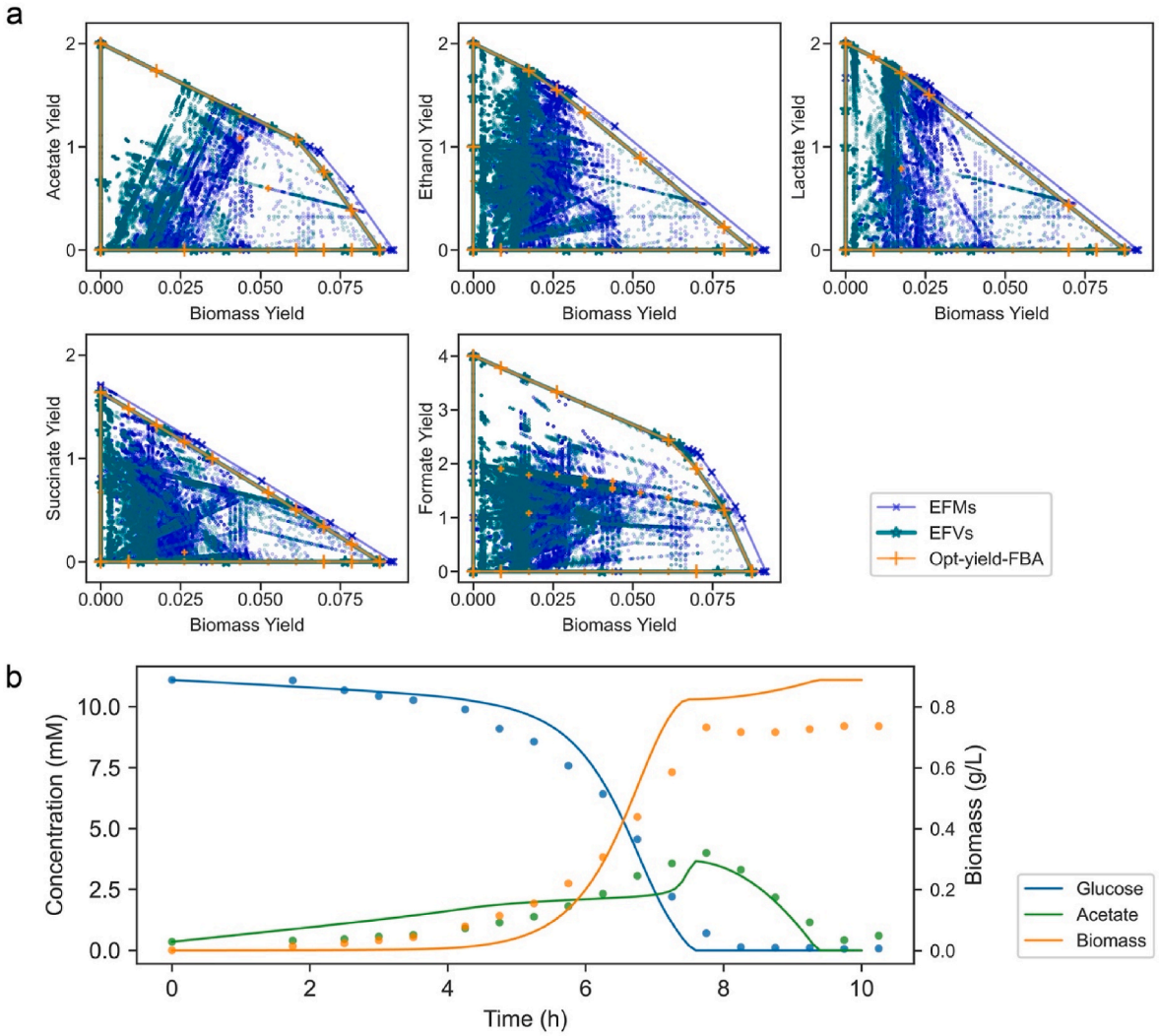

Fig. 4. The yield spaces from EFMs, EFVs and opt-yield-FBA, and corresponding simulation using the *E. coli* core metabolic model. a) Yield spaces of acetate, ethanol, lactate, succinate, and formate by EFMs, EFVs, and opt-yield-FBA. The X-axis represents the biomass yields, and the Y-axis represents the product yields. The dots and markers denote the yield values distribution, and the line denotes the convex hull in a yield space. The colors of blue, green, and orange correspond to the EFM, EFV, and opt-yield-FBA respectively. The EFVs' yield space is considered as a theoretically correct result. The yield space from opt-yield-FBA is the same as that of EFVs. EFMs have a larger yield space because inhomogeneous constraints in the model are not considered. b) The simulation of fermentation. The dots denote the experimental observations, while the lines denote the simulation from the cybernetic model.

metabolic network under anaerobic conditions (Kim et al., 2008). This small-scale network contains 12 internal reactions and 19 metabolites (Table S1) (Kim et al., 2008). The main metabolic pathways under anaerobic conditions are shown in Fig. 3a, which include one substrate (glucose), five internal metabolites, and eight products. The simulation of the cybernetic model relies on seven external metabolites including glucose, succinate, formate, acetate, lactate, ethanol, and biomass. The first step of the strategy is yield space calculation. The biomass yield is defined as the flux of biomass divided by the flux of glucose consumption, written as $Y_{biomass} = r_{biomass}/r_{\text{glucose}}$. The range of biomass yield was calculated by opt-yield-FBA. When calculating the minimum yield value, the optimization direction was set as 'min'. The minimum and maximum yield values are 0.0112 and 0.027, respectively. Then, we divided the biomass yield range into ten intervals. For each biomass yield interval value, we set the biomass yield as a constraint of the model, and search for the minimum and maximum acetate yield (Y_{ac}.). For the biomass-acetate yield space, we obtained 22 pathways (10 equal intervals with both max and min, grew points in Fig. 3b), which form a convex set (dashed lines region in Fig. 3b). Furthermore, the smallest set of extreme points was obtained by the convex hull process and these extreme points can be represented as the entire yield space (blue '+' in Fig. 3b. Finally, we obtained 22 pathways for each targeted metabolite, and 38 pathways at extreme points were selected to cover the five complete yield spaces. Compared to the yield space from EFM, the shape of yield space from opt-yield-FBA is the same as that from EFMs except for formate. The formate decomposition reaction (FOR => $CO_2 + H_2$) is an independent mode of EFMs without glucose consumption. Therefore, the formate yield of that EFM cannot be calculated without glucose consumption (the denominator of formate yield), but the formate can be decomposed freely in the model. With opt-yield-FBA, the glucose consumption was bounded greater than zero to avoid this issue. After employing experimental data, six activity pathways were selected by MYA for cybernetic modeling. After estimating the parameters using experimental data by least squares, the simulation showed well agreement with the experimental observations (Fig. 3c).

3.3. *The opt-yield-FBA based performance is robust for the medium scale network*

To test the availability and correctness of opt-yield-FBA for a larger

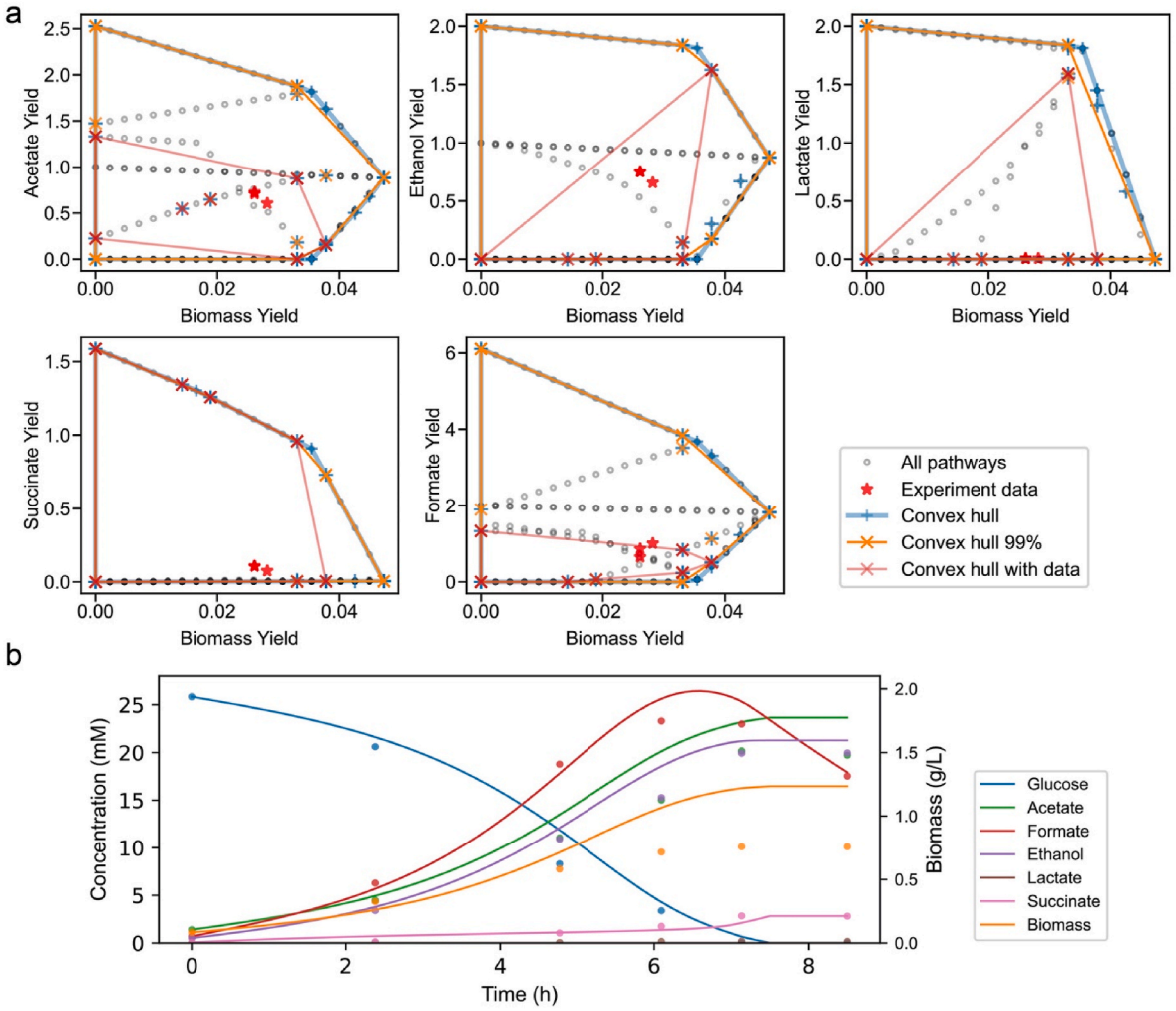

Fig. 5. The yield space analysis and simulation using the *E. coli* *i*ML1515 model. a) The yield space of acetate, ethanol, lactate, succinate, and formate. The X-axis represents the biomass yields, and the Y-axis represents the product yields. The dots and markers denote the yield distribution of pathways, and the lines denote the yield space. The grey dots are all pathways generated by opt-yield-FBA (the dark grey dots are overlapping yield values from different pathways) and the red stars denote experimental data. The blue and orange lines are convex hulls which covering complete and 99% yield spaces. The red line denotes the convex hull of active pathways identified by MYA with experimental data. b) The cybernetic model simulation is under anaerobic conditions. The dots denote the experimental observations, and the lines denote simulations of cybernetic models.

metabolic network, a core model of *E. coli* containing 95 reactions and 60 metabolites was applied (Orth et al., 2010a). The oxygen exchange reaction was set to be capable to simulate aerobic conditions. The model calculation generated 100,273 EFMs and 95,106 EFVs. We also calculated the pairwise yield space of five metabolites and biomass by opt-yield-FBA, and 110 opt-yield-FBA pathways were generated (Fig. 4a). As previously mentioned, the FEVs could consider inhomogeneous constraints better than EFMs, and the *E. coli* core model contains many such constraints like minimal maintain energy bounds. The yield space from EFVs is considered as a theoretically correct result which is smaller than that from EFMs (Fig. 4a). The difference between the EFMs yield space (blue lines) and the EFVs yield space (green lines) may be related to the inhomogeneous constraints. The pathways from opt-yield-FBA are significantly less than the other two methods and are mainly located at the boundaries of the yield space. The yield space of opt-yield-FBA is almost the same as the EFVs' with 99.888% ± 0.098% and 100% similarity in the convex hull area and parameters (Table S3 and Fig. 4a), which indicates similar performance of opt-yield-FBA as EFVs'. More quantity comparison of EFMs, EFV and opt-yield-FBA are shown in Table S3. During the calculation, 110 pathways from opt-yield-FBA could represent the same yield space of 95,106 EFVs and a more accurate yield space of 100,273 EFMs. After that, we performed MYA based on the convex hull to further reduce the membership of the pathways and identified the smallest set of 38 pathways that cover the entire yield space. Finally, we identified five active pathways by experimental data to construct a cybernetic model to simulate the metabolic dynamics under aerobic conditions (A. Varma and Palsson, 1994). The simulation result was shown in Fig. 4b, the dots are the experimental concentrations, and the lines are the simulation results of the cybernetic model. All model parameters can be found in Table S4.

3.4. The metabolic dynamics prediction at the genome scale

One of the challenges of the traditional hybrid cybernetic model approach is that the EFMs based methods are hard to process the complex metabolic networks at the genome-scale. Here, we applied opt-yield-FBA with *E. coli* *i*ML1515, which contains 1515 genes, 2719 reactions, and 1192 metabolites (Monk et al., 2017). In this case, the calculation of EFMs or EFVs using tools like efmtools (Terzer and Stelling, 2008) is challenging due to the large computational demand. As previously described, opt-yield-FBA is an FBA based method and can be used to calculate *i*ML1515 yield space of succinate, formate, acetate, lactate, and ethanol (Fig. 5a). There are 220 pathways generated in total. A following convex hull yield analysis was used to reduce the

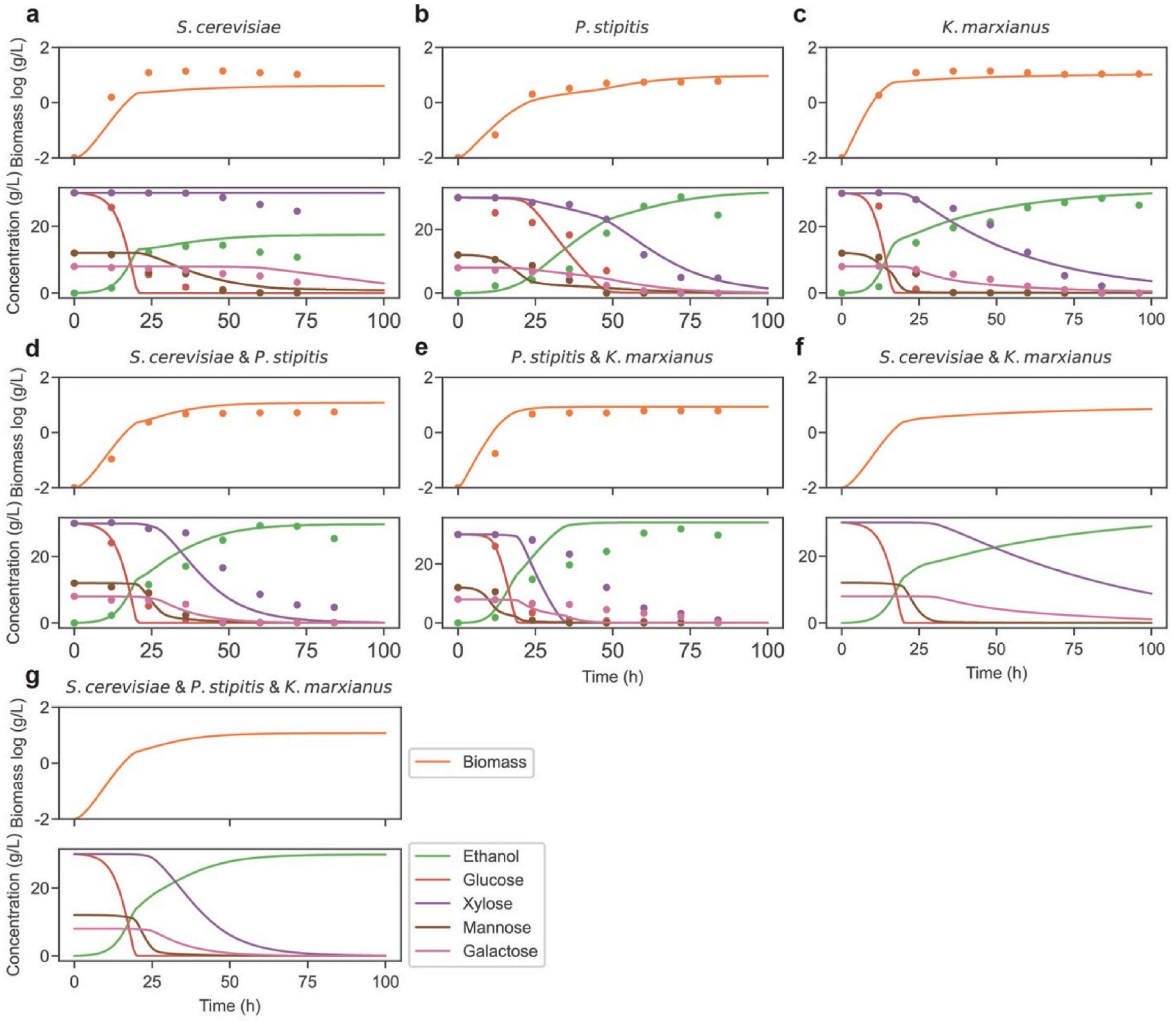

Fig. 6. Use of the HCM strategy for predicting three yeast species interactions. a-c) The simulation of monoculture including *Saccharomyces cerevisiae, Pichia stipites* and *Kluyveromyces marxianus*. Both the biomass (log g/L) and metabolites concentrations (g/L) were shown. d) The simulation of two species co-culture of *S. cerevisiae* and *P. stipites.* e) The simulation of two species co-culture of *P. stipites* and *K. marxianus.* f) The prediction of three species co-culture. The dots denote the experimental observations, while the lines denote results generated by the cybernetic model simulation.

number of pathways and 37 pathways were selected to represent the five pairwise yield spaces. We also calculated a set of pathways that covers 99% of the volume of the original convex hull (orange line and * in Fig. 5a). As reference suggested (Song and Ramkrishna, 2009), the threshold was set as 99% and that threshold could reduce the number of pathways and has little effect on the yield space. We also tried different thresholds of the convex hull to test the robustness of yield spaces calculation shown in Fig. S3 and Table S6. Based on the experimental data (red stars in Fig. 5a) (Kim et al., 2008), we identified the final active pathways by MYA (red lines and 'x' in Fig. 5a), these are a minimal set of pathways that could cover the experimental yield. We collected the experimental data of six metabolites and biomass concentrations under anaerobic conditions. After estimation of the parameters, the cybernetic model was used to simulate dynamic conditions as shown in Fig. 5b. All metabolites concentration simulations agree with experimental data, especially formate degradation. The biomass formula and its molar mass should be considered, and the trend of biomass also agrees with experimental data. All the model parameters can be found in Table S5.

3.5. Prediction of multi-species interactions

With the advancement of fermentation technology and the increasing recurrence of research systems, dynamic models are also widely applied to multi-species fermentation as well as for studying the gut microbiota. To evaluate our modeling approach, we tested multi-species models to test whether our method can be used to simulate the dynamics of microbial consortia. The first case includes a system consisting of three yeast, *Saccharomyces cerevisiae, Pichia stipites,* and *Kluyveromyces marxianus* (Geng et al., 2012). This simple system was simulated by a reduced model by the traditional HCM, and we performed our HCM strategy with a complete yeast GEM (Geng et al., 2012) (Lu et al., 2019). These three yeast have similar fermentation characteristics, and most interspecific relationships are competitive for absorbing the same carbon source (Rouhollah et al., 2007). For simplicity we used the same GEM for the three species like done in the traditional HCM analysis (Geng et al., 2012) and the GEM were adopted from the Yeast8 model, which contains 3989 reactions and 2693metabolites (Lu et al., 2019). 88 pathways were first identified to provide the master yield spaces and after the convex hull yield analysis, 17 pathways were selected to cover 99% of the master yield space. We identified different sets of active pathways for the three species using MYA with different fermentation data. For *S. cerevisiae* mono-culture, nine active pathways are selected to cover the experimental yield space; for *P. stipites* 10 active pathways were selected, and 10 active pathways were selected for *K. marxianus*. As shown in Fig. 6a–e, we estimated a set of the parameters that could simulate the mono-culture

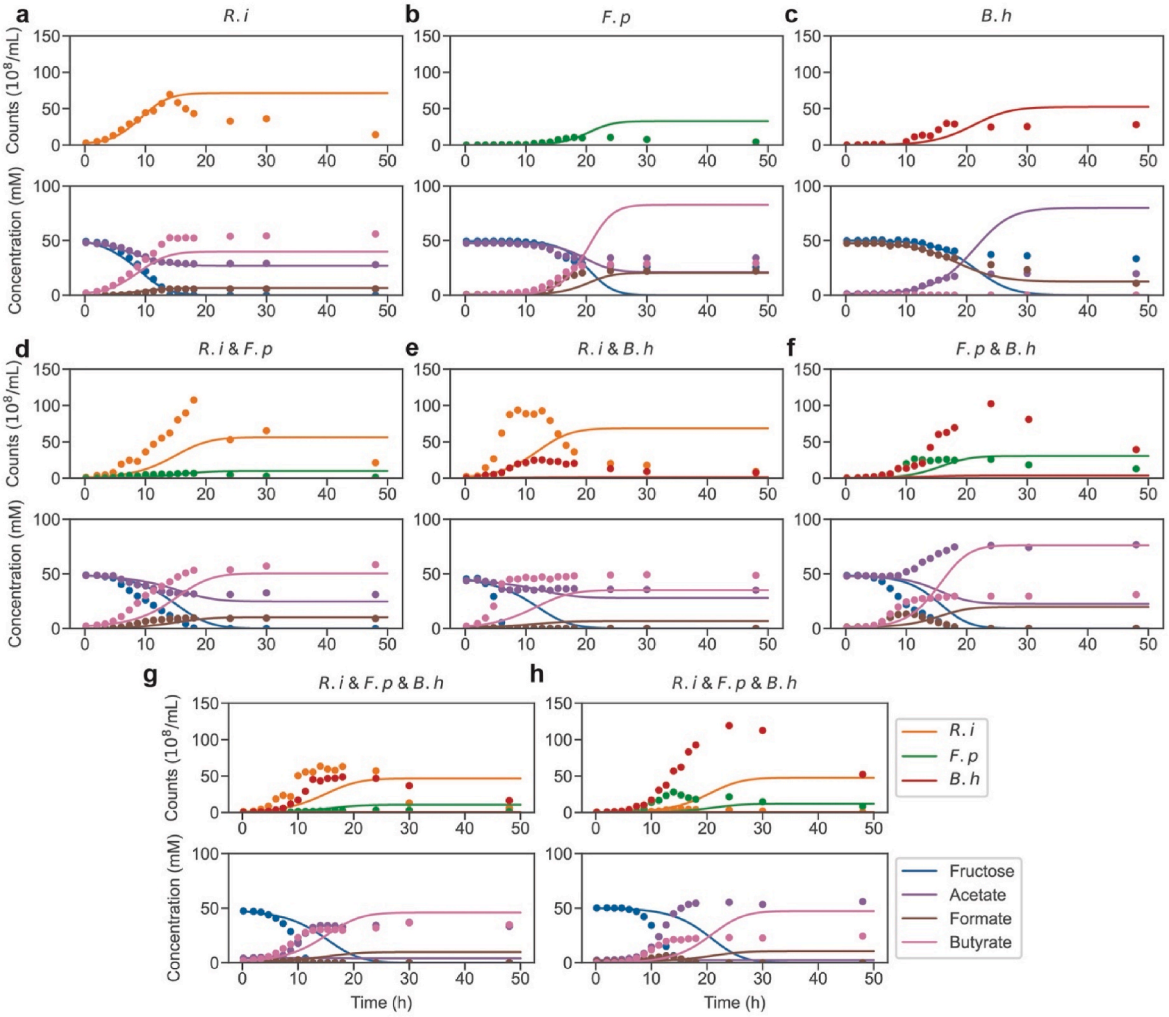

Fig. 7. The simulation of human gut microbial consortia with three members. a-c) The simulation of single species mono-culture of *Roseburia intestinalis (R.i), Faecalibacterium prausnitzii (F.p)*, and *Blautia hydrogenotrophica (B.h)* d-f) The simulation of two species co-cultures. g) The simulation of three species co-cultures. The dots denote the experimental measurements, while the lines denote model simulations. Both the cells count (upper part), and metabolites concentrations (mM, lower part) are shown.

and two species co-culture. Next the same parameters were used to predict three-species co-cultures. Because the experimental date only observed the total biomass, for the culture of muti species, the biomass of different species was assumed to be the same.

Complex systems such as the gut microbiota system are a challenge for mathematical modeling. Models of these complex systems often have high requirements for both modeling mechanisms and complete experimental data. To validate the ability of our strategy in gut microbiota and understand its limitations for complex microbial communities, a three-bacterial cross-feeding community was selected (D'hoe et al., 2018). These three species co-culture including representative gut microbial species *Roseburia intestinalis* L1-82, *Faecalibacterium prausnitzii* A2-165 and *Blautia hydrogenotrophica* S5a33 (D'hoe et al., 2018). All three species are capable of using fructose as the main carbon source. *B. hydrogenotrophica* is an acetate producer and formate consumer, while *R. intestinalis* and *F prausnitzii* are acetate consumers. All three species sequences are available on NCBI, the sequences numbers are NZ_GG657710.1, NZ_GG697168.2 NZ_LR027880.1. Three draft GEMs were constructed by Carveme (Machado et al., 2018). The *B. hydrogenotrophica* GEM contains 1719 reactions and 1195 metabolites; the *F prausnitzii* GEM contains 1358 reactions and 989 metabolites; the *R. intestinalis* GEM contains 1724 reactions and 1215 metabolites. Because the three species have different substrate usage, the production yields were generated based on different substrates and different rates. All the yield calculations and pathways selection considered fructose as the carbon source, and acetate or formate as an alternative carbon source accordingly. After analyzing the convex hull yield, we finally selected 4, 5, and 5 active pathways for *B. hydrogenotrophica, R. intestinalis,* and *F prausnitzii* separately.

Because of the complexity of the experimental data and the differences in fermentation phenotypes of different strains, we did not get parameters to make the model perform as well as the previous cases. Therefore, we try to split the experimental data and use different combinations of experimental data to narrow down the parameters to improve the model performance. We tried to use only mono-culture data (Fig. 7, Table S8), only two-species co-culture data (Fig. S1), and both mono-culture and two-species co-culture data (Fig. S2) to estimate parameters separately and check the performance. First, the model parameters were estimated using mono-culture experimental data, and then applied to check two species and three species co-cultures (Table S8). For mono-cultures, shown in Fig. 7 a-c, some biomass and metabolites concentration agree with experimental data, but some mismatches were identified for *F prausnitzii* mono-culture (Fig. 7 b) where the acetate concentration was not simulated well. And in *R. intestinalis,* and *F prausnitzii* (Fig. 7b and **c**) mono-culture, during the stationary and decline phases, the carbon source is still present in the

Table 2
Yield space and its perimeter and area.

GEMs	reactions	metabolites	EFM/EFV	Opt-Yield-FBA	Convex hull	Active path/final path
Receded *E. coli*	12	19	8/-	110	38	6
E. coli core	95	60	100,273/95,106	110	37	5
*i*ML1515	2712	1192	–/–[a]	210	20 (99% convex hull)	8
Yeast *S.* cerevisiae	3989	2693	–/–[a]	88	17 (99%)	9
P. stipitis	3989	2693	–/–[a]	88	17 (99%)	10
K. marxianus	3989	2693	–/–[a]	88	17 (99%)	10
B. hydrogenotrophica	1719	1195	–/–[a]	132	14 (99%)	4
F prausnitzii	1358	989	–/–[a]	132	22 (99%)	5
R. intestinalis	1724	1215	–/–[a]	132	16 (99%)	5

[a] -: EFMs/EFVs cannot be calculated.

medium, but the number of cells no longer increases. In model results, there is no carbon source during the stationary and decline phases. The experimental growth medium data used in the original study (D'hoe et al., 2018) includes unknown carbon sources and that is one of the reasons why some metabolites are mismatched. For two species co-culture, shown in Fig. 7d–f, parts of biomass and metabolites concentration agree well with experimental data. The relationship between *R. intestinalis* and *F prausnitzii* is competition and their dual-culture simulation in Fig. 7d was better than the other two dual-culture in Fig. 7e and **f**. For three species co-culture, shown in Fig. 7 g-h, there are two different states are observed in the experiments, which are dominated by *R. intestinalis* and *B. hydrogenotrophica* (Fig. 7g) or by *F. prausnitzii* and *B. hydrogenotrophica* (Fig. 7h), respectively. That two different states may be due to dependency on the initial species abundance. Even though the model is not fitting quantitatively to all the experimental data, some species and metabolites are matching the experimental data or at least trends for the three species culture simulations such as mono-culture and *R. intestinalis* and *F prausnitzii* dual-culture. We tried to improve model performance by integrating different experimental data sets to estimate the parameters (Figs. S1–S2). However, the prediction for three species co-culture was not as good as expected. Although the model performance can be improved by increasing the number of paths and parameters, or even more radical fitting methods. We believe that there are some limitations or conditions that are important for cybernetic model performance. The first one is the need for well-quantified experimental data (especially for the substrate uptake) because both the assumptions and pathways selection steps of cybernetic modeling are based on yield. The experimental growth medium data from reference (D'hoe et al., 2018) includes unknown carbon sources that may affect our model performance. The second limitation is that the model method is not sensitive to initial species abundance and cannot have two different results for three species co-culture. As shown in Fig. 7g and **h**, two different states are obtained and our model only simulated one state and failed to obtain a good fit. Since the initial abundances are very low for the three species, the slight change in the initial species abundance did not result in a significant change in the model results. The last point that might improve the performance of the model is the need to adjust the weights according to the different systems when estimating the parameters. For example in this system, there are four metabolites and biomass for three species and eight 8 scenarios to simulate (Fig. 7 a-h), the weight of acetate can be increased because it is a key metabolite for this cross-feeding system. Moreover, we believe that interspecies effects may be also an important factor affecting the performance of the model, which is also a reason the model of the three-bacterial cross-feeding system did not perform as well as the previous three-yeast model.

4. Discussions

The concept of yield is deeply rooted in metabolic engineering, and the optimization difference between rate and yield had been clearly articulated by Klamt and Schuster (Klamt et al., 2018; Schuster et al., 2008). When the growth environment is stressed and resources are constrained, the survival goal for an organism may be is how to use the substrate effectively, rather than maximizing the growth rate. This assumption is consistent with the hypothesis underlying cybernetic models, which regulates the allocation of resources to achieve the maximum return on investment. In industrial production, the yield requirements of strains are sometimes more important than the rate requirements, because the production rate can be compensated by increasing fermentation facilities, however, the yield cannot be compensated. Especially when the fermentation substrate is expensive, the cost of the substrate will be an essential economic factor. In modeling part, with the improvement of modeling techniques and more information can be integrated into GEMs such as incorporated enzymatic constraints like GECKO (Sánchez et al., 2017; Yang et al., 2021). EFMs and EFVs face challenges when processing large-scale models with many inhomogeneous constraints, requiring extensive computational resources to perform a large number of pathways calculations. The opt-yield-FBA can be an alternative of EFMs in the HCM approach because it can obtain yield space using GEMs with lower computational demands. As shown in Table 2, opt-yield-FBA calculated 110 pathways for the *E. coli* core model, while EFMs and EFVs calculated ~100,000 pathways.

When computing yield space, a potential question is how to deal with multi products yield space at the same time, not only pairwise products yield space. The yield space for more than two products is a multi-

Table 3
Multi-dimensional (multi-products) yield space.

Relative Values[a] (Mean ± std)	EFVs volume[b]	opt-yield-FBA volume	EFMs volume	EFVs area[c]	opt-yield-FBA area	EFMs area
2-D biomass-related	1	0.999 ± 0.001	1.064 ± 0.024	1	1.000 ± 0.000	1.009 ± 0.019
2-D All	1	0.957 ± 0.081	1.025 ± 0.032	1	0.997 ± 0.006	1.007 ± 0.012
3-D	1	0.898 ± 0.109	1.036 ± 0.040	1	0.930 ± 0.078	1.006 ± 0.006
4-D	1	0.831 ± 0.113	1.049 ± 0.040	1	0.841 ± 0.101	1.005 ± 0.003
All	1	0.891 ± 0.115	1.037 ± 0.038	1	0.918 ± 0.101	1.006 ± 0.008

[a] Relative values: divided by EFV values.
[b] Volume: volume of the convex hull when input dimension >2. When the space is 2-dimensional, this is the area of the convex hull.
[c] Area: surface area of the convex hull when input dimension >2. When the space is 2-dimensional, this is the perimeter of the convex hull.

dimensional polygon and vector in geometric and algebraic separately. We compared the properties of multi-dimensional yield space from EFMs, EFVs and opt-yield-FBA for a reduced *E. coli* core model and the EFVs' yield space is still considered the theoretically correct result (Table 3 and Table S9). For biomass-related 2-dimensional yield space (one dimension is biomass yield the other dimension is other product yield), the opt-yield-FBA convex hull covered 99.9% volume and 100% area of the EFVs' convex hull. And for all multi-dimensional yield space, the opt-yield-FBA convex hull covered 89.1% volume and 91.8% area of EFVs' convex hull. The coverage could be improved by increasing the step number of calculations with smaller sampling intervals. Even though the opt-yield-FBA could provide most coverage (89.1%) of multi-dimensional yield space, how to improve the coverage or even provide a theoretically correct result like EFVs' still is promising, especially helpful for pathway selections.

A challenge for advancing HCM is the need for more precise substrate data because both yield analysis and cybernetic modeling rely on the analysis of substrate uptake. Especially for multi-species cross-feeding or human gut microbiota community simulation. As the cybernetic modeling is a parametric modulation model, the balance of accuracy and over-fitting should also be considered. Some random sampling methods, such as the Bayesian-based method (Machado et al., 2018), may be an alternative for parameter estimation in cybernetic modeling because the result will be a normally distributed range rather than taking a specific value. The distribution of different combinations of parameters for the sampling estimation is also a direction for further study.

In conclusion we present an optimized yield analysis algorithm (opt-yield-FBA) and a HCM framework to simulate the metabolic dynamics at the genome-scale without the need for model reduction and EFMs calculation. The opt-yield-FBA can search for the maximal yield by solving a series of linear programming problems. The opt-yield-FBA algorithm could be implemented by the flux balance analysis (FBA) framework and is compatible with most current modeling tools. With the introduction of opt-yield-FBA, HCM can directly get the yield spaces and avoid the bottleneck of EFMs, and it will therefore can possible to apply the HCM strategy on genome-scale metabolic networks.

Author contributions

Jens Nielsen: Conceptualization and supervision. Hao Luo: Methodology, software, and writing- Original draft preparation. Boyang Ji: Data curation, validation and visualization. Peishun Li: Formal analysis and visualization. All authors: Writing – Review, Editing.

Declaration of competing interest

The authors declare no competing financial interests.

Data availability

Data will be made available on request.

Acknowledgments

This work was funded by the Novo Nordisk Foundation (NNF15OC0016798 and NNF10CC1016517), the Knut and Alice Wallenberg Foundation and BioGaia AB.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ymben.2022.12.001.

References

Aboulmouna, L., Raja, R., Khanum, S., Gupta, S., Maurya, M.R., Grama, A., Subramaniam, S., Ramkrishna, D., 2020. Cybernetic modeling of biological processes in mammalian systems: cybernetic modeling of mammalian bioprocesses. Curr. Opin. Chem. Eng. https://doi.org/10.1016/j.coche.2020.100660.

Agrawal, A., Verschueren, R., Diamond, S., Boyd, S., 2018. A Rewriting System for Convex Optimization Problems, pp. 42–60. https://doi.org/10.1080/23307706.2017.1397554 5.

Ahamed, F., Song, H., Ho, Y.K., 2021. Modeling coordinated enzymatic control of saccharification and fermentation by Clostridium thermocellum during consolidated bioprocessing of cellulose. Biotechnol. Bioeng. bit. 27705 https://doi.org/10.1002/bit.27705.

Bordbar, A., Monk, J.M., King, Z.A., Palsson, B.O., 2014. Constraint-based models predict metabolic and associated cellular functions. Nat. Rev. Genet. 15, 107–120. https://doi.org/10.1038/nrg3643.

Chen, Y., Li, F., Nielsen, J., 2022. Genome-scale modeling of yeast metabolism: retrospectives and perspectives. FEMS Yeast Res. 22, 1–9. https://doi.org/10.1093/FEMSYR/FOAC003.

D'hoe, K., Vet, S., Faust, K., Moens, F., Falony, G., Gonze, D., Lloréns-Rico, V., Gelens, L., Danckaert, J., De Vuyst, L., Raes, J., 2018. Integrated culturing, modeling and transcriptomics uncovers complex interactions and emergent behavior in a three-species synthetic gut community. Elife 7. https://doi.org/10.7554/eLife.37090.

Dhurjati, P., Ramkrishna, D., Flickinger, M.C., Tsao, G.T., 1985. A cybernetic view of microbial growth: modeling of cells as optimal strategists. Biotechnol. Bioeng. 27, 1–9. https://doi.org/10.1002/BIT.260270102.

Ebrahim, A., Lerman, J.A., Palsson, B.O., Hyduke, D.R., 2013. COBRApy: COnstraints-based reconstruction and analysis for Python. BMC Syst. Biol. 7, 1–6. https://doi.org/10.1186/1752-0509-7-74/FIGURES/2.

Geng, J., Ji, B., Li, G., López-Isunza, F., Nielsen, J., 2021. CODY enables quantitatively spatiotemporal predictions on in vivo gut microbial variability induced by diet intervention. Proc. Natl. Acad. Sci. U.S.A. 118 https://doi.org/10.1073/PNAS.2019336118/-/DCSUPPLEMENTAL.

Geng, J., Song, H.-S., Yuan, J., Ramkrishna, D., 2012. On enhancing productivity of bioethanol with multiple species. Biotechnol. Bioeng. 109, 1508–1517. https://doi.org/10.1002/bit.24419.

Gu, C., Kim, G.B., Kim, W.J., Kim, H.U., Lee, S.Y., 2019. Current status and applications of genome-scale metabolic models. Genome Biol. 20, 1–18. https://doi.org/10.1186/S13059-019-1730-3.

Henry, C.S., Dejongh, M., Best, A.A., Frybarger, P.M., Linsay, B., Stevens, R.L., 2010. High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat. Biotechnol. 28, 977–982. https://doi.org/10.1038/nbt.1672.

Kim, J. Il, Varner, J.D., Ramkrishna, D., 2008. A hybrid model of anaerobic *E. coli* GJT001: combination of elementary flux modes and cybernetic variables. Biotechnol. Prog. 24, 993–1006. https://doi.org/10.1002/btpr.73.

Klamt, S., Müller, S., Regensburger, G., Zanghellini, J., 2018. A mathematical framework for yield (vs. rate) optimization in constraint-based modeling and applications in metabolic engineering. Metab. Eng. 47, 153–169results. https://doi.org/10.1016/j.ymben.2018.02.001.

Klamt, S., Regensburger, G., Gerstl, M.P., Jungreuthmayer, C., Schuster, S., Mahadevan, R., Zanghellini, J., Müller, S., 2017. From elementary flux modes to elementary flux vectors: metabolic pathway analysis with arbitrary linear flux constraints. PLoS Comput. Biol. 13, e1005409 https://doi.org/10.1371/journal.pcbi.1005409.

Klamt, S., Saez-Rodriguez, J., Gilles, E.D., 2007. Structural and functional analysis of cellular networks with CellNetAnalyzer. BMC Syst. Biol. 1 https://doi.org/10.1186/1752-0509-1-2.

Kompala, D.S., Ramkrishna, D., Jansen, N.B., Tsao, G.T., 1986. Investigation of bacterial growth on mixed substrates: experimental evaluation of cybernetic models. Biotechnol. Bioeng. 28, 1044–1055. https://doi.org/10.1002/bit.260280715.

Lieven, C., Beber, M.E., Olivier, B.G., Bergmann, F.T., Ataman, M., Babaei, P., Bartell, J. A., Blank, L.M., Chauhan, S., Correia, K., Diener, C., Dräger, A., Ebert, B.E., Edirisinghe, J.N., Faria, J.P., Feist, A.M., Fengos, G., Fleming, R.M.T., García-Jiménez, B., Hatzimanikatis, V., van Helvoirt, W., Henry, C.S., Hermjakob, H., Herrgård, M.J., Kaafarani, A., Kim, H.U., King, Z., Klamt, S., Klipp, E., Koehorst, J.J., König, M., Lakshmanan, M., Lee, D.Y., Lee, S.Y., Lee, S., Lewis, N.E., Liu, F., Ma, H., Machado, D., Mahadevan, R., Maia, P., Mardinoglu, A., Medlock, G.L., Monk, J.M., Nielsen, J., Nielsen, L.K., Nogales, J., Nookaew, I., Palsson, B.O., Papin, J.A., Patil, K. R., Poolman, M., Price, N.D., Resendis-Antonio, O., Richelle, A., Rocha, I., Sánchez, B.J., Schaap, P.J., Malik Sheriff, R.S., Shoaie, S., Sonnenschein, N., Teusink, B., Vilaça, P., Vik, J.O., Wodke, J.A.H., Xavier, J.C., Yuan, Q., Zakhartsev, M., Zhang, C., 2020. MEMOTE for standardized genome-scale metabolic model testing. Nat. Biotechnol. 38, 272–276. https://doi.org/10.1038/s41587-020-0446-y.

Lu, H., Li, F., Sánchez, B.J., Zhu, Z., Li, G., Domenzain, I., Marcišauskas, S., Anton, P.M., Lappa, D., Lieven, C., Beber, M.E., Sonnenschein, N., Kerkhoven, E.J., Nielsen, J., 2019. A consensus S. cerevisiae metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. Nat. Commun. 10, 1–13. https://doi.org/10.1038/s41467-019-11581-3.

Machado, D., Andrejev, S., Tramontano, M., Patil, K.R., 2018. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. Nucleic Acids Res. 46, 7542–7553. https://doi.org/10.1093/NAR/GKY537.

Martínez, J.A., Bulté, D.B., Contreras, M.A., Palomares, L.A., Ramírez, O.T., 2020. Dynamic modeling of CHO cell metabolism using the hybrid cybernetic approach

with a novel elementary mode analysis strategy. Front. Bioeng. Biotechnol. 8, 279. https://doi.org/10.3389/fbioe.2020.00279.
Monk, J.M., Lloyd, C.J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., Feist, A.M., Palsson, B.O., 2017. iML1515, a knowledgebase that computes Escherichia coli traits. Nat. Biotechnol. 35, 904–908. https://doi.org/10.1038/nbt.3956.
Müller, S., Regensburger, G., 2016. Elementary vectors and conformal sums in polyhedral geometry and their relevance for metabolic pathway analysis. Front. Genet. 7, 90. https://doi.org/10.3389/FGENE.2016.00090/BIBTEX.
Nielsen, J., Keasling, J.D., 2016. Engineering cellular metabolism. Cell 164, 1185–1197. https://doi.org/10.1016/J.CELL.2016.02.004.
Orth, J.D., Conrad, T.M., Na, J., Lerman, J.A., Nam, H., Feist, A.M., Palsson, B., 2011. A comprehensive genome-scale reconstruction of Escherichia coli metabolism—2011. Mol. Syst. Biol. 7, 535. https://doi.org/10.1038/MSB.2011.65.
Orth, J.D., Fleming, R.M.T., Palsson, B.Ø., 2010a. Reconstruction and use of microbial metabolic networks: the core Escherichia coli metabolic model as an educational guide. EcoSal Plus 4. https://doi.org/10.1128/ECOSALPLUS.10.2.1/ASSET/AF2F042E-12AF-48A1-967D-9627A8D81014/ASSETS/GRAPHIC/10.2.1_FIG_022.GIF.
Orth, J.D., Thiele, I., Palsson, B.O., 2010b. What is flux balance analysis? Nat. Biotechnol. https://doi.org/10.1038/nbt.1614.
Perrin, E., Ghini, V., Giovannini, M., Di Patti, F., Cardazzo, B., Carraro, L., Fagorzi, C., Turano, P., Fani, R., Fondi, M., 2020. Diauxie and co-utilization of carbon sources can coexist during bacterial growth in nutritionally complex environments. Nat. Commun. 11, 1–16. https://doi.org/10.1038/s41467-020-16872-8.
Ramkrishna, D., Song, H.-S., 2012. Dynamic models of metabolism: Review of the cybernetic approach. AIChE J. 58, 986–997. https://doi.org/10.1002/aic.13734.
Rouhollah, H., Iraj, N., Giti, E., Sorah, A., 2007. Mixed sugar fermentation by Pichia stipitis, Sacharomyces cerevisiaea, and an isolated xylosefermenting Kluyveromyces marxianus and their cocultures. Afr. J. Biotechnol. 6.
Sánchez, B.J., Zhang, C., Nilsson, A., Lahtvee, P.-J., Kerkhoven, E.J., Nielsen, J., 2017. Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. Mol. Syst. Biol. 13, 935. https://doi.org/10.15252/MSB.20167411.
Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M., Sauer, U., 2012. Multidimensional optimality of microbial metabolism. Science 84 336, 601–604. https://doi.org/10.1126/SCIENCE.1216882/SUPPL_FILE/SCHUETZ.SM.PDF.
Schuster, S., Dandekar, T., Fell, D.A., 1999. Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. Trends Biotechnol. 17, 53–60. https://doi.org/10.1016/s0167-7799(98)01290-6.
Schuster, S., Fell, D.A., Dandekar, T., 2000. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nat. Biotechnol. 18, 326–332. https://doi.org/10.1038/73786.
Schuster, S., Pfeiffer, T., Fell, D.A., 2008. Is maximization of molar yield in metabolic networks favoured by evolution? J. Theor. Biol. 252, 497–504. https://doi.org/10.1016/J.JTBI.2007.12.008.
Song, H.-S., Morgan, J.A., Ramkrishna, D., 2009. Systematic development of hybrid cybernetic models: application to recombinant yeast co-consuming glucose and xylose. Biotechnol. Bioeng. 103, 984–1002. https://doi.org/10.1002/bit.22332.
Song, H.-S., Ramkrishna, D., 2011. Cybernetic models based on lumped elementary modes accurately predict strain-specific metabolic function. Biotechnol. Bioeng. 108, 127–140. https://doi.org/10.1002/bit.22922.
Song, H.-S., Ramkrishna, D., 2010. Prediction of metabolic function from limited data: lumped hybrid cybernetic modeling (L-HCM). Biotechnol. Bioeng. 106 https://doi.org/10.1002/bit.22692 n/a-n/a.
Song, H.-S., Ramkrishna, D., 2009. Reduction of a set of elementary modes using yield analysis. Biotechnol. Bioeng. 102, 554–568. https://doi.org/10.1002/bit.22062.
Straight, J.V., Ramkrishna, D., 1994. Cybernetic modeling and regulation of metabolic pathways. Growth on Complementary Nutrients. Biotechnol. Prog. 10, 574–587. https://doi.org/10.1021/bp00030a002.
Terzer, M., Stelling, J., 2008. Large-scale computation of elementary flux modes with bit pattern trees. Bioinformatics 24, 2229–2235. https://doi.org/10.1093/bioinformatics/btn401.
Teusink, B., Smid, E.J., 2006. Modelling strategies for the industrial exploitation of lactic acid bacteria. Nat. Rev. Microbiol. 4, 46–56. https://doi.org/10.1038/nrmicro1319.
Varma, Amit, Palsson, B.O., 1994a. Metabolic flux balancing: basic concepts, scientific and practical use. Bio Technol. 12, 994–998. https://doi.org/10.1038/nbt1094-994.
Varma, A., Palsson, B.O., 1994b. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. Appl. Environ. Microbiol. 60, 3724. https://doi.org/10.1128/AEM.60.10.3724-3731.1994.
Vilkhovoy, M., Minot, M., Varner, J.D., 2016. Effective dynamic models of metabolic networks. IEEE Life Sci. Lett. 2, 51–54. https://doi.org/10.1109/lls.2016.2644649.
Yang, X., Mao, Z., Zhao, X., Wang, R., Zhang, P., Cai, J., Xue, C., Ma, H., 2021. Integrating thermodynamic and enzymatic constraints into genome-scale metabolic models. Metab. Eng. 67, 133–144. https://doi.org/10.1016/J.YMBEN.2021.06.005.
Young, J.D., 2015. Learning from the steersman: a natural history of cybernetic models. Ind. Eng. Chem. Res. https://doi.org/10.1021/acs.iecr.5b01315.
Young, J.D., Ramkrishna, D., 2007. On the matching and proportional laws of cybernetic models. Biotechnol. Prog. 23, 83–99. https://doi.org/10.1021/bp060176q.