Dynamic optimization of metabolic networks coupled with gene expression

Alexander Bockmayr

Freie Universität Berlin

Joint work with Steffen Waldherr and Diego A. Oyarzún

DFG Research Center MATHEON
Mathematics for key technologies

Metabolic networks, FU Berlin, SS 2014

Outline

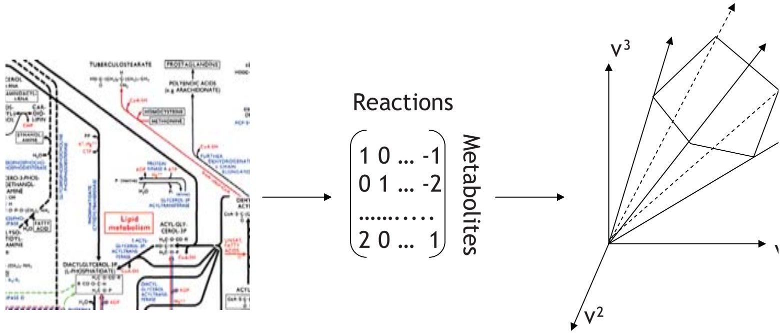
- ▷ Constraint-based modeling of metabolic networks
- ▷ Metabolism and enzyme production
- ▷ Dynamic model
- ▷ Time-scale separation
- ▷ Steady-state and dynamic optimisation: FBA, dFBA, RBA
- ▷ Dynamic enzyme-cost FBA: deFBA
- ▷ Applications

Constraint-based modeling

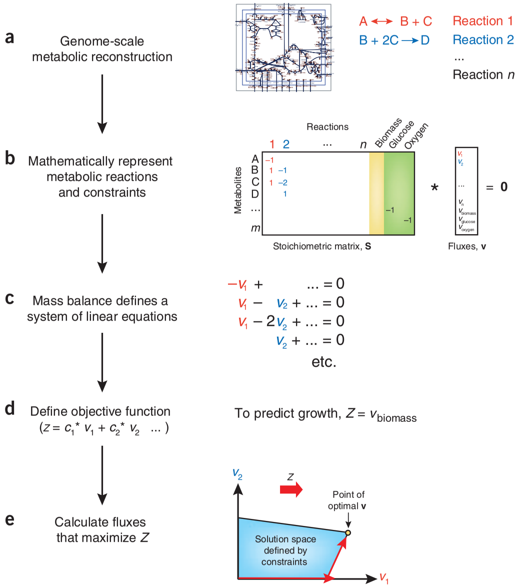
Set of possible behaviors of a metabolic network at steady state

$$C = \{v \in \mathbb{R}^n \mid Sv = 0,\ v_i \geq 0,\ i \in Irr\}$$

$\rightsquigarrow$ steady-state flux cone



Flux balance analysis (FBA)


a Genome-scale metabolic reconstruction

b Mathematically represent metabolic reactions and constraints

c Mass balance defines a system of linear equations

d Define objective function ($z = c_1^* v_1 + c_2^* v_2 \ \dots$)

e Calculate fluxes that maximize Z

- ▷ *E. coli* metabolism
- ▷ Genome-scale reconstruction (*i*JO1366)
- ▷ 1336 metabolites, 2251 reactions
- ▷ Objective function: biomass
- ▷ Glucose and oxygen uptake reactions
- ▷ Aerobic and anaerobic growth
- ▷ Software: e.g. COBRA Toolbox 2.0

Orth/Thiele/Palsson 10

Metabolism and enzyme production

$$\begin{array}{ccc} Y \xleftrightarrow{V_y} X, & X \xrightarrow{V_x} \tilde{X}, & X \xrightarrow{V_p} P \\ \uparrow P & \uparrow P & \uparrow P \end{array}$$

- ▷ Molecular species
  - Extracellular nutrients and/or waste Y
  - Intracellular metabolites X
  - Macromolecules/enzymes P
- ▷ Reaction fluxes
  - Exchange reactions V_y
  - Internal metabolic reactions V_x
  - Biomass reactions V_p
- ▷ Stoichiometric matrices S_j^i (species i, reactions j)

Dynamic model

- ▷ Mass balance

$$\begin{aligned} \dot{Y} &= -S_y^y V_y \\ \dot{P} &= S_p^p V_p \\ \dot{X} &= S_y^x V_y + S_x^x V_x - S_p^x V_p \end{aligned}$$

- ▷ Macromolecule production is slow: small ε
- ▷ Macromolecules are made from many components: large α

$$\begin{aligned} \dot{Y} &= -S_y^y V_y \\ \dot{P} &= \varepsilon\, S_p^p V_p \\ \dot{X} &= S_y^x V_y + S_x^x V_x - \varepsilon \alpha\, S_p^x V_p \end{aligned}$$

Quasi steady state approximation

- ▷ Time-scale separation (using Tikhonov's theorem)

$$\begin{aligned} \dot{Y} &= -S_y^y V_y, \\ \dot{P} &= \varepsilon\, S_p^p V_p, \\ 0 &= S_y^x V_y + S_x^x V_x - \alpha\, \varepsilon\, S_p^x V_p. \end{aligned}$$

- ▷ Exchange reactions and biomass production coupled via quasi steady-state constraint for intracellular metabolism.
- ▷ Model reduction

Steady-state vs. dynamic optimisation

	Metabolism	+ Enzyme production
Steady state	Flux Balance Analysis (FBA) Varma/Palsson 94	Resource Balance Analysis (RBA) Goelzer et al. 11
Dynamic	Dynamic FBA (dFBA) Mahadevan et al. 02	Dynamic Enzyme Cost Analysis (deFBA) – THIS TALK –

Flux balance analysis (FBA)

- Goal: Determine metabolic fluxes maximizing a cellular objective.
- Network model: Metabolism & steady state

$$0 = S_y v_y + S_x v_x - S_{bm} v_{bm} \tag{1}$$

- Constraints: Bounds on fluxes

$$v_{i,min} \le v_i \le v_{i,max} \tag{2}$$

- Optimisation: Linear programming (LP)

$$\max_{v} v_{bm} \;\text{ such that }\; (1) \text{ and } (2)$$

Dynamic flux balance analysis (dFBA)

- Goal: Determine metabolic fluxes maximizing a cellular objective over time (based on biomass concentration $P(t)$).
- Network model: Metabolism & dynamic

$$\begin{aligned}\dot{y}(t) &= -S^y_y v_y(t)\, P(t)\\ \dot{x}(t) &= S^x_y v_y(t)\, P(t) + S^x_x v_x(t)\, P(t) - S^x_{bm} v_{bm}(t)\, P(t)\\ \dot{P}(t) &= v_{bm}(t)\, P(t)\end{aligned} \tag{3}$$

- Constraints: Bounds on fluxes & flux changes

$$v_{i,min}(y,P) \le v_i(t) \le v_{i,max}(y,P), \qquad |\dot{v}_i(t)| \le \dot{v}_{i,max} \tag{4}$$

- Optimisation: Non-linear dynamic

$$\max_{v(t)} P(t_{end}) \;\text{ such that }\; (3) \text{ and } (4)$$

Resource balance analysis (RBA)

- Goal: Determine cell composition (protein concentrations p) and metabolic fluxes v maximizing the growth rate μ.
- Network model: Metabolism+enzyme production & steady st.

$$\begin{aligned}0 &= S^x_y v_y + S^x_x v_x - \alpha\varepsilon S^x_p v_p\\ 0 &= \varepsilon S^p_p v_p - \mu p\end{aligned} \tag{5}$$

- Constraints: Enzyme capacity & cellular composition

$$\sum_{j \in \mathscr{V}_i} |v_j(t)/k_j| \le p_i, \qquad \sum_i c_i p_i \le 1 \tag{6}$$

- Optimization: Iteratively solving LPs

$$\max_{v,p} \mu \;\text{ such that }\; (5) \text{ and } (6)$$

Dynamic enzyme-cost FBA (deFBA)

- Goal: Determine the dynamic cell composition and metabolic fluxes to maximize a cellular objective over a time interval
- Network model: Metabolism+enzyme production & dynamic

$$\begin{aligned}\dot{Y} = -S^y_y V_y, \qquad \dot{P} &= \varepsilon\, S^p_p V_p,\\ 0 &= S^x_y V_y + S^x_x V_x - \alpha\, \varepsilon\, S^x_p V_p\end{aligned} \tag{7}$$

- Constraints: Enzyme capacity & cellular composition

$$\sum_{j \in \mathscr{V}_i} |V_j(t)/k_j| \le P_i(t), \qquad \sum_i c_i P_i(t) \le 1 \tag{8}$$

- Optimization: Linear dynamic

$$\max_{V,Y,P} \int_0^{t_{end}} c^{\mathrm{T}} P(t) dt \;\text{ such that }\; (7) \text{ and } (8)$$

Constraints

- Enzyme capacity constraints

$$\left|\frac{v_1}{c_1}\right| + \cdots + \left|\frac{v_m}{c_m}\right| \;\leq\; P_E$$

- Biomass-independent flux bounds

$$v_{min} \leq v \leq v_{max}$$

- Non-negativity of molecular species

$$Y \geq 0, \quad P \geq 0$$

- Biomass composition constraints

$$H_B P \leq h_B$$

Dynamic optimization problem

Let $z = (Y, P)$.

$$\max_{\mathscr{V}(\mathscr{Z}, z_0)} \int_0^{t_{end}} \Phi(z(t), v(t))dt \;+\; \Psi(z(t_{end}))$$

$$\begin{aligned}
\text{s.t. } \dot{Y} &= -S_y^y V_y, \quad \dot{P} = \varepsilon\, S_p^p V_p,\\
& S_y^x V_y + S_x^x V_x - \alpha\,\varepsilon\, S_p^x V_p = 0,\\
& z(0) = z_0,\\
& z(t) \geq 0,\\
& v_{min} \leq v(t) \leq v_{max},\\
& H_C v(t) \leq H_E P(t), \quad H_B P(t) \leq h_B.
\end{aligned}$$

Minimal example

- Components: Nutrient Y, metabolite X, generic enzyme P
- Reactions

$$V_y : Y \rightarrow X \;\text{ (uptake)}, \quad V_p : \alpha X \rightarrow P \;\text{ (biomass)}$$

- Enzymatic constraint

$$\frac{V_y}{k_y} + \frac{\varepsilon V_p}{k_p} \leq P$$

- Approximate model (on long time-scales)

$$\dot{Y} = -V_y, \quad \dot{P} = \varepsilon\, V_p, \qquad V_y = \alpha\,\varepsilon\, V_p$$

Objective functionals

- Maximization of terminal biomass

$$J_1 = P(t_{end})$$

- Maximization of discounted biomass integral

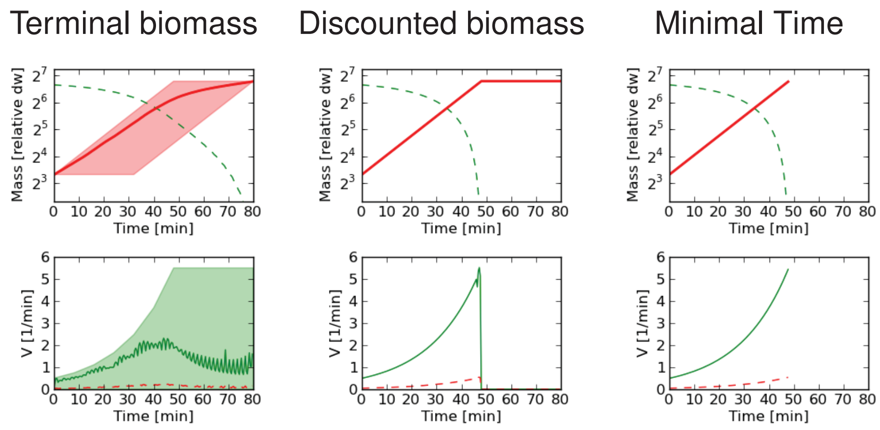
$$J_2 = \int_0^{t_{end}} P(\tau) e^{-\mu\tau} d\tau$$

- Minimization of time to consume nutrients

$$J_3 = -\int_0^{t_{end}} d\tau = -t_{end}$$

with $Y(t_{end}) = 0$.

Optimization results for minimal network



Analytical proof of existence and uniqueness for J_2 and J_3
⇝ mathematical optimum biologically meaningful?

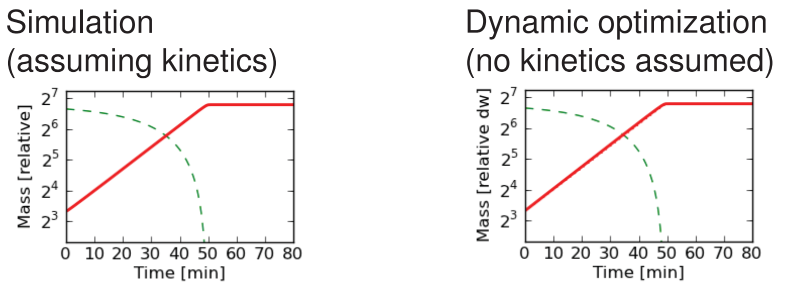
Minimal network yields Monod model

- ▷ Assume Michaelis-Menten kinetics

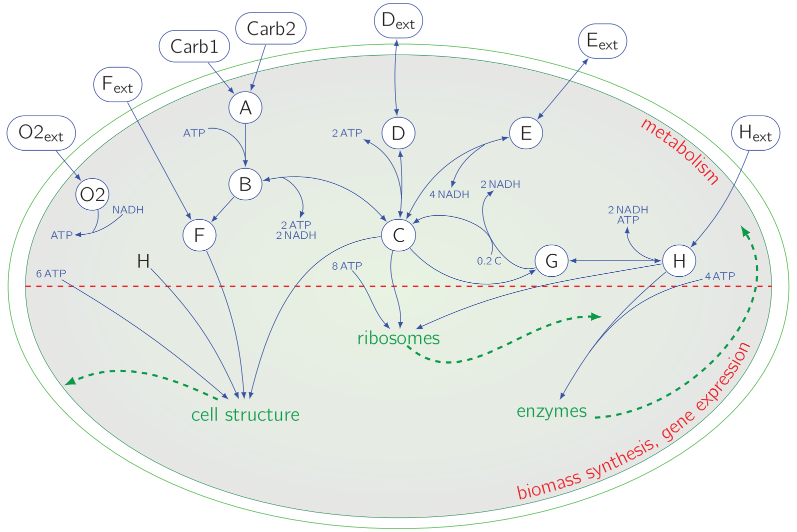
$$V_y = \frac{V_{m,y} P Y}{K_y \vartheta_e + Y}, \qquad \varepsilon V_p = \frac{V_{m,p} P X}{K_p + X}.$$

- ▷ Choose consistent parameters

$$V_{m,y} = \left(\frac{1}{k_y} + \frac{1}{\alpha k_p}\right)^{-1}, \qquad V_{m,p} = \left(\frac{\alpha}{k_y} + \frac{1}{k_p}\right)^{-1}.$$



Core cellular network



Covert et al. 2001 (for the metabolic part)

Network specification

Reaction	Enz	k_{cat}
Exchange reactions		
Carb1 → A	T_{C1}	$3000\ \frac{1}{\min}$
Carb2 → A	T_{C2}	$2000\ \frac{1}{\min}$
$F_{ext} \to F$	T_F	$3000\ \frac{1}{\min}$
$O2_{ext} \to O2$	S	$1000\ \frac{1}{\min}$
$D \leftrightarrow D_{ext}$	S	$1000\ \frac{1}{\min}$
$E \leftrightarrow E_{ext}$	S	$1000\ \frac{1}{\min}$
$H_{ext} \to A$	T_H	$3000\ \frac{1}{\min}$
Metabolic reactions		
$A + ATP \to B$	E_B	$1800\ \frac{1}{\min}$
$B \to C + 2\,ATP + 2\,NADH$	E_C	$1800\ \frac{1}{\min}$
$B \to F$	E_F	$1800\ \frac{1}{\min}$
$C \to G$	E_G	$1800\ \frac{1}{\min}$
$G \to 0.8\,C + 2\,NADH$	E_N	$1800\ \frac{1}{\min}$
$C \leftrightarrow 2\,ATP + 3\,D$	E_D	$1800\ \frac{1}{\min}$
$C + 4\,NADH \leftrightarrow 3\,E$	E_E	$1800\ \frac{1}{\min}$
$G + ATP + 2\,NADH \leftrightarrow H$	E_H	$1800\ \frac{1}{\min}$
$NADH + O \to ATP$	E_T	$1800\ \frac{1}{\min}$

Reaction	Enz	k_{cat}
Biomass reactions		
$400\,H + 1600\,ATP \to T_{C1}$	R	$2.5\ \frac{1}{\min}$
$1500\,H + 6000\,ATP \to T_{C2}$	R	$0.67\ \frac{1}{\min}$
$400\,H + 1600\,ATP \to T_F$	R	$2.5\ \frac{1}{\min}$
$400\,H + 1600\,ATP \to T_O$	R	$2.5\ \frac{1}{\min}$
$400\,H + 1600\,ATP \to T_D$	R	$2.5\ \frac{1}{\min}$
$400\,H + 1600\,ATP \to T_E$	R	$2.5\ \frac{1}{\min}$
$400\,H + 1600\,ATP \to T_H$	R	$2.5\ \frac{1}{\min}$
$500\,H + 2000\,ATP \to E_B$	R	$2\ \frac{1}{\min}$
$500\,H + 2000\,ATP \to E_C$	R	$2\ \frac{1}{\min}$
$1000\,H + 4000\,ATP \to E_D$	R	$1\ \frac{1}{\min}$
$1000\,H + 4000\,ATP \to E_E$	R	$1\ \frac{1}{\min}$
$1500\,H + 6000\,ATP \to E_F$	R	$0.67\ \frac{1}{\min}$
$500\,H + 2000\,ATP \to E_G$	R	$2\ \frac{1}{\min}$
$2500\,H + 10000\,ATP \to E_H$	R	$0.4\ \frac{1}{\min}$
$500\,H + 2000\,ATP \to E_N$	R	$2\ \frac{1}{\min}$
$500\,H + 2000\,ATP \to E_T$	R	$2\ \frac{1}{\min}$
$2000\,H + 4000\,C + 16000\,ATP \to R$	R	$0.2\ \frac{1}{\min}$
$250\,H + 250\,C + 250\,F + 1500\,ATP \to S$	R	$3\ \frac{1}{\min}$

⇝ preferred carbon source Carb1

Case study: Carbon switch

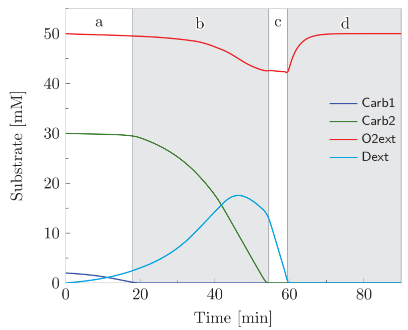
Growth scenario:
- Low amount of preferred carbon source C_1
- High amount of non-preferred carbon source C_2
- Ample oxygen supply

Objective: Discounted biomass

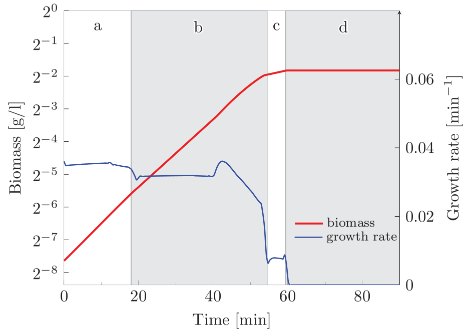
$$J = \int_0^{t_{end}} c_{bm}^{\mathrm{T}} P(t)\, e^{-\mu t} dt$$

Dynamic optimization results I

Substrates



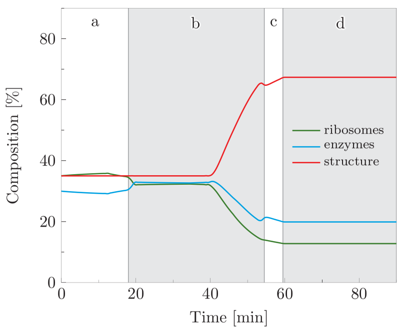
Biomass & Growth



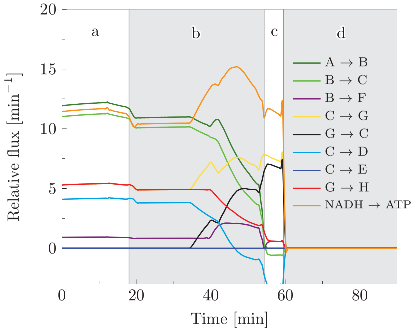
Four cellular growth phases

Dynamic optimization results II

Cell composition



Metabolic fluxes



Cellular reorganisation at the end of phase (b)

Discussion

Modeling metabolism including enzyme costs
- Mass balance ODE model for cellular metabolism and biomass production.
- Time-scale separation yields dynamic biomass model with quasi steady-state metabolic constraints.
- Dynamic optimisation framework: deFBA

Case studies
- Minimal network yields Monod growth model.
- Core network shows different exponential growth phases and pre-adaptation to impending nutrient depletion.
- Choice of the objective functional crucial for the results.