Dynamic optimization of metabolic networks coupled with gene expression

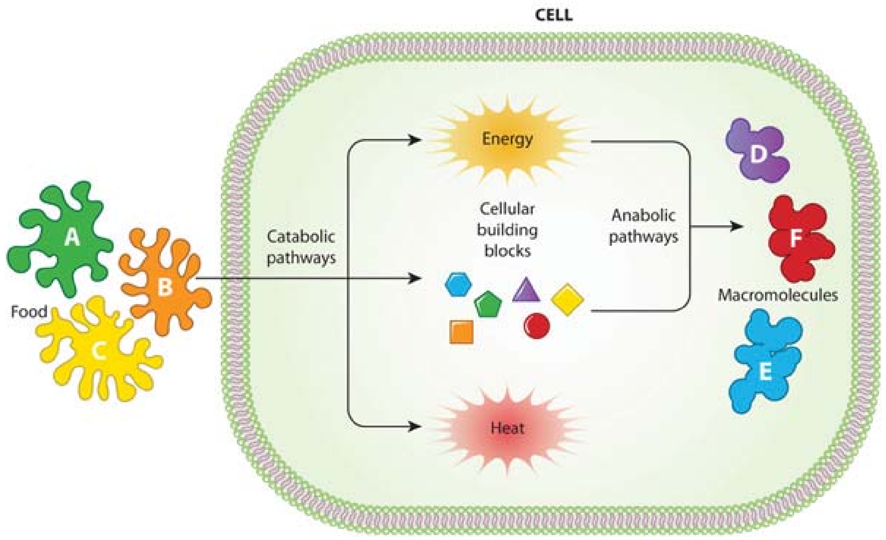
Alexander Bockmayr
Freie Universität Berlin

Joint work with Steffen Waldherr and Diego A. Oyarzún

Research Center MATHEON
Mathematics for key technologies

Metabolic networks, FU Berlin, SS 2015

Cellular metabolism



© 2010 Nature Education

Metabolism and enzyme production

$$Y \underset{\uparrow P}{\xleftrightarrow{V_y}} X, \quad X \underset{\uparrow P}{\xrightarrow{V_x}} \tilde{X}, \quad X \underset{\uparrow P}{\xrightarrow{V_p}} P$$

- 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

Tikhonov's theorem

$$\begin{aligned}\dot{z} &= f(x,z,\varepsilon), \qquad (*)\\ \varepsilon\dot{x} &= g(x,z,\varepsilon),\\ (z(0),x(0)) &= (z_0,x_0),\end{aligned}$$

with f, g sufficiently smooth, under the following hypotheses:

(H1) there exists a unique solution $x = \bar{g}(z)$, sufficiently smooth, of $g(x,z,0) = 0$; the matrix $\partial g/\partial x(\bar{g}(z),z,0)$ has all eigenvalues with strictly negative real part;

(H2) the reduced system

$$\begin{aligned}\dot{z} &= f(\bar{g}(z),z,0), \qquad (**)\\ z(0) &= z_0,\end{aligned}$$

has a solution $z^0(t)$ on an interval $[0,T]$, for some $T > 0$;

(H3) x_0 is in the basin of attraction of the steady state $\bar{g}(z_0)$ of the fast system $\dot{\xi} = g(\xi, z_0, 0)$.

If these hypotheses are satisfied, system (*) admits a solution $(x^\varepsilon(t), z^\varepsilon(t))$ on $[0,T]$. In addition,

$$\lim_{\varepsilon\to 0+} z^\varepsilon(t) = z^0(t)$$

and

$$\lim_{\varepsilon\to 0+} x^\varepsilon(t) = x^0(t) = \bar{g}(z_0)(t),$$

uniformly on any closed interval contained in $(0,T]$.

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) – TODAY –

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

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

- 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

$$\begin{aligned} \dot{y}(t) &= -S_y^y v_y(t)\, P(t) \\ \dot{x}(t) &= S_y^x v_y(t)\, P(t) + S_x^x v_x(t)\, P(t) - S_{bm}^x v_{bm}(t)\, P(t) \\ \dot{P}(t) &= v_{bm}(t)\, P(t) \end{aligned} \tag{3}$$

- Constraints: Bounds on fluxes & flux changes

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

- 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.

$$\begin{aligned} 0 &= S_y^x v_y + S_x^x v_x - \alpha\varepsilon S_p^x v_p \\ 0 &= \varepsilon S_p^p v_p - \mu p \end{aligned} \tag{5}$$

- Constraints: Enzyme capacity & cellular composition

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

- 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

$$\begin{aligned} \dot{Y} = -S_y^y V_y, \qquad \dot{P} &= \varepsilon\, S_p^p V_p, \\ 0 = S_y^x V_y + S_x^x V_x &- \alpha\, \varepsilon\, S_p^x V_p \end{aligned} \tag{7}$$

- Constraints: Enzyme capacity & cellular composition

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

- 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

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

- Biomass-independent flux bounds

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

- Non-negativity of molecular species

$$Y \ge 0, \quad P \ge 0$$

- Biomass composition constraints

$$H_B P \le h_B$$

Dynamic optimization problem

Let $z = (Y, P)$.

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

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

Minimal example

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

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

- Enzymatic constraint

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

- 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

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

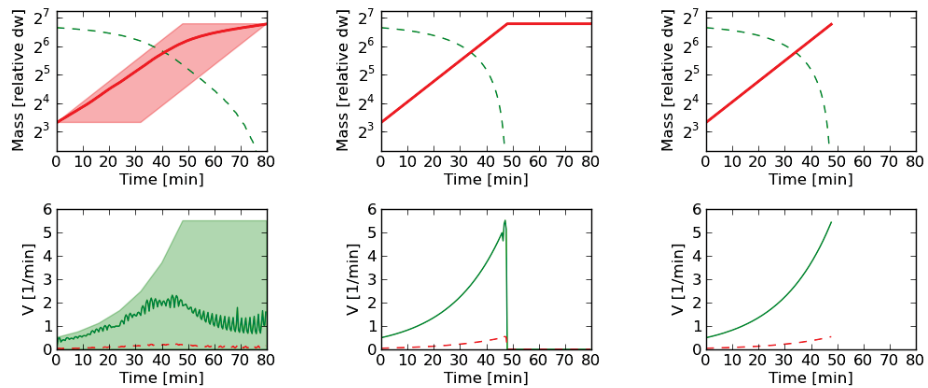
- Minimization of time to consume nutrients

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

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

Optimization results for minimal network

Terminal biomass | Discounted biomass | Minimal Time



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

Minimal network yields Monod model

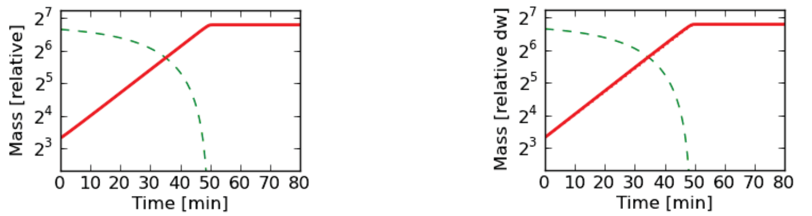
- Assume Michaelis-Menten kinetics

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

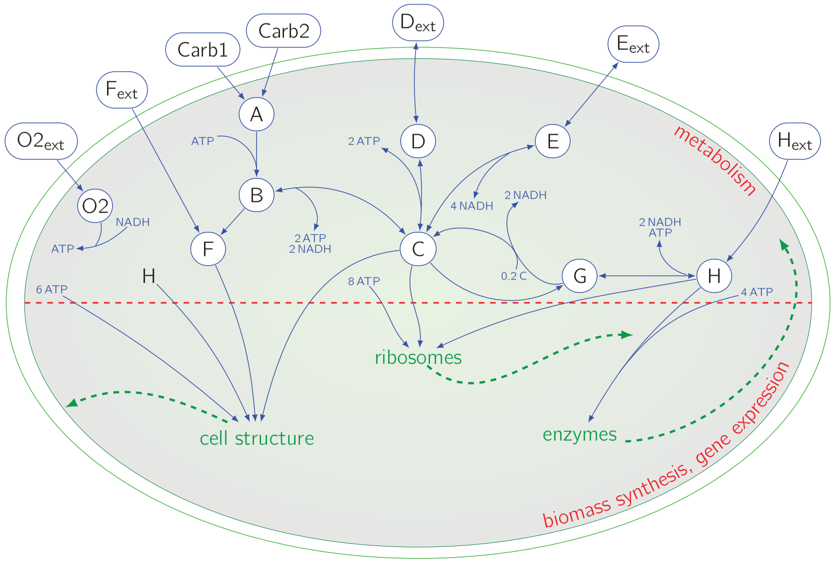
- Choose consistent parameters

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

Simulation (assuming kinetics) | Dynamic optimization (no kinetics assumed)



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} → F	T_F	$3000\ \frac{1}{min}$
$O2_{ext}$ → O2	S	$1000\ \frac{1}{min}$
D ↔ D_{ext}	S	$1000\ \frac{1}{min}$
E ↔ E_{ext}	S	$1000\ \frac{1}{min}$
H_{ext} → A	T_H	$3000\ \frac{1}{min}$
Metabolic reactions		
A + ATP → B	E_B	$1800\ \frac{1}{min}$
B → C + 2 ATP + 2 NADH	E_C	$1800\ \frac{1}{min}$
B → F	E_F	$1800\ \frac{1}{min}$
C → G	E_G	$1800\ \frac{1}{min}$
G → 0.8 C + 2 NADH	E_N	$1800\ \frac{1}{min}$
C ↔ 2 ATP + 3 D	E_D	$1800\ \frac{1}{min}$
C + 4 NADH ↔ 3 E	E_E	$1800\ \frac{1}{min}$
G + ATP + 2 NADH ↔ H	E_H	$1800\ \frac{1}{min}$
NADH + O → ATP	E_T	$1800\ \frac{1}{min}$

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

⇝ preferred carbon source Carb1

Case study: Carbon switch

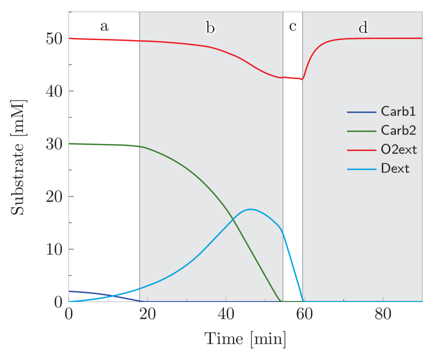
Growth scenario:
- Low amount of preferred carbon source C_1
- High amount of non-preferred carbon source C_2
- Ample oxygen supply

Objective: Discounted biomass

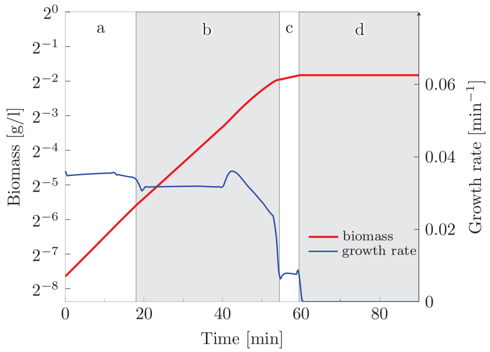
$$J = \int_0^{t_{end}} c_{bm}^{\mathrm{T}} P(t)\, e^{-\mu t} dt$$

Dynamic optimization results I

Substrates



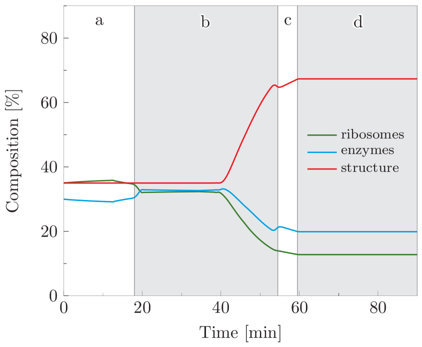
Biomass & Growth



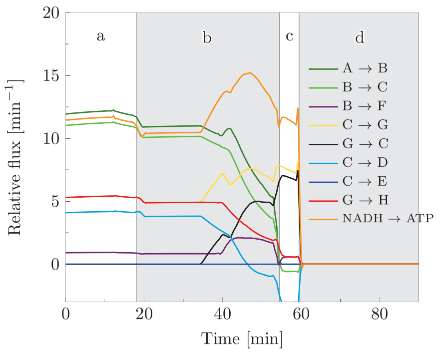
Four cellular growth phases

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.