12. Dynamic enzyme-cost FBA (deFBA)

Dynamic optimization of metabolic networks coupled with gene expression
Waldherr S, Oyarzún DA, Bockmayr A.
Journal of Theoretical Biology, 365, 469-485, 2015
http://dx.doi.org/10.1016/j.jtbi.2014.10.035

Metabolism and enzyme production

Y ←→ X,  X ←→ Y,  X ←→ 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_{ij} (species i, reactions j)

Cellular metabolism

Dynamic model

▷ Mass balance
  \[ \dot{Y} = -S_{yY} V_y \]
  \[ \dot{P} = S_{pP} V_p \]
  \[ \dot{X} = S_{xX} V_y + S_{xX} V_x - S_{xP} V_p \]

▷ Macromolecule production is slow: small \( \varepsilon \)
▷ Macromolecules are made from many components: large \( \alpha \)
Quasi steady state approximation

- Time-scale separation (using Tikhonov’s theorem)
  \[ \dot{Y} = -S^\varepsilon_y V_y, \]
  \[ \dot{P} = \varepsilon S^\varepsilon_p V_p, \]
  \[ 0 = S^\varepsilon_y V_y + S^\varepsilon_x V_x - \alpha \varepsilon S^\varepsilon_p V_p. \]
- Exchange reactions and biomass production coupled via quasi steady-state constraint for intracellular metabolism.
- Model reduction

\[ \begin{align*}
\dot{z} &= f(x, z, \varepsilon), \\
\varepsilon \dot{x} &= g(x, z, \varepsilon), \\
(z(0), x(0)) &= (z_0, x_0),
\end{align*} \]

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 \( \frac{\partial g}{\partial x}(\bar{g}(z), z, 0) \) has all eigenvalues with strictly negative real part;

(H2) the reduced system
\[ \begin{align*}
\dot{z} &= f(\bar{g}(z), z, 0), \\
z(0) &= z_0,
\end{align*} \]
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

<table>
<thead>
<tr>
<th></th>
<th>Metabolism</th>
<th>+ Enzyme production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steady state</td>
<td>Flux Balance Analysis (FBA)</td>
<td>Resource Balance Analysis (RBA)</td>
</tr>
<tr>
<td></td>
<td>Varma/Palsson 94</td>
<td>Goelzer et al. 11</td>
</tr>
<tr>
<td>Dynamic</td>
<td>Dynamic FBA (dFBA)</td>
<td>Dynamic Enzyme Cost Analysis (deFBA)</td>
</tr>
<tr>
<td></td>
<td>Mahadevan et al. 02</td>
<td>– TODAY –</td>
</tr>
</tbody>
</table>
 Flux balance analysis (FBA)

- **Goal**: Determine metabolic fluxes maximizing a cellular objective.
- **Network model**: Metabolism & steady state
  \[ 0 = S_y v_y + S_x v_x - S_{bm} v_{bm} \]  
  \[ 0 = S_y^p v_y + S_x^p v_x - \alpha \varepsilon S_x^p v_p \]  
- **Constraints**: Bounds on fluxes
  \[ v_{i,\text{min}} \leq v_i \leq v_{i,\text{max}} \]  
- **Optimisation**: Linear programming (LP)

  \[ \max_v v_{bm} \text{ such that (1) and (2)} \]

 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
  \[ \dot{y}(t) = -S_y v_y(t) P(t) \]  
  \[ \dot{x}(t) = S_y^p v_y(t) P(t) + S_x v_x(t) P(t) - S_{bm} v_{bm}(t) P(t) \]  
  \[ \dot{P}(t) = v_{bm}(t) P(t) \]  
- **Constraints**: Bounds on fluxes & flux changes
  \[ v_{i,\text{min}}(y,P) \leq v_i(t) \leq v_{i,\text{max}}(y,P), \quad |\dot{v}_i(t)| \leq \dot{v}_{i,\text{max}} \]  
- **Optimisation**: Non-linear dynamic

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

 Resource balance analysis (RBA)

- **Goal**: Determine cell composition (protein concentrations \( p \)) and metabolic fluxes \( v \) maximizing the growth rate \( \mu \).
- **Network model**: Metabolism + enzyme production & steady state
  \[ 0 = S_y v_y + S_x v_x - \alpha \varepsilon S_x v_p \]  
  \[ 0 = S_y^p v_y - \mu p \]  
- **Constraints**: Enzyme capacity & cellular composition
  \[ \sum_{j \in \gamma_i} |v_j(t)/k_j| \leq p_i, \quad \sum_i c_i p_i \leq 1 \]  
- **Optimisation**: Iteratively solving LPs

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

 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
  \[ \dot{Y} = -S_y v_y, \quad \dot{P} = \varepsilon S_x^p v_p, \quad 0 = S_y v_y + S_x v_x - \alpha \varepsilon S_x v_p \]  
- **Constraints**: Enzyme capacity & cellular composition
  \[ \sum_{j \in \gamma_i} |v_j(t)/k_j| \leq P_i(t), \quad \sum_i c_i P_i(t) \leq 1 \]  
- **Optimisation**: Linear dynamic

  \[ \max_{v,Y,P} \int_0^{t_{\text{end}}} c^T P(t) dt \text{ such that (7) and (8)} \]
**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**
  \[
  v_{\text{min}} \leq v \leq v_{\text{max}}
  \]

- **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 \( z = (Y, P) \).

\[
\begin{align*}
\max_{r(z, z_0)} & \quad \int_0^{t_{\text{end}}} \Phi(z(t), v(t)) \, dt + \Psi(z(t_{\text{end}})) \\
\text{s.t.} & \quad \dot{Y} = -S^y_y V_y, \quad \dot{P} = \varepsilon S^p_y V_y, \\
& \quad S^x_y V_y + S^x_x V_x - \alpha \varepsilon S^p_y V_p = 0, \\
& \quad z(0) = z_0, \\
& \quad z(t) \geq 0, \\
& \quad v_{\text{min}} \leq v(t) \leq v_{\text{max}}, \\
& \quad H_C v(t) \leq H_E P(t), \quad H_B P(t) \leq h_B.
\end{align*}
\]

**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)
  \[
  \dot{Y} = -v_y, \quad \dot{P} = \varepsilon v_p, \quad v_y = \alpha \varepsilon v_p
  \]

**Objective functionals**

- **Maximization of terminal biomass**
  \[
  J_1 = P(t_{\text{end}})
  \]

- **Maximization of discounted biomass integral**
  \[
  J_2 = \int_0^{t_{\text{end}}} P(\tau) e^{-\mu \tau} \, d\tau
  \]

- **Minimization of time to consume nutrients**
  \[
  J_3 = -\int_0^{t_{\text{end}}} d\tau = -t_{\text{end}}
  \]

with \( Y(t_{\text{end}}) = 0 \).
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?

Core cellular network

Covert et al. 2001 (for the metabolic part)

Minimal network yields Monod model

- Assume Michaelis-Menten kinetics

$$V_Y = \frac{V_{m,Y} P Y}{K_Y \delta Y + Y}, \quad \epsilon V_P = \frac{V_{m,P} P X}{K_P + X}.$$

- Choose consistent parameters

$$V_{m,Y} = \left(\frac{1}{k_Y} + \frac{1}{\alpha k_p}\right)^{-1}, \quad V_{m,P} = \left(\frac{\alpha}{k_Y} + \frac{1}{k_P}\right)^{-1}.$$ 

Simulation (assuming kinetics)

Dynamic optimization (no kinetics assumed)

Network specification

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Enz</th>
<th>$k_{cat}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exchange reactions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carb1 $\rightarrow$ A</td>
<td>$T_C$</td>
<td>3000</td>
</tr>
<tr>
<td>Carb2 $\rightarrow$ A</td>
<td>$T_C$</td>
<td>2000</td>
</tr>
<tr>
<td>Fext $\rightarrow$ F</td>
<td>$T_F$</td>
<td>3000</td>
</tr>
<tr>
<td>O2ext $\rightarrow$ O2</td>
<td>$S$</td>
<td>1000</td>
</tr>
<tr>
<td>D $\rightarrow$ Dext</td>
<td>$S$</td>
<td>1000</td>
</tr>
<tr>
<td>E $\rightarrow$ Eext</td>
<td>$S$</td>
<td>1000</td>
</tr>
<tr>
<td>Hext $\rightarrow$ A</td>
<td>$T_H$</td>
<td>2000</td>
</tr>
<tr>
<td>Metabolic reactions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A + ATP $\rightarrow$ B</td>
<td>$E_B$</td>
<td>1800</td>
</tr>
<tr>
<td>B + G $\rightarrow$ 2ATP + 2NADH</td>
<td>$E_B$</td>
<td>1800</td>
</tr>
<tr>
<td>B $\rightarrow$ F</td>
<td>$E_F$</td>
<td>1800</td>
</tr>
<tr>
<td>C $\rightarrow$ G</td>
<td>$E_G$</td>
<td>1800</td>
</tr>
<tr>
<td>G $\rightarrow$ 0.6G + 2NADH</td>
<td>$E_G$</td>
<td>1800</td>
</tr>
<tr>
<td>C $\rightarrow$ 2ATP + 3D</td>
<td>$E_G$</td>
<td>1800</td>
</tr>
<tr>
<td>C $\rightarrow$ 4NADH $\rightarrow$ 3E</td>
<td>$E_G$</td>
<td>1800</td>
</tr>
<tr>
<td>G $\rightarrow$ ATP + 2NADH $\rightarrow$ H</td>
<td>$E_G$</td>
<td>1800</td>
</tr>
<tr>
<td>NADH + O $\rightarrow$ ATP</td>
<td>$E_T$</td>
<td>1800</td>
</tr>
</tbody>
</table>

Biomass reactions

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Enz</th>
<th>$k_{cat}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>400H + 1600ATP $\rightarrow$ T_C</td>
<td>R</td>
<td>2.5</td>
</tr>
<tr>
<td>1500H + 6000ATP $\rightarrow$ T_C</td>
<td>R</td>
<td>0.67</td>
</tr>
<tr>
<td>400H + 1600ATP $\rightarrow$ T_F</td>
<td>R</td>
<td>2.5</td>
</tr>
<tr>
<td>400H + 1600ATP $\rightarrow$ T_S</td>
<td>R</td>
<td>2.5</td>
</tr>
<tr>
<td>400H + 1600ATP $\rightarrow$ T_E</td>
<td>R</td>
<td>2.5</td>
</tr>
<tr>
<td>400H + 1600ATP $\rightarrow$ T_D</td>
<td>R</td>
<td>2.5</td>
</tr>
<tr>
<td>500H + 2000ATP $\rightarrow$ E_B</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>500H + 2000ATP $\rightarrow$ E_F</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>100H $\rightarrow$ 4000ATP $\rightarrow$ E_D</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>100H $\rightarrow$ 4000ATP $\rightarrow$ E_E</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>1500H + 6000ATP $\rightarrow$ E_F</td>
<td>R</td>
<td>0.67</td>
</tr>
<tr>
<td>500H + 2000ATP $\rightarrow$ E_B</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>500H + 2000ATP $\rightarrow$ E_F</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>200H + 4000C + 16000ATP $\rightarrow$ R</td>
<td>R</td>
<td>0.2</td>
</tr>
<tr>
<td>250H + 250C + 250F + 1500ATP $\rightarrow$ S</td>
<td>R</td>
<td>3</td>
</tr>
</tbody>
</table>

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

$$J = \int_{0}^{t_{\text{end}}} c_{bm} P(t) e^{-\mu t} dt$$

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.