Escherichia coli HGT – a novel high glucose throughput chassis especially designed for typical production conditions in large scale based on comprehensive systems biology studies

Ralf Takors
July 20th, 2017
Scaling-up *E. coli*
Modelling Strain Responses
Conclusions for Strain Engineering
Transferring Results

- Joana Simen
- Michael Löffler
- Alexander Niess
Scale-up is the process that enables lab-scale performance in large-scale, ensuring economic feasibility.
Large Scale Constraints

- Oxygen transfer rate (OTR) is limited (~150 – 180 mmol/Lh)
- Cooling capacities are limited

But: productivities (g/L_{reactor} h) should be maximized

Michalowski et al, 2017 Metabolic Engineering
Ideal Strain for Large-Scale Application

The ideal producer should:

• enable high glucose uptake rates, even under resting condition
• be blind with respect to extracellular heterogeneities

Goal: to create a novel *E. coli* chassis with fundamentally changed properties.

Drawback: transcriptional responses under large-scale conditions not known

Approach:

To perform systems biology studies on large-scale performance for deriving guidelines for synthetic strain construction.
Successful Bioprocess Development Needs to Pass Multiple Scales

What is the impact of repeated triggering on cellular performance?

Simulated glucose distribution in 900 L stirred tank reactor
Lapin et al., 2004

- cell-free protein synthesis
- metabolic analysis
- regulatory networks & strain engineering
- population dynamics
- scale-up analysis
Scaling-up *E. coli*

- Impact of Glucose Gradients on Cellular Performance -
Development of a STR-PFR system to simulate oscillating gradients

- Volume PFR/ total volume: 25%
- Online: temperature and oxygen
- Thermally insulated

Plug-flow like behavior ($\tau_{PFR} = 125\ s, \ Bo=83$)
STR-PFR Experiments With *E. coli*: Glucose Limitation

**I Batch:**
- $\mu = \mu_{\text{max}}$
- $V_{\text{total}} = V_{\text{STR}} = 1.5 \text{ l}$

**II Chemostat ($\mu = 0.2 \text{ h}^{-1}$)**
- Steady-state (feed)
- $V_{\text{total}} = V_{\text{STR}} = 1.5 \text{ l}$

**III Chemostat ($\mu = 0.2 \text{ h}^{-1}$) +PFR**
- PFR: glucose starvation zone
- $V_{\text{total}} = V_{\text{STR}} + V_{\text{PFR}} = 1.12 \text{ l} + 0.38 \text{ l}$
- Residence time PFR 125 s

**Diagram:**
- Well mixed
- Inhomogeneously mixed

**Long-term responses:**
- 25 min
- 120 min
- 28 h

Simen, Löfler et al. 2016 Metabolic Engineering
Glucose:

Short-Term Responses
Intracellular Nucleotides (Carbon Limitation-Starvation)

- Decreasing intracellular ATP-Concentration
- Increasing ADP-Level
- Fast Energy charge reduction (30 sec)
- Decreasing GTP-Level
- Sigmoidal ppGpp accumulation

Sequential response along the PFR
< 30 sec: Reduced Energy charge
< 70 sec: Decreasing GTP-Level
30-90 sec: Increasing ppGpp-Level

Transcriptional Response
Short-term Response on Glucose Limitation: Differentially Expressed (DE) Genes

- Increasing limitation inside PFR
- Filter for DE genes
  - 1% false discovery rate
  - P5 down: 836
  - P5 up: 955
  - Fold change >1.5 fold
  - P5 down: 369
  - P5 up 266

→ Immediate DE response after entering PFR
  - ’slow‘ dynamics < 70 s
  - Transcriptional boost > 70 s

Simen, Löffler et al. 2016 Metabolic Engineering
Differentially Expressed Genes (PFR outlet vs STR)

Information storage and processing
K Transcription
L Replication, recombination and repair
J Translation, ribosomal structure and biogenesis

Cellular processes and signaling
O Posttranslational modification, protein turnover and chaperones
D Cell cycle control, cell division and chromosome partitioning
T Signal transduction mechanisms
M Cell wall, membrane, envelope biogenesis
V Defense mechanisms
U Intracellular trafficking, secretion, vesicular transport
N Cell motility

Metabolism
G Carbohydrate transport and metabolism
C Energy production and conversion
E Amino acid transport and metabolism
I Lipids transport and metabolism
P Inorganic ion transport and metabolism
Q Secondary metabolites biosynthesis, transport and metabolism
F Nucleotide transport and metabolism
H Coenzyme transport and metabolism

Poorly characterized
R General function prediction only
S Function unknown
**General finding:**

→ induction of carbon import, metabolism and energy generation

→ repression of energy intensive processes
Glucose:

Long-Term Responses
Long-Term Response on Glucose Gradients (triplicate results):
Reference – before connection with PFR

Simen, Löffler et al. 2016 Metabolic Engineering

- distinct initial state $S_0$ in STR
Long-Term Response on Glucose Gradients (triplicate results):
Phase I after connection with PFR

- distinct initial state $S_0$ in STR
- strong transcriptional changes after PFR connection
Long-Term Response on Glucose Gradients (triplicate results):
Phase II after connection with PFR

- distinct initial state $S_0$ in STR
- strong transcriptional changes after PFR connection
- transcriptional changes converge to a new steady-state
Long-Term Response on Glucose Gradients (triplicate results):
Repeated Transcriptional Shifts from STR to PFR

- distinct initial state $S_0$ in STR
- strong transcriptional changes after PFR connection
- transcriptional changes converge to a new steady-state
- gradients in PFR cause trackable transcriptional changes leading to a stable distribution of transcriptional patterns

Gradients cause massive, periodic on-/off-switching of genes finally leading to stable transcriptional patterns coexisting next to each other
What can we learn from modelling?
What can we learn from modelling?

Predicting the Transcriptional Dynamics
Agent-Based Modelling For Predicting Transcript Dynamics

Agent (single cell)

- polymerase movement: constant elongation rate
  - constraint: minimum distance between ribosomes
- attenuation modelling neglected for simplification
  - Simplification: start only valid under nitrogen limitation
  - Transcription and translation are closely coupled
- Number of ribosomes per gene is variable and specific for the gene
- mRNA degradation considered with constant degradation elongation rate, initiated at start codon of transcription
- No protein degradation, only dilution by growth

Population (STR-PFR operating in continuous mode)

- STR = ideally mixed; PFR = plugflow reactor
- Population balances:
  - (1) PFR entering
  - (2) drained off by efflux
  - (3) cell division, no initiation of transcription
- Fate of 10,000 cells was tracked

Number of likely events per time:

\[
\begin{align*}
\alpha_1 &= N_{STR} \frac{V_{PFR}}{V_{STR}} \\
\alpha_2 &= N_{STR} \frac{V_{Feed}}{V_{STR}} \\
\alpha_3 &= N_{STR}^0 D
\end{align*}
\]

Gillespie algorithm:

\[
\tau = \frac{1}{\sum \alpha_i} \ln \left( \frac{1}{r_1} \right)
\]

\[
\sum_{j=1}^{i-1} \alpha_j \leq r_2 \sum_{j=1}^{3} \alpha_j \leq \sum_{j=1}^{i} \alpha_j
\]
### Simulation Parameters – No Regression

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_{R_{\text{NAP \ elo}}}$</td>
<td>21</td>
<td>Nucleotides per second</td>
</tr>
<tr>
<td>$v_{\text{Ribosome \ elo}}$</td>
<td>21</td>
<td>Nucleotides per second</td>
</tr>
<tr>
<td>$v_{R_{\text{Nase \ elo}}}$</td>
<td>21</td>
<td>Nucleotides per second</td>
</tr>
<tr>
<td>$\Delta x$</td>
<td>100</td>
<td>Nucleotides</td>
</tr>
<tr>
<td>$\Delta y$</td>
<td>100</td>
<td>Nucleotides</td>
</tr>
<tr>
<td>$\Delta z$</td>
<td>100</td>
<td>Nucleotides</td>
</tr>
<tr>
<td>$t_{\text{ind}}$</td>
<td>[30 125]</td>
<td>Seconds</td>
</tr>
<tr>
<td>$V_{PFR}$</td>
<td>180</td>
<td>mL min$^{-1}$</td>
</tr>
<tr>
<td>$V_{\text{Feed}}$</td>
<td>5</td>
<td>mL min$^{-1}$</td>
</tr>
<tr>
<td>$V_{STR}$</td>
<td>1,120</td>
<td>mL</td>
</tr>
<tr>
<td>$D$</td>
<td>0.2</td>
<td>h$^{-1}$</td>
</tr>
<tr>
<td>$N_{STR}^0$</td>
<td>10,000</td>
<td>cells</td>
</tr>
</tbody>
</table>

- Elongation rate
- Ribosome movement on mRNA
- RNAse degradation rate
- Minimum distance between two RNAPs
- Minimum distance between ribosomes
- Closest distance of RNAse to ribosome
- Induction period
- Flow through PFR
- Feed
- STR volume
- Dilution rate of the system
- Total number of tracked cells
Once Initiated, Transcriptional (TC) Stimulus Propagates From PFR into STR Followed by Translation (TL): Example TRP Operon

Induction of *trp* operon during N-limitation

Niess et al. 2017. *Frontiers in Microbiology*

**Key findings:**

- Once initiated in PFR, transcription and translation continues in STR
- Mayor burden is located in STR
- Cells with different transcriptional patterns co-exist next to each other: population heterogeneity
Prediction of population transcript levels in STR
Agent based modeling tracking the fate of 10,000 cells

Niess et al. 2017. Frontiers in Microbiology

High prediction quality of short – and long-term transcriptional dynamics
What can we learn from modelling?

Deriving Constraints for Creating Robust Strains
For Large-Scale Application
Modeling Assumptions for Estimating Cellular Efforts Periodically Switching Genes ON/OFF

Simen, Löffler et al. 2016 Metabolic Engineering

- Select transcripts with FDR < 1%
- Group in up- and downregulated genes
- De novo precursor synthesis:
  - calculate individual ATP needs for G, A, T, C formation from metabolic precursors
  - calculate ATP needs for amino acid production assuming average protein composition

- **Transcription**: For each mRNA with individual G, A, T, C content: 2 ATP/nucleotide
- **Translation**: 4 ATP/amino acid
- Considering mRNA dynamics and PFR residence time: 11 ribosomes per mRNA
- **Completing translation** after recycling in STR
Modeling Additional ATP Costs for Transcription and Translation

Individual ATP balances for each gene, reference: native maintenance

ATP cost reference: *E. coli* maintenance (Taymaz-Nikerel et al. 2010)

Maintenance demands are increased by 40 – 50 % in large-scale!
Represents missing energy for hyperproducers.
## Exploiting the Results: Identifiers for Smart Genome Reduction

### Top 20 Gene Targets for Deletion/Modulation Saving ATP

**Simen, Löffler et al. 2016 Metabolic Engineering**

---

<table>
<thead>
<tr>
<th>Gene†</th>
<th>Percentage increasing the growth-independent maintenance, %</th>
<th>COG</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>De novo mRNA synthesis</td>
<td>Translation</td>
<td>∑</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene†</th>
<th>Percentage increasing the growth-independent maintenance, %</th>
<th>COG</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>tliC</td>
<td>2.70</td>
<td>0.40</td>
<td>3.10</td>
</tr>
<tr>
<td>aroF</td>
<td>0.67</td>
<td>0.10</td>
<td>0.77</td>
</tr>
<tr>
<td>aldA</td>
<td>0.48</td>
<td>0.07</td>
<td>0.55</td>
</tr>
<tr>
<td>cstA</td>
<td>0.35</td>
<td>0.05</td>
<td>0.40</td>
</tr>
<tr>
<td>aceA</td>
<td>0.34</td>
<td>0.05</td>
<td>0.39</td>
</tr>
<tr>
<td>cspD</td>
<td>0.31</td>
<td>0.05</td>
<td>0.36</td>
</tr>
<tr>
<td>aceB</td>
<td>0.27</td>
<td>0.04</td>
<td>0.31</td>
</tr>
<tr>
<td>trg</td>
<td>0.27</td>
<td>0.04</td>
<td>0.31</td>
</tr>
<tr>
<td>groL</td>
<td>0.26</td>
<td>0.04</td>
<td>0.30</td>
</tr>
<tr>
<td>dnaK</td>
<td>0.24</td>
<td>0.03</td>
<td>0.27</td>
</tr>
<tr>
<td>yfIA</td>
<td>0.18</td>
<td>0.03</td>
<td>0.21</td>
</tr>
<tr>
<td>gatC</td>
<td>0.17</td>
<td>0.03</td>
<td>0.20</td>
</tr>
<tr>
<td>flgL</td>
<td>0.16</td>
<td>0.02</td>
<td>0.19</td>
</tr>
<tr>
<td>flgK</td>
<td>0.13</td>
<td>0.02</td>
<td>0.15</td>
</tr>
<tr>
<td>acs</td>
<td>0.13</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>mdh</td>
<td>0.12</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>kgtP</td>
<td>0.11</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>tliA</td>
<td>0.11</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>glnH</td>
<td>0.10</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>yjdA</td>
<td>0.10</td>
<td>0.01</td>
<td>0.11</td>
</tr>
</tbody>
</table>

*Only genes which expression was always significantly changed between STR and PFR PS (FDR < 0.01) were selected for the calculations (core genes).*

---

**Ongoing research**

---

29 Takors, SysSyn Meeting, Munich
Transferring Results:

*E. coli* HGT
The **ideal producer** should:

- be blind with respect of extracellular heterogeneities
- enable high glucose uptake rates, even under resting condition

Goal:

**to create a novel E. coli chassis** with fundamentally changed properties.

$q_s$ is increased **but** $m_s$ stays constant
Engineering \textit{E. coli} HGT Following Two Guidelines

HGT = high glucose throughput

Modulating stringent response to make \textit{E. coli} blind

Modulating central metabolism

Integrating \textit{aceE*}
(pyruvate dehydrogenase subunit)

WO patent filed:
Bacterial strain and method for high throughput of sugar in the microbial conversion into biosynthetic products

Michalowski et al, 2017 \textit{Metabolic Engineering}
E. coli HGT

Michalowski et al, 2017 Metabolic Engineering

spoT[R290E;K292D] ΔrelA aceE[G267C]
Installing ammonia limitation under glucose saturated conditions to test engineered strains
**E. coli HGT**

Includes *aceE* + changes in *relA* and *spoT*

---

*E. coli* HGT shows no ppGpp changes under ammonia limited conditions, i.e. no stringent response initiated

---

*E. coli* HGT provides a surplus of pyruvate for downstream use.
E. coli HGT – glucose uptake kinetics

E. coli HGT shows about 10 fold higher glucose uptake for non-growing conditions than needed for maintenance demands and reaches maximum uptake with 0.3 1/h. The surplus of carbon is available as pyruvate predominately.
Summary & Thanks

- Mixing times of max 70 s should be installed preventing massive transcriptional responses.
- Massive transcriptional dynamics are induced by substrate heterogeneity causing maintenance increase (1.4-1.5 fold).
- Maintenance dynamics and transcriptional adaptation can be well modelled/predicted.
- Large-scale performance can be simulated.
- Novel chassis *E. coli* HGT created.

Thanks

Joana Simen, Michael Löffler,
Annette Michalowski, Alexander Niess

Thanks for Cooperation:
- IMG (AG Riess, Tübingen) for transcript measurements
- IMB (AG Sprenger) for reporter strain construction

Funding: Thanks to BMBF (eBio, ReCOgNice), EU (ST-Flow)
Different Time Scales of Response

Niess et al. 2017. Frontiers in Microbiology