Network-wide Thermodynamic Constraints Shape NAD(P)H Cofactor Specificity of Metabolic Reactions

Steffen Klamt and Pavlos Stephanos Bekiaris

Max Planck Institute for Dynamics of Complex Technical Systems
Research Group „Analysis and Redesign of Biological Networks“

11/15/2023
NADH / NADPH: Two ubiquitous redox cofactors

Central metabolism of *E. coli*:

Nicotinamide adenine dinucleotide (NAD\(^+\))

\[
\text{NAD}^+ + 2\ e^- + H^+ \leftrightarrow \text{NADH} \quad (\Delta E'^\circ = -320 \text{ mV})
\]

Nicotinamide adenine dinucleotide phosphate (NADP\(^+\))

\[
\text{NADP}^+ + 2\ e^- + H^+ \leftrightarrow \text{NADPH} \quad (\Delta E'^\circ = -320 \text{ mV})
\]

*E. coli* (iML1515): 128 reactions use NAD(H), 110 reactions use NADP(H), 6 reactions use both.
NADH / NADPH: Two ubiquitous redox cofactors

Why two pools of (very similar) redox cofactors?

Simultaneous operation of oxidation and reduction reactions!

\[
[NAD^+] \gg [NADH]  \quad \text{in vivo } NADH/NAD^+ \text{ ratio of } \approx 0.03 \quad \text{in } E. \col (\text{aerobic, Bennet et al., 2009})
\]

\[
[NADP^+] \ll [NADPH]  \quad \text{in vivo } NADPH/NADP^+ \text{ ratio of } \approx 57 \quad \text{in } E. \col (\text{aerobic, Bennet et al., 2009})
\]

**Catabolism:**

\[
A + NAD^+ \quad \text{favored} \quad \rightarrow \quad B + NADH
\]

\[
D + NAD^+ \quad \text{less favored!} \quad \leftarrow \quad C + NADH
\]

**Anabolism:**

\[
V + NADPH \quad \text{favored} \quad \rightarrow \quad X + NADP^+
\]

\[
Z + NADPH \quad \text{less favored!} \quad \leftarrow \quad Y + NADP^+
\]

- Are two pools of redox cofactors really advantageous?
- What shapes the NAD(P)(H) reaction specificities in the network?

\[ \text{Hypothesis: } \text{NAD(H)} \text{ and NADP(H)} \text{ reaction specificities are distributed such that the network-wide thermodynamic driving force for growth is optimized.} \]
Driving Forces of Reactions and Pathways

Example network with (given) standard Gibbs free energies of reactions:

Driving force of a single reaction: \( f_i = -\Delta_r G'_i = -\Delta_r G'i^o - R \cdot T \cdot \sum_j N_{ji} \cdot \ln(c_j) \)

Example: \( R_1 (A \rightarrow B) \) with fixed concentrations: \( f_{R1} = -\Delta_r G'_{R1} - R \cdot T \cdot (-\ln(c_A) + \ln(c_B)) \)

Driving force of a specific pathway: minimum driving force of all reactions of the pathway

\( f_{\text{pathway}} = \min(f_{R1}, f_{R2}, f_{R3}, f_{R4}, f_{R8}) \) with fixed concentrations
Max-min Driving Force (MDF) of a Pathway and Network

Max-min driving force (MDF) of a pathway: maximal achievable driving force of the pathway

\[ MDF_{\text{Pathway}} = \max(f_{\text{Pathway}}) = \max(\min(f_{R1}, f_{R2}, f_{R3}, f_{R4}, f_{R5})) \] within given concentration ranges

Red: given
Blue: calculated

MDF of network (also called OptMDF): find flux distribution \( \mathbf{r} \) maximizing MDF of active reactions

\[ MDF_{\text{Network}} = \max(MDF(\mathbf{r})) \] with given concentration ranges and flux constraints

(Noor et al., 2014)

(Hädicke et al., 2018)

\( \rightarrow \) Mixed-Integer Linear Program (MILP) within constraint-based metabolic model
Max-min Driving Force (MDF) in a Network

(MDF in a network
(Hädicke et al., 2018)

→ Mixed-Integer Linear Program (MILP)
within constraint-based metabolic model

\[
\begin{align*}
\text{Maximize} & \quad B \\
\text{s.t.} & \quad Nr = 0 \\
& \quad \alpha_i \leq r_i \leq \beta_i \\
& \quad \ln(c_{\text{min}}) \leq x \leq \ln(c_{\text{max}}) \\
& \quad r_i \leq z_i \cdot \beta_i \\
& \quad z_i \in \{0,1\} \\
& \quad f_i = -\Delta_r G'_i = -\Delta_r G^\circ - RT \cdot (N_{*,i})^T \cdot x \\
& \quad B \leq f_i + M \cdot (1 - z_i), \quad (M \text{ very large})
\end{align*}
\]
Max-min Driving Force (MDF) of a Pathway and Network

Max-min driving force (MDF) of a pathway: maximal achievable driving force of the pathway

\[ MDF_{\text{Pathway}} = \max(f_{\text{Pathway}}) = \max(\min(f_{R1}, f_{R2}, f_{R3}, f_{R4}, f_{R5})) \] within given concentration ranges

\[
\begin{align*}
A & \rightarrow B & f_{R1} & \rightarrow C & f_{R2} & \rightarrow D & f_{R3} & \rightarrow E & f_{R4} & \rightarrow F & f_{R5} & \rightarrow G & f_{R6} & \rightarrow H & f_{R7} & \rightarrow I & f_{R8} & \rightarrow J & f_{R9} & \rightarrow K
\end{align*}
\]

Metabolite concentrations under MDF minimize enzyme costs (neglecting saturation effects)

(Noor et al., 2013)

MDF of network (also called OptMDF): find flux distribution \( \mathbf{\pi} \) maximizing MDF of active reactions

\[ MDF_{\text{Network}} = \max(MDF(\mathbf{\pi})) \] with given concentration ranges and flux constraints

Here: SubMDF with respect to NAD(P)(H)-dependent reactions

(Hädicke et al., 2018)

Mixed-Integer Linear Program (MILP) within constraint-based metabolic model

(This work, 2023)

Mixed-Integer Linear Program (MILP) within constraint-based metabolic model

Here: SubMDF with respect to NAD(P)(H)-dependent reactions
Max-min Driving Force in a Subnetwork (SubMDF)

### MDF in a network
(Hädicke et al., 2018)

- Mixed-Integer Linear Program (MILP) within constraint-based metabolic model

\[
\begin{align*}
\text{Maximize} & \quad B \\
\text{s.t.} & \quad N_r = 0 \\
& \quad \alpha_i \leq r_i \leq \beta_i \\
& \quad \ln(c_{\text{min}}) \leq x \leq \ln(c_{\text{max}}) \\
& \quad r_i \leq z_i \cdot \beta_i \\
& \quad z_i \in \{0,1\} \\
& \quad f_i = -\Delta r G'_i = -\Delta_r G'^\circ - RT \cdot (N_{*,i})^T \cdot x \\
& \quad B \leq f_i + M \cdot (1 - z_i), \quad (M \text{ very large})
\end{align*}
\]

### SubMDF in a network
(Bekiaris and Klamt, 2023)

- Mixed-Integer Linear Program (MILP) within constraint-based metabolic model

\[
\begin{align*}
\text{Maximize} & \quad B_{sub} \\
\text{s.t.} & \quad N_r = 0. \\
& \quad \alpha_i \leq r_i \leq \beta_i \\
& \quad \ln(c_{\text{min}}) \leq x \leq \ln(c_{\text{max}}) \\
& \quad r_i \leq z_i \cdot \beta_i \\
& \quad z_i \in \{0,1\} \\
& \quad f_i = -\Delta r G'_i = -\Delta_r G'^\circ - RT \cdot (N_{*,i})^T \cdot x \\
& \quad B \leq f_i + M \cdot (1 - z_i), \quad (M \text{ very large}) \\
& \quad B \geq 0.1
\end{align*}
\]

Thermodynamic feasibility of entire flux vector
Minimum Driving force of the subset \( S \) of selected reactions

\[
B_{sub} \leq f_j + M \cdot (1 - z_j), \quad \forall j \in S
\]
**Max-min Driving Force (MDF) of a Pathway and Network**

**Max-min driving force (MDF) of a pathway**: maximal achievable driving force of the pathway

\[ MDF_{\text{Pathway}} = \max(f_{\text{Pathway}}) = \max(\min(f_{R1}, f_{R2}, f_{R3}, f_{R4}, f_{R5})) \text{ within given concentration ranges} \]

![Diagram of a pathway with metabolites A to P and fluxes f_{R1} to f_{R5}](image)

**MDF of network**: also called OptMDF: find flux distribution \( \mathbf{r} \) maximizing MDF of active reactions

\[ MDF_{\text{Network}} = \max(MDF(\mathbf{r})) \text{ with given concentration ranges and flux constraints} \]

![Diagram of a network with metabolites A to P and fluxes f_{R1} to f_{R5}](image)

**MDF of subnetwork (SubMDF)**: find flux distribution \( \mathbf{r} \) maximizing MDF of selected set of reactions

\[ \text{e.g., } \text{SubMDF} = MDF(R_1, R_5) \text{ with given concentration ranges and flux constraints} \]

![Diagram of a subnetwork with metabolites A to P and fluxes f_{R1} to f_{R5}](image)

*(Noor et al., 2013)*

Metabolite concentrations under MDF minimize enzyme costs (neglecting saturation effects)

*(Hädicke et al., 2018)*

\( \Rightarrow \) Mixed-Integer Linear Program (MILP) within constraint-based metabolic model

What NAD(P)(H) specificities maximize the MDF/SubMDF for growth-related flux distributions and how close is the wild-type specificity to this optimal specificity?

*(This work, 2023)*

\( \Rightarrow \) Mixed-Integer Linear Program (MILP) within constraint-based metabolic model

Here: SubMDF with respect to NAD(P)(H)-dependent reactions
Reconfiguration of a given (stoichiometric) metabolic model for TCOSA:

**Application to E. coli**

**Resulting model:** iML1515_TCOSA (derived from genome-scale E. coli model iML1515; Monk et al., 2017).

- Substrate: glucose. Aerobic (+O₂) and anaerobic (-O₂) conditions.
- Metabolite concentration ranges: [10⁻⁶… 0.02 M]
- ΔrG° values: from eQuilibrator (Flamholz et al., 2012) via its Python API (Beber et al., 2021)
**Wild-type specificity:** Use original NAD(P)(H) specificity for the NAD(P)(H)-dependent reactions

**Flexible specificity:** NAD(P)(H) specificity can be freely selected for each reaction (but only one at a time for each reaction)

**Single cofactor pool:** Only NAD(H)-dependent reactions can be used (NADP(H) not allowed)

**Random specificity:** 1'000 random specificities (stochastic coin flip to select NAD(H) or NADP(H) specificity for each redox-cofactor-dependent reaction)
Conclusion #1: The wild-type NAD(P)H specificity enables high thermodynamic potentials that are (a) close to the theoretical maximum and (b) significantly better than random specificities or using a single redox cofactor pool.
Analysis 2: Necessary Swaps in Wild-type Specificity to Reach the Theoretical Maximal (Sub)MDF of the Flexible Specificity

<table>
<thead>
<tr>
<th>Oxygen availability</th>
<th>Growth rate [h⁻¹]</th>
<th>Number of necessary swaps in wild type to reach (Sub)MDF of flexible specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aerobic</td>
<td>0.868</td>
<td>MDF: 6, SubMDF: 2</td>
</tr>
<tr>
<td>Aerobic</td>
<td>0.818</td>
<td>MDF: 0, SubMDF: 3</td>
</tr>
<tr>
<td>Aerobic</td>
<td>0.768</td>
<td>MDF: 0, SubMDF: 2</td>
</tr>
<tr>
<td>Aerobic</td>
<td>0.718</td>
<td>MDF: 0, SubMDF: 2</td>
</tr>
<tr>
<td>Aerobic</td>
<td>0.668</td>
<td>MDF: 0, SubMDF: 0</td>
</tr>
<tr>
<td>Aerobic</td>
<td>0.618</td>
<td>MDF: 0, SubMDF: 0</td>
</tr>
<tr>
<td>Aerobic</td>
<td>0.568...0.518</td>
<td>MDF: 0, SubMDF: 0</td>
</tr>
<tr>
<td>Aerobic</td>
<td>0.468...0.118</td>
<td>MDF: 0, SubMDF: 0</td>
</tr>
<tr>
<td>Aerobic</td>
<td>0.068</td>
<td>MDF: 0, SubMDF: 0</td>
</tr>
<tr>
<td>Aerobic</td>
<td>0.05</td>
<td>MDF: 0, SubMDF: 0</td>
</tr>
<tr>
<td>Anaerobic</td>
<td>0.371</td>
<td>MDF: 1, SubMDF: 1</td>
</tr>
<tr>
<td>Anaerobic</td>
<td>0.321</td>
<td>MDF: 0, SubMDF: 0</td>
</tr>
<tr>
<td>Anaerobic</td>
<td>0.271</td>
<td>MDF: 9, SubMDF: 14</td>
</tr>
<tr>
<td>Anaerobic</td>
<td>0.221</td>
<td>MDF: 0, SubMDF: 2</td>
</tr>
<tr>
<td>Anaerobic</td>
<td>0.171</td>
<td>MDF: 0, SubMDF: 5</td>
</tr>
<tr>
<td>Anaerobic</td>
<td>0.121</td>
<td>MDF: 0, SubMDF: 4</td>
</tr>
<tr>
<td>Anaerobic</td>
<td>0.071</td>
<td>MDF: 0, SubMDF: 3</td>
</tr>
<tr>
<td>Anaerobic</td>
<td>0.05</td>
<td>MDF: 0, SubMDF: 3</td>
</tr>
</tbody>
</table>
Two frequently suggested cofactor swaps to increase (Sub)MDF:

1) Pyruvate dehydrogenase

CoA + pyruvate + NAD$^+$ $\rightarrow$ acetyl-CoA + CO$_2$ + NADH \hspace{1cm} (ΔG° of -34.37 kJ/mol)

$\rightarrow$ CoA + pyruvate + NADP$^+$ $\rightarrow$ acetyl-CoA + CO$_2$ + NADPH

Synthesis of NADPH (thermodynamically unfavorable) in a reaction that has very negative ΔG°.

2) Isocitrate dehydrogenase

isocitrate + NADP$^+$ $\rightarrow$ 2-oxoglutarate + CO$_2$ + NADPH \hspace{1cm} (ΔG° of +5.13 kJ/mol)

$\rightarrow$ isocitrate + NAD$^+$ $\rightarrow$ 2-oxoglutarate + CO$_2$ + NADH

Use NAD$^+$ (thermodynamically favorable) instead of NADP$^+$ to overcome the positive ΔG° of this reaction.
(But: unfavorable when using acetate as substrate!).
Analysis 3: Trends of NAD(P)(H) Concentration Ratios

Observed trends in *E. coli*:

\[
\frac{[\text{NADH}]}{[\text{NAD}^+]} \ll 1 \\
\frac{[\text{NADPH}]}{[\text{NADP}^+]} \gg 1
\]

*in vivo NADH/NAD*⁺* ratio of ≈0.03* 
*in vivo NADPH/NADP*⁺* ratio of ≈57* 

in *E. coli* (aerobic, Bennet et al., 2009)

\[ Q = \frac{\frac{[\text{NADH}]}{[\text{NAD}^+]}}{\frac{[\text{NADPH}]}{[\text{NADP}^+]}} \ll 1 \]

*in vivo* \(\approx 0.00053\) in *E. coli* 
(aerobic, Bennet et al., 2009)
Conclusion #2: Qualitative trends of relative NAD(P)(H) concentrations can be predicted
Analysis 4: Effect of a Third Redox Cofactor Pool (Flexible Specificity)

### NAD(H) vs NADP(H)

- NAD → Malate → NADP
- NADH → OAA + H⁺ → NADPH

- **MDH_FWD_ORIGINAL_NAD**
- **MDH_FWD_VARIANT_NADP**

### “NADX(H)”

- NADX → Malate → NADXH
- NADXH → OAA + H⁺

- **MDH_FWD_VARIANT_NADX**

### 3 Redox Potential Scenarios:

1. **As for NAD(P)(H)**
   - Standard redox potential (ΔE°) of -320 mV
   - ΔG° difference of ca. 61 kJ/mol

2. **Lower Potential**
   - ΔE° = -475 mV

3. **Higher Potential**
   - ΔE° = -165 mV

---

Standard redox potential (ΔE°) of -320 mV

ΔfG° difference of ca. 61 kJ/mol
Conclusion #3: A third redox cofactor pool could be advantageous if it has a low standard redox potential!
Several autotrophic organisms like acetogens use ferredoxin ($\Delta E^{\circ}$ of -420 mV) as a third major redox cofactor in many redox reactions.

→ Additional degree of freedom to maintain high thermodynamic driving forces in their complicated redox metabolism.

**Analysis 5: Robustness of the Results**

A) Robustness against random variations of $\Delta_r G^{\circ}$

(implemented by random variations of the $\Delta_l G^{\circ}$ of each metabolite)
Analysis 5: Robustness of the Results

B) Robustness against assumed metabolite concentration ranges

→ in vivo concentration values from Bennett et al., 2009 (aerobic conditions)

For MDF: single bottleneck
(independent of NAD(P)(H) specificities)
C) Changing the substrate: acetate instead of glucose (aerobic conditions only)

Conclusion #4: Results are robust against different variations
Conclusion

✓ TCOSA framework for analyzing the thermodynamic effects of (redox) cofactor swaps.

✓ Our analysis indicates that evolution shaped **the NAD(P)(H) specificity of reactions to enable high thermodynamic potentials** in the metabolic network.
  • minimizes enzyme demand for redox reactions (cf. also Goldford et al., 2022)

✓ We used **MDF as a measure** for the (network-wide) thermodynamic **potential**:
  
  **Caveat:** A cell is likely not in a state close to a computed MDF (e.g., enzyme kinetics affects feasible metabolite concentrations and thus the MDF).

  **But** the higher the (theoretical) MDF, the larger the **thermodynamic flexibility** of the network (broader ranges of feasible metabolite concentration)!

✓ TCOSA can be used for other species and/or other cofactor pairs (e.g., ATP/GTP) and even for **predicting optimal cofactor specificities** (e.g. metabolic engineering).

Acknowledgements

Thank you for your attention!