MAX PLANCK INSTITUTE FOR DYNAMICS OF COMPLEX TECHNICAL SYSTEMS MAGDEBURG

ARB ANALYSIS AND REDESIGN OF BIOLOGICAL NETWORKS

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

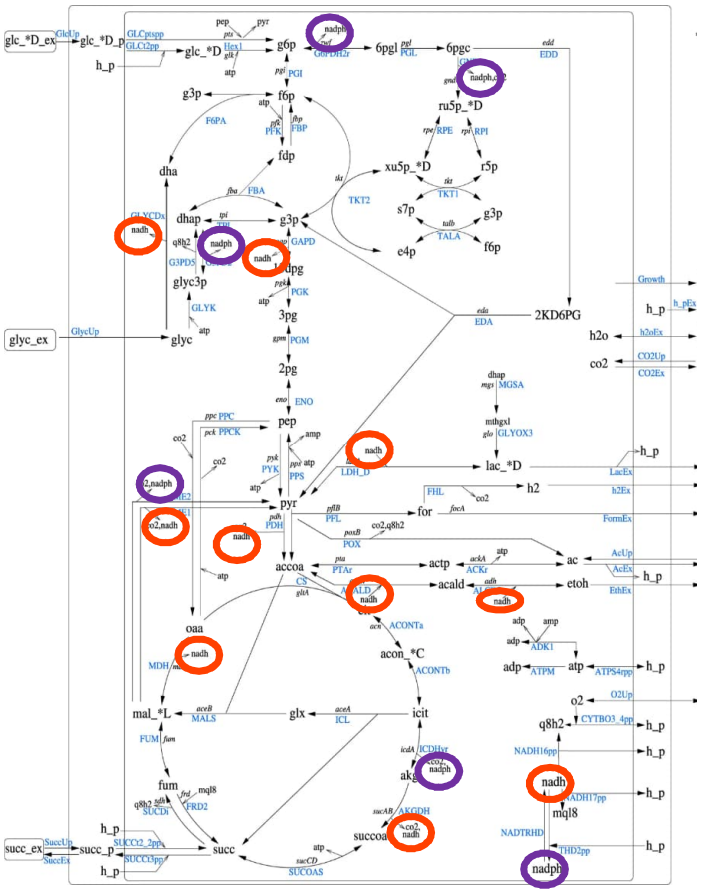
Steffen Klamt and Pavlos Stephanos Bekiaris

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

11/15/2023

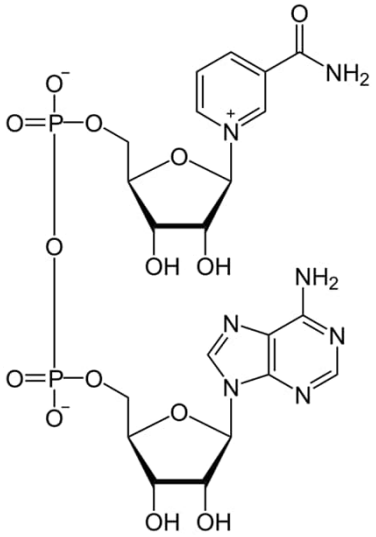
NADH / NADPH: Two ubiquitous redox cofactors

Central metabolism of E. coli:



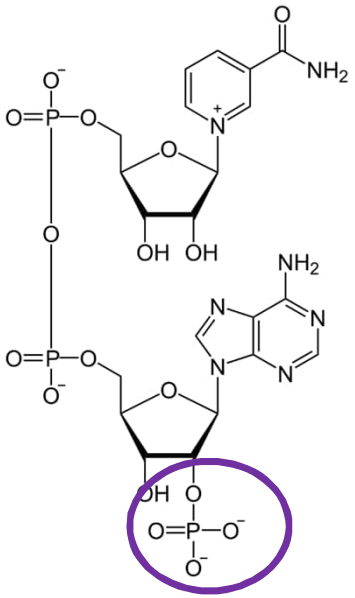
E. coli (iML1515): 128 reactions use NAD(H), 110 reactions use NADP(H), 6 reactions use both.

Nicotinamide adenine dinucleotide (NAD^+)



$NAD^+ + 2\ e^- + H^+ \leftrightarrow NADH$
$(\Delta E'° = -320\ mV)$

Nicotinamide adenine dinucleotide phoshpate ($NADP^+$)



$NADP^+ + 2\ e^- + H^+ \leftrightarrow NADPH$
$(\Delta E'° = -320\ mV)$

NADH / NADPH: Two ubiquitous redox cofactors

Why two pools of (very similar) redox cofactors?

Simultaneous operation of oxidation and reduction reactions!

$[NAD^+] \gg [NADH]$

in vivo $NADH/NAD^+$ **ratio of ≈0.03**
in *E. coli* (aerobic, Bennet et al., 2009)

Catabolism:

$A + NAD^+ \xrightarrow{favored} B + NADH$

$D + NAD^+ \xleftarrow[less\ favored!]{} C + NADH$

$[NADP^+] \ll [NADPH]$

in vivo $NADPH/NADP^+$ **ratio of ≈57**
in *E. coli* (aerobic, Bennet et al., 2009)

Anabolism:

$V + NADPH \xrightarrow{favored} X + NADP^+$

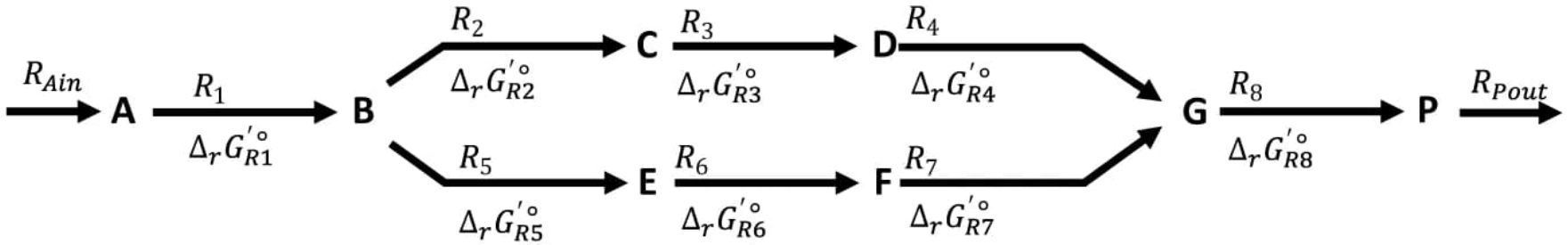
$Z + NADPH \xleftarrow[less\ favored!]{} Y + NADP^+$

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

→ *Hypothesis*: **NAD(H)** and **NADP(H)** 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:



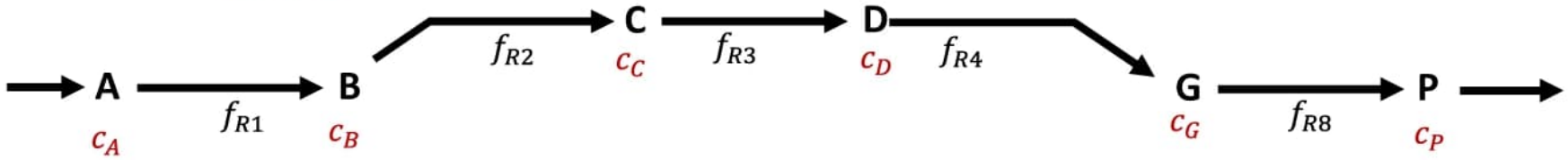
Driving force of a single reaction: $f_i = -\Delta_r G_i' = -\Delta_r G_i'^{\circ} - R \cdot T \cdot \sum_j N_{ji} \cdot \ln(c_j)$

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

Red: given
Blue: calculated

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

$f_{Pathway} = \min(f_{R1}, f_{R2}, f_{R3}, f_{R4}, f_{R8})$ with fixed concentrations

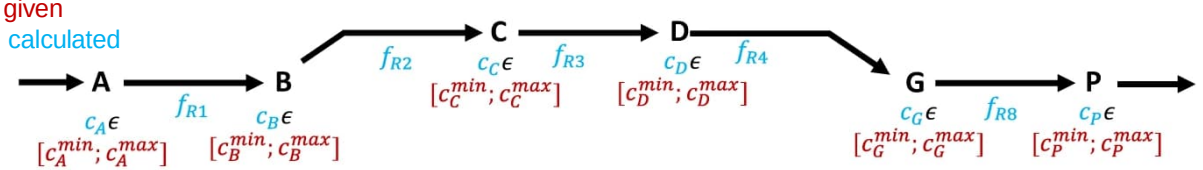


Max-min Driving Force (MDF) of a Pathway and Network

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

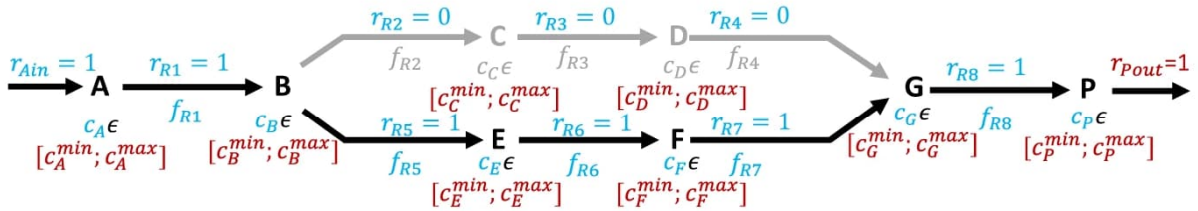
$MDF_{Pathway} = \max(f_{Pathway}) = \max(\min(f_{R1}, f_{R2}, f_{R3}, f_{R4}, f_{R8}))$ within given concentration ranges

Red: given
Blue: calculated



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

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



(Noor et al., 2014)

(Hädicke et al., 2018)
→ 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

$$\underset{\mathbf{x},\mathbf{r},\mathbf{z}}{Maximize} \quad B$$

$$s.t.$$

$$\mathbf{Nr} = \mathbf{0}$$

$$\alpha_i \leq r_i \leq \beta_i$$

$$\ln(\mathbf{c_{min}}) \leq \mathbf{x} \leq \ln(\mathbf{c_{max}})$$

$$r_i \leq z_i \cdot \beta_i$$

$$z_i \in \{0,1\}$$

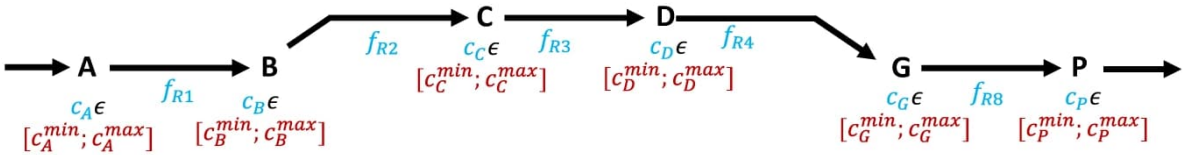
$$f_i = -\Delta_r G_i' = -\Delta_r G'^{\circ} - RT \cdot (\mathbf{N}_{*,i})^T \cdot \mathbf{x}$$

$$B \leq f_i + M \cdot (1 - z_i), \qquad (M\ very\ large)$$

Max-min Driving Force (MDF) of a Pathway and Network

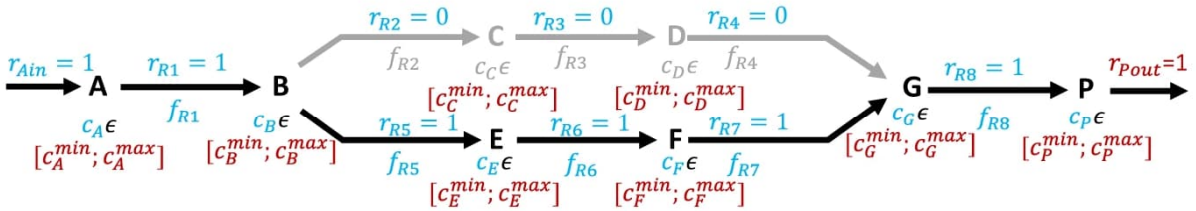
Max-min driving force (MDF) of a pathway: maximal achievable driving force of the pathway

$MDF_{Pathway} = \max(f_{Pathway}) = \max(\min(f_{R1}, f_{R2}, f_{R3}, f_{R4}, f_{R8}))$ within given concentration ranges



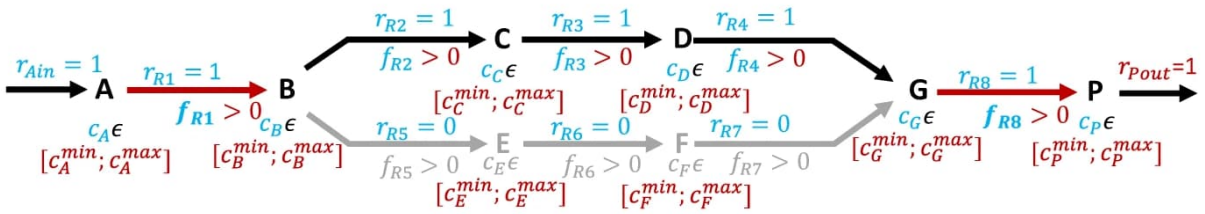
MDF of network (also called OptMDF): find flux distribution $\mathbf{r}$ maximizing MDF of active reactions

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



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

e.g., SubMDF = MDF(R_1, R_8) with given concentration ranges and flux constraints
and demanded minimal MDF (e.g., >0 kJ/mol) of other active reactions



(Noor et al., 2013)

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

(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

$$\underset{\mathbf{x},\mathbf{r},\mathbf{z}}{Maximize} \quad B$$

$$s.t.$$

$$\mathbf{Nr} = \mathbf{0}$$

$$\alpha_i \leq r_i \leq \beta_i$$

$$\ln(\mathbf{c_{min}}) \leq \mathbf{x} \leq \ln(\mathbf{c_{max}})$$

$$r_i \leq z_i \cdot \beta_i$$

$$z_i \in \{0,1\}$$

$$f_i = -\Delta_r G_i' = -\Delta_r G'^{\circ} - RT \cdot (\mathbf{N}_{*,i})^T \cdot \mathbf{x}$$

$$B \leq f_i + M \cdot (1 - z_i), \qquad (M\ very\ large)$$

SubMDF in a network (Bekiaris and Klamt, 2023)

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

Maximize MDF for selected reactions → $\underset{\mathbf{x},\mathbf{r},\mathbf{z},B}{Maximize} \quad B_{sub}$

$$s.t.$$

$$\mathbf{Nr} = \mathbf{0}.$$

$$\alpha_i \leq r_i \leq \beta_i$$

$$\ln(\mathbf{c_{min}}) \leq \mathbf{x} \leq \ln(\mathbf{c_{max}})$$

$$r_i \leq z_i \cdot \beta_i$$

$$z_i \in \{0,1\}$$

$$f_i = -\Delta_r G_i' = -\Delta_r G'^{\circ} - RT \cdot (\mathbf{N}_{*,i})^T \cdot \mathbf{x}$$

$$B \leq f_i + M \cdot (1 - z_i), \qquad (M\ very\ large)$$

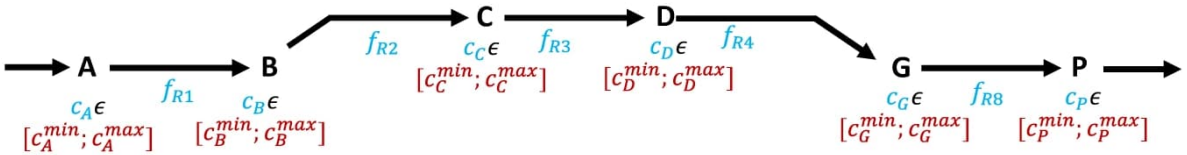
Thermodynamic feasibility of entire flux vector → $B \geq 0.1$

Minimum Driving force of the subset *S* of selected reactions → $B_{sub} \leq f_j + M \cdot (1 - z_j), \qquad \forall j \in S$

Max-min Driving Force (MDF) of a Pathway and Network

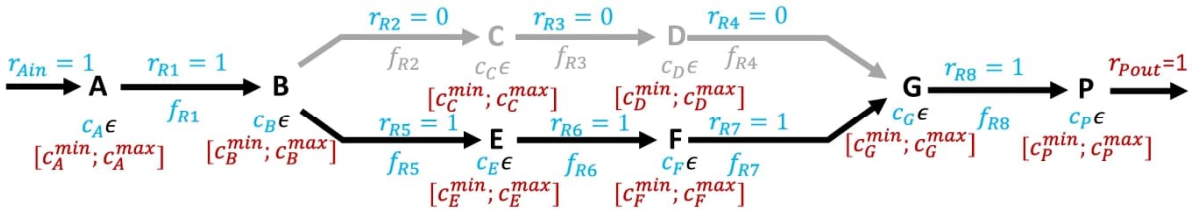
Max-min driving force (MDF) of a pathway: maximal achievable driving force of the pathway

$MDF_{Pathway} = \max(f_{Pathway}) = \max(\min(f_{R1}, f_{R2}, f_{R3}, f_{R4}, f_{R8}))$ within given concentration ranges



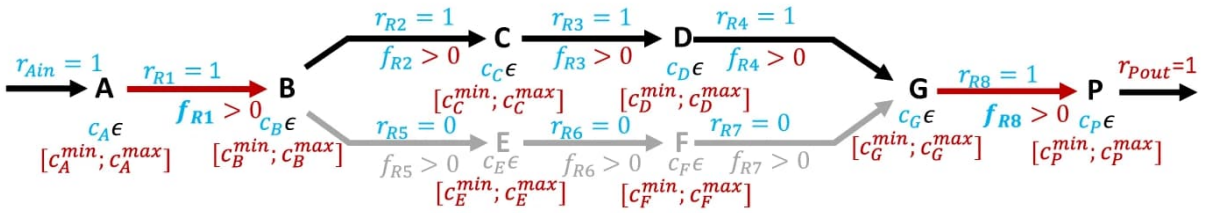
MDF of network (also called OptMDF): find flux distribution $\mathbf{r}$ maximizing MDF of active reactions

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



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

e.g., SubMDF = MDF(R_1, R_8) with given concentration ranges and flux constraints
and demanded minimal MDF (e.g., >0 kJ/mol) of other active reactions



(Noor et al., 2013)

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

(Hädicke et al., 2018)
→ Mixed-Integer Linear Program (MILP) within constraint-based metabolic model

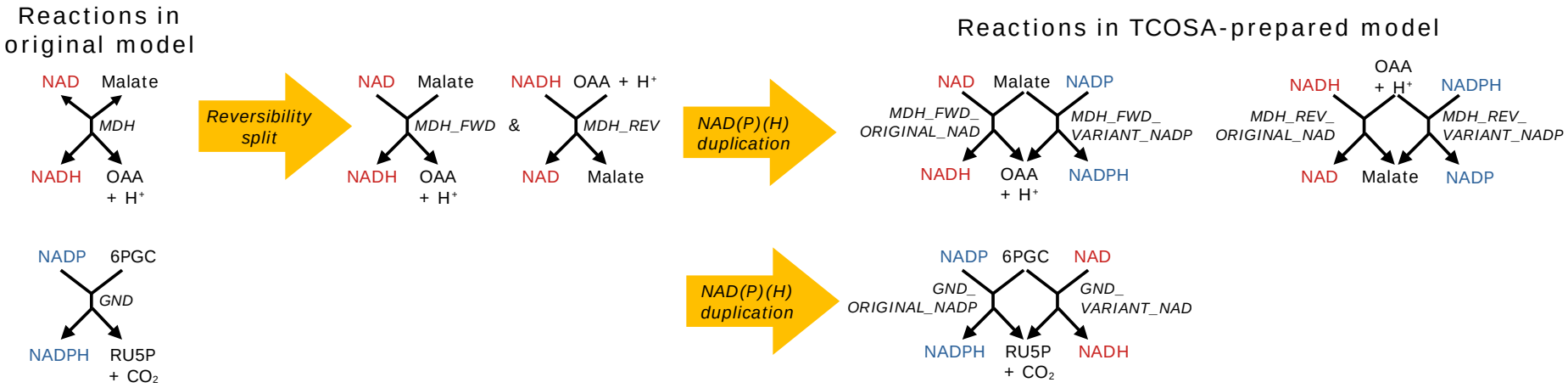
What NAD(P)(H) specificities maximize the MDF/SubMDF for growth-related flux distributions and how close is the wild-type specificity to this optimal specificty?

(This work, 2023)
→ Mixed-Integer Linear Program (MILP) within constraint-based metabolic model

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

TCOSA: Thermodynamic Cofactor Swapping Analysis

Reconfiguration of a given (stoichiometric) metabolic model for TCOSA:



Application to *E. coli*

Resulting model: iML1515_TCOSA (derived from genome-scale *E. coli* model *i*ML1515; Monk *et al.*, 2017).

- Substrate: glucose. Aerobic ($+O_2$) and anaerobic ($-O_2$) conditions.
- Metabolite concentration ranges: [10^{-6}… 0.02 M]
- $\Delta_rG'^{\circ}$ values: from eQuilibrator (Flamholz *et al.*, 2012) via its Python API (Beber *et al.*, 2021)

Computing (Sub)MDF with Different NAD(P) Specificity Scenarios

Wild-type specificity: Use original NAD(P)(H) specificity for the NAD(P)(H)-dependent reactions

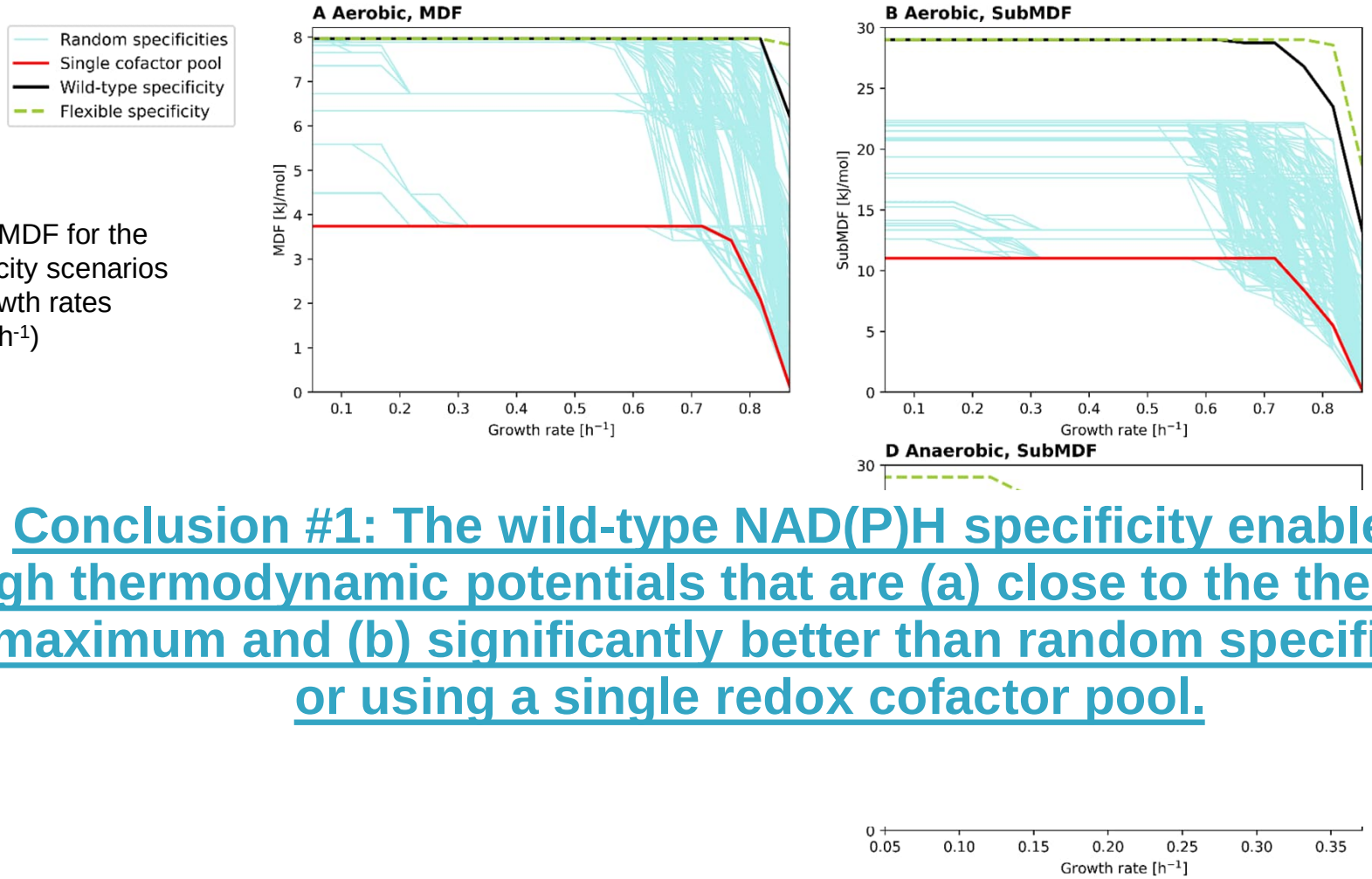
Flexible specificity: NAD(P)(H) specificity can be freely selected for each reaction (but only one at a time for each reaction)

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)

Analysis 1: (Sub)MDF Results for the Different Specificity Scenarios

Compute (Sub)MDF for the different specificity scenarios for different growth rates (step size 0.05 h^{-1})



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

Oxygen availability	Growth rate [h^{-1}]	Number of necessary swaps in wild type to reach (Sub)MDF of flexible specificity	
		MDF	SubMDF
Aerobic	0.868	6	2
Aerobic	0.818	0	3
Aerobic	0.768	0	2
Aerobic	0.718	0	2
Aerobic	0.668	0	0
Aerobic	0.618	0	0
Aerobic	0.568...0.518	0	0
Aerobic	0.468...0.118	0	0
Aerobic	0.068	0	0
Aerobic	0.05	0	0
Anaerobic	0.371	1	1
Anaerobic	0.321	0	0
Anaerobic	0.271	9	14
Anaerobic	0.221	0	2
Anaerobic	0.171	0	5
Anaerobic	0.121	0	4
Anaerobic	0.071	0	3
Anaerobic	0.05	0	3

Analysis 2: Necessary Swaps in Wild-type Specificity to Reach the Theoretical Maximal (Sub)MDF of the Flexible Specificity

Two frequently suggested cofactor swaps to increase (Sub)MDF:

1) Pyruvate dehydrogenase

CoA + pyruvate + NAD^+ → acetyl-CoA + CO_2 + NADH (ΔG'° of -34.37 kJ/mol)

⇨ CoA + pyruvate + $NADP^+$ → acetyl-CoA + CO_2 + NADPH

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

2) Isocitrate dehydrogenase

isocitrate + $NADP^+$ → 2-oxoglutarate + CO_2 + NADPH (ΔG'° of +5.13 kJ/mol)

⇨ isocitrate + NAD^+ → 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*:

$[NADH] << [NAD^+]$ $[NADPH] >> [NADP^+]$

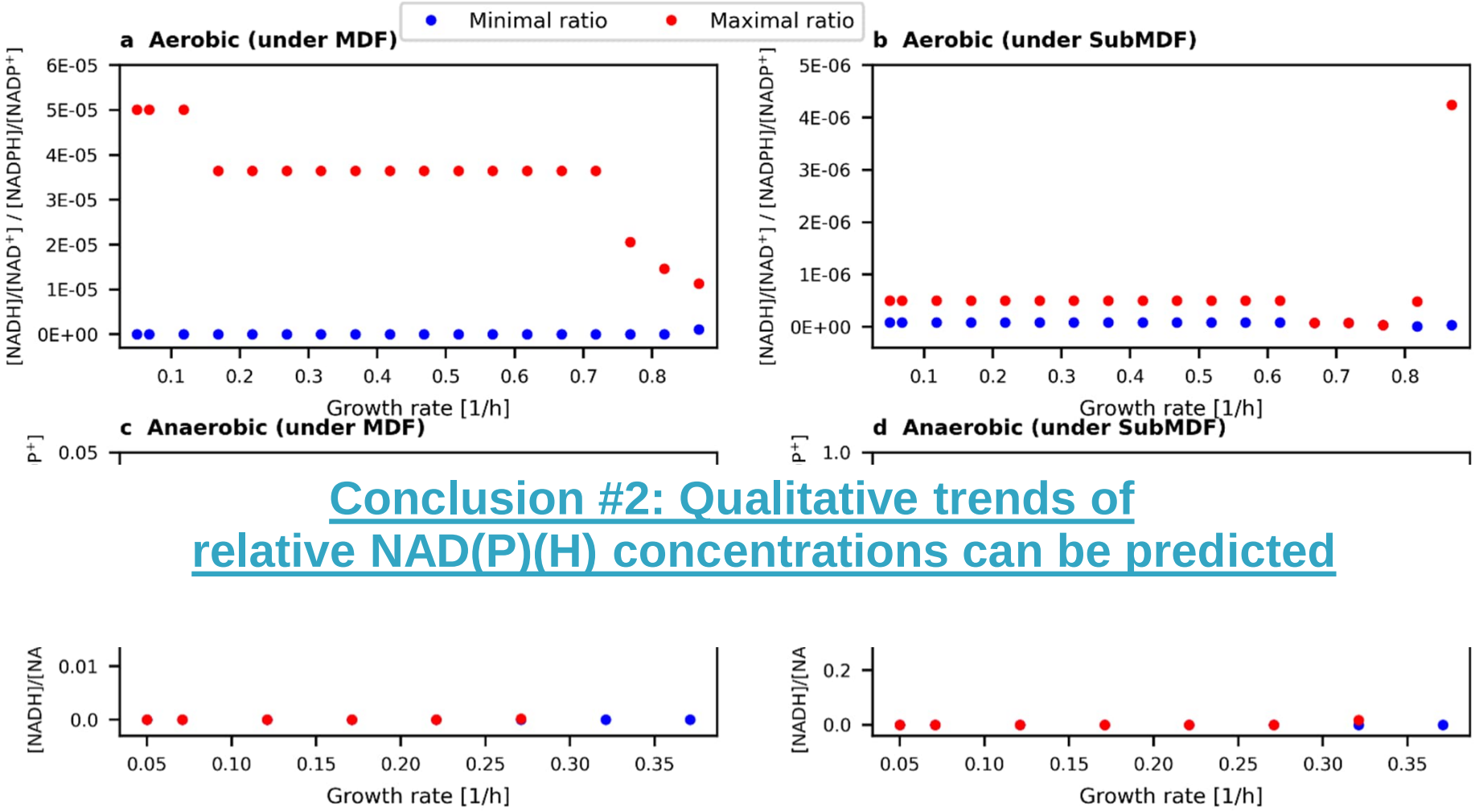
in vivo $NADH/NAD^+$ **ratio of ≈0.03** in *E. coli* (aerobic, Bennet et al., 2009)

in vivo $NADPH/NADP^+$ **ratio of ≈57** in *E. coli* (aerobic, Bennet et al., 2009)

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

in vivo **≈0.00053** in *E. coli* (aerobic, Bennet et al., 2009)

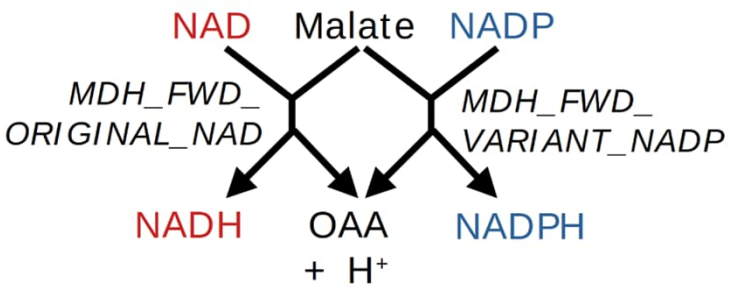
Analysis 3: Minimal/Maximal Concentration Ratios at Optimal (Sub)MDF (Wildtype Specificity)



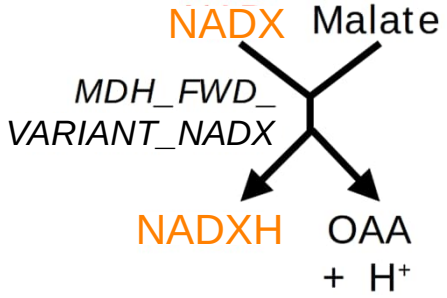
Conclusion #2: Qualitative trends of relative NAD(P)(H) concentrations can be predicted

Analysis 4: Effect of a Third Redox Cofactor Pool (Flexible Specificity)

NAD(H) NADP(H)



"NADX(H)"



NAD → NADH

Δ_fG'° difference of ca. 61 kJ/mol → Standard redox potential (ΔE'°) of -320 mV

NADP → NADPH

3 redox potential scenarios:

1) As for NAD(P)(H)

NADX → **ΔE'°=-320 mV** → NADXH

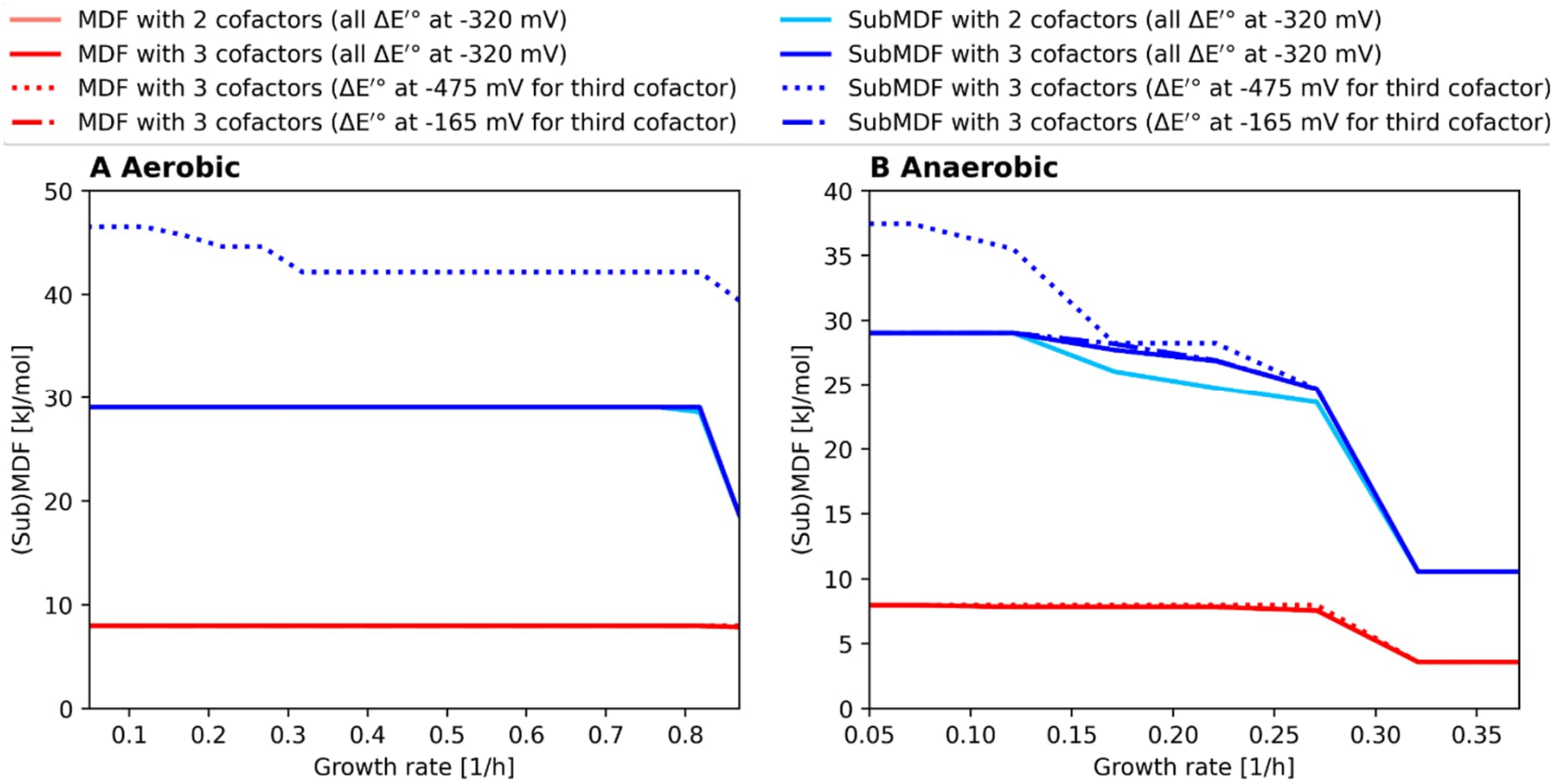
2) Lower potential

NADX → **ΔE'°=-475 mV** → NADXH

3) Higher potential

NADX → **ΔE'°=-165 mV** → NADXH

Analysis 4: (Sub)MDF Results with 3 Cofactors (Flexible Specificity)

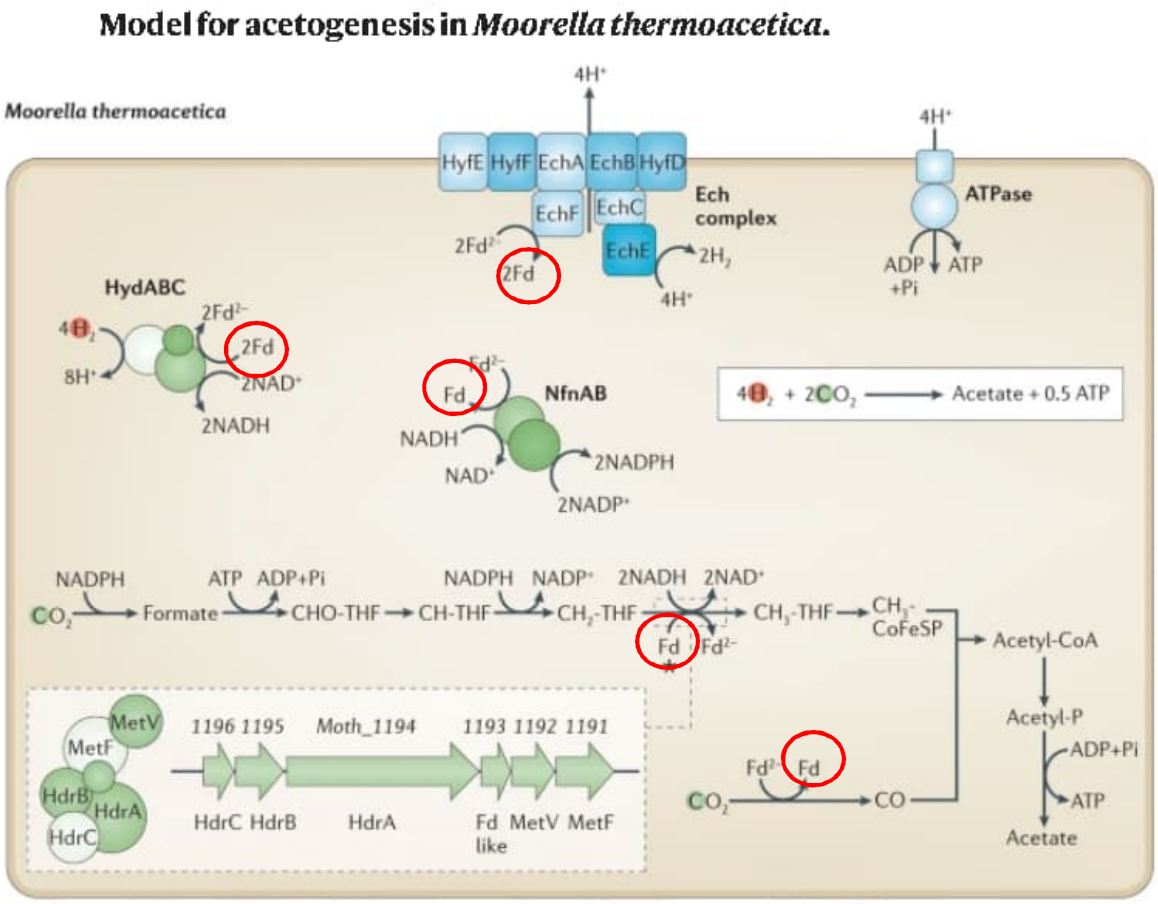


Conclusion #3: A third redox cofactor pool could be advantageous if it has a low standard redox potential!

Analysis 4: (Sub)MDF Results with 3 Cofactors (Flexible Specificity)

Several autotrophic organisms like acetogens use ferredoxin (ΔE'° 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.

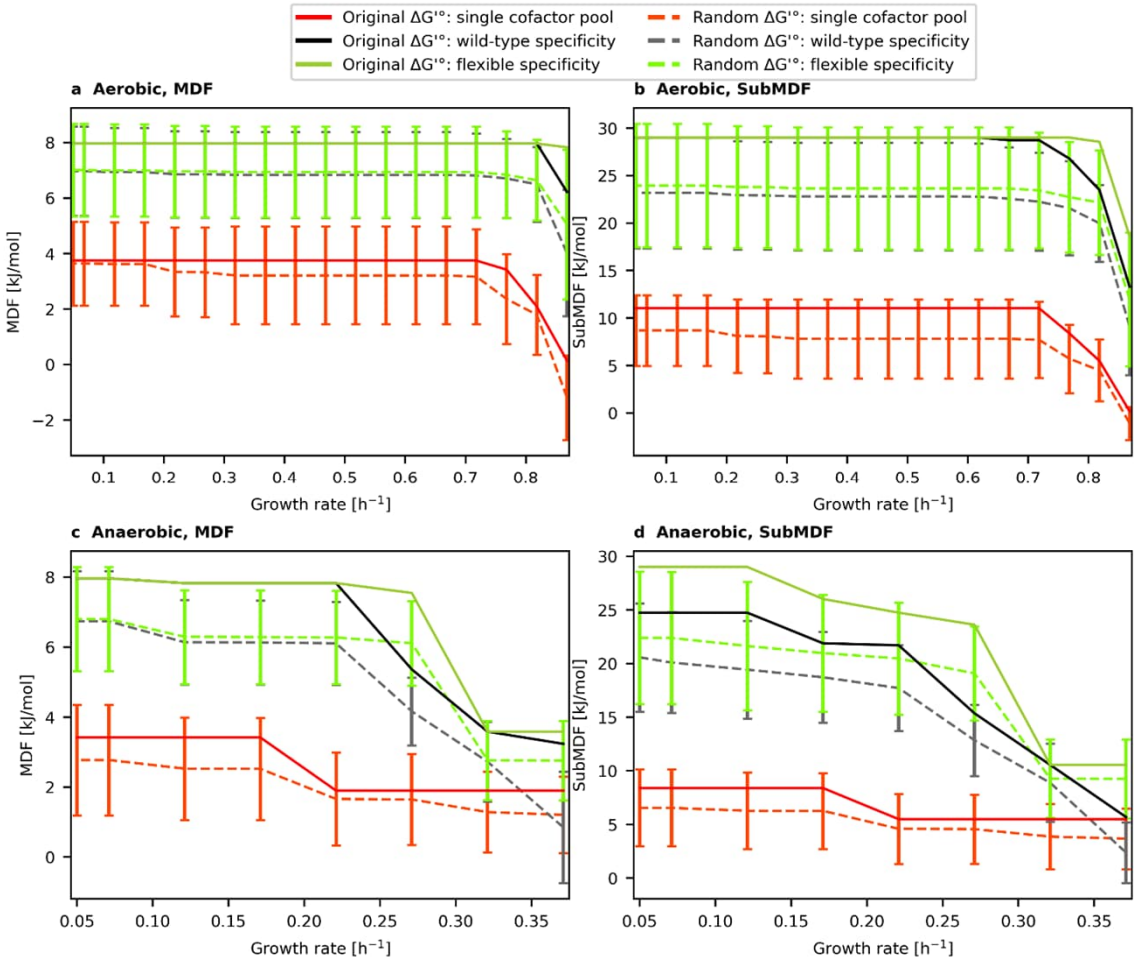


Schuchmann et al. *Nat Rev Microbiol* **12**, 809–821 (2014).

Analysis 5: Robustness of the Results

A) Robustness against random variations of $\Delta_r G'°$

(implemented by random variations of the $\Delta_f G'°$ of each metabolite)

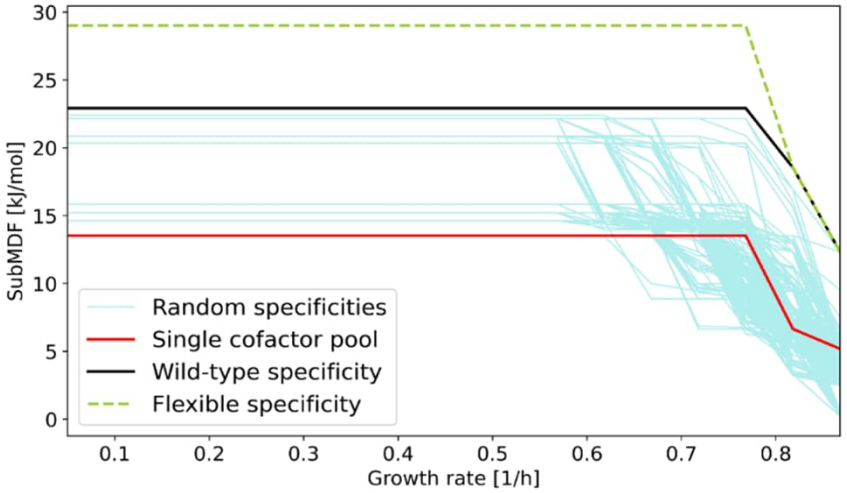


Analysis 5: Robustness of the Results

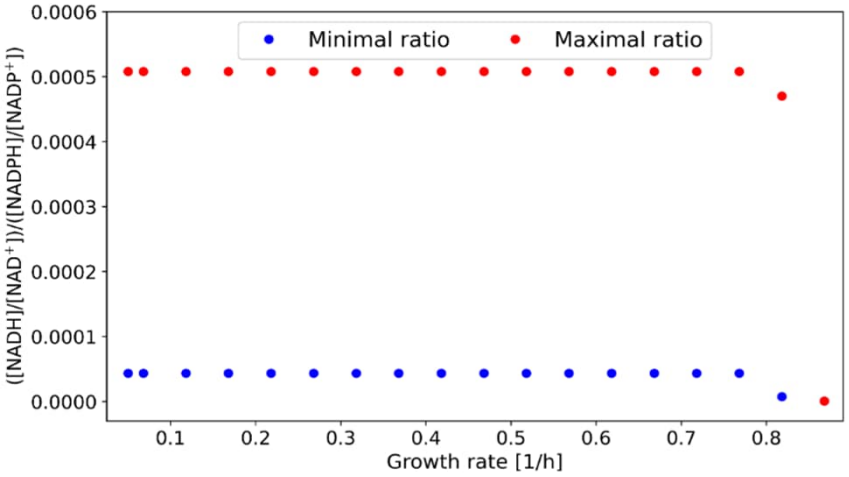
B) Robustness against assumed metabolite concentration ranges

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

SubMDF (aerobic)



Ratios under maximal SubMDF (aerobic)



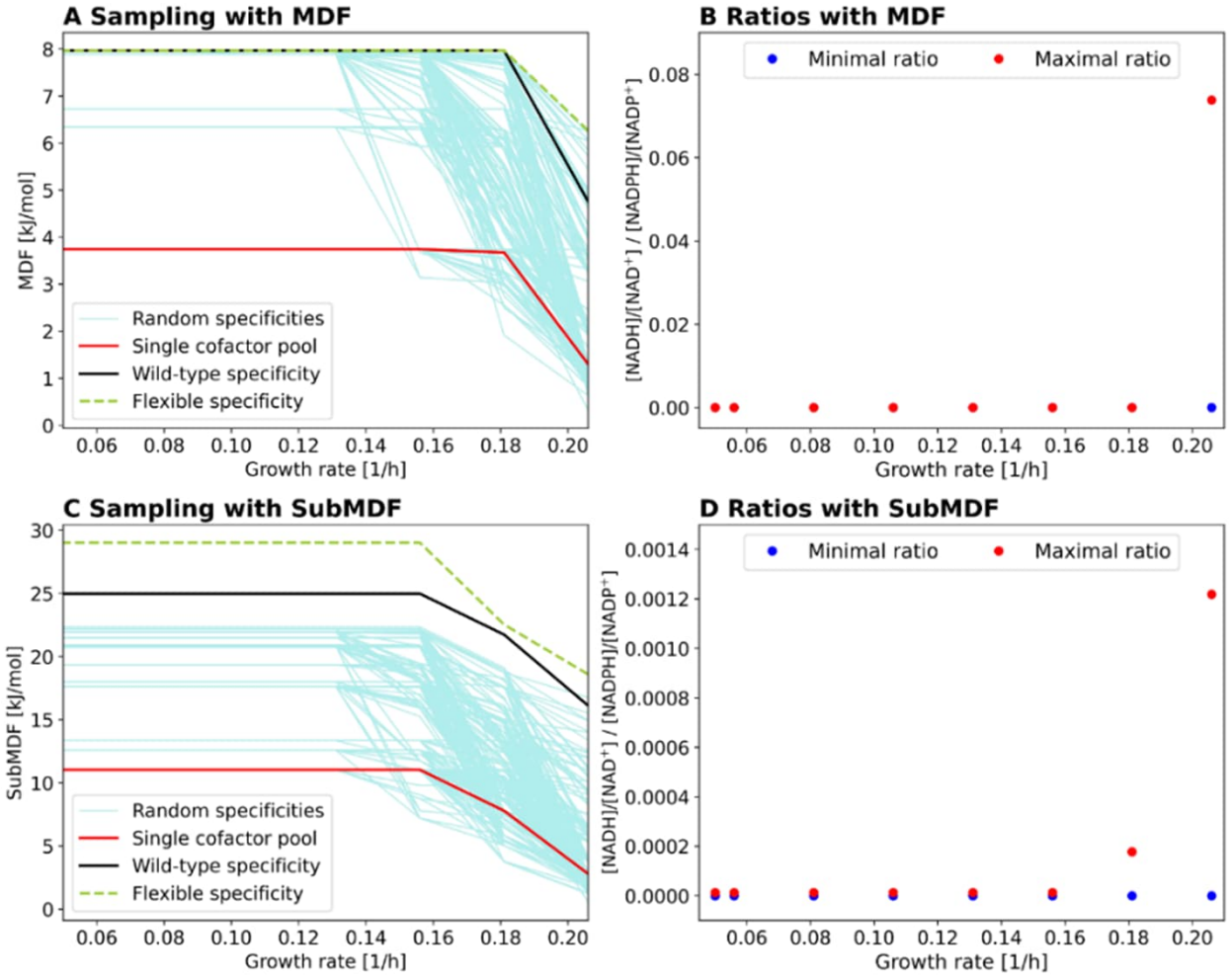
For MDF: single bottleneck
(independent of NAD(P(H) specificities)

Analysis 5: Robustness of the Results

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

MPI Magdeburg
Research Group Analysis and Redesign of Biological Networks

Thank you for your attention!

*Bekiaris PS, Klamt S (2023) Network-wide Thermodynamic Constraints Shape NAD(P)H Cofactor Specificity of Biochemical Reactions. Nature Communications **14**:4660.*