THERMODYNAMICS BASED FLUX BALANCE ANALYSIS

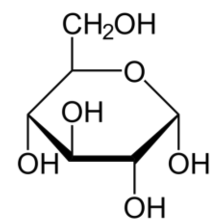
Principles and Applications of Systems Biology

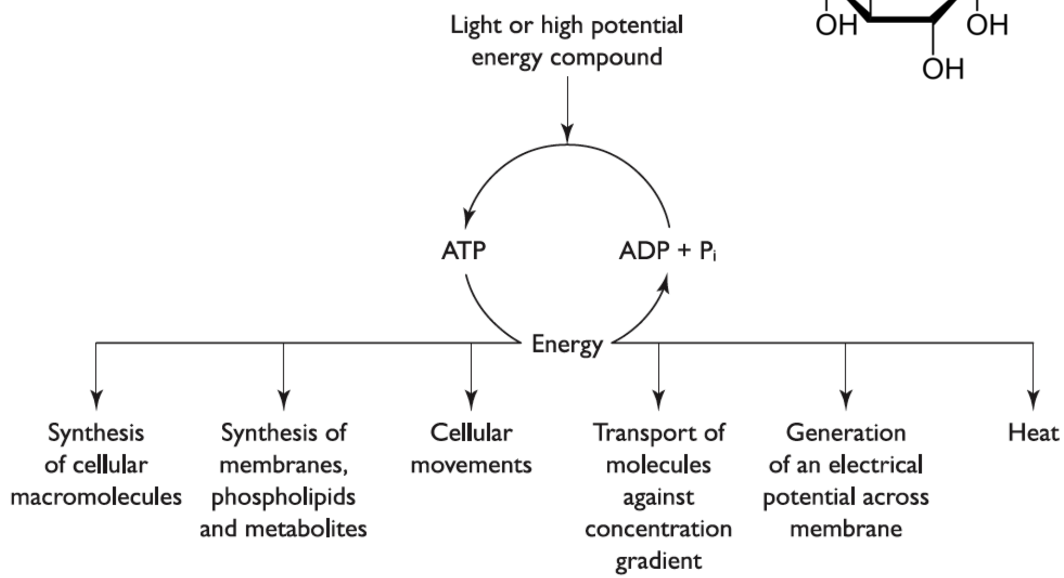
EPFL

Vassily Hatzimanikatis
September 2025

Bioenergetics

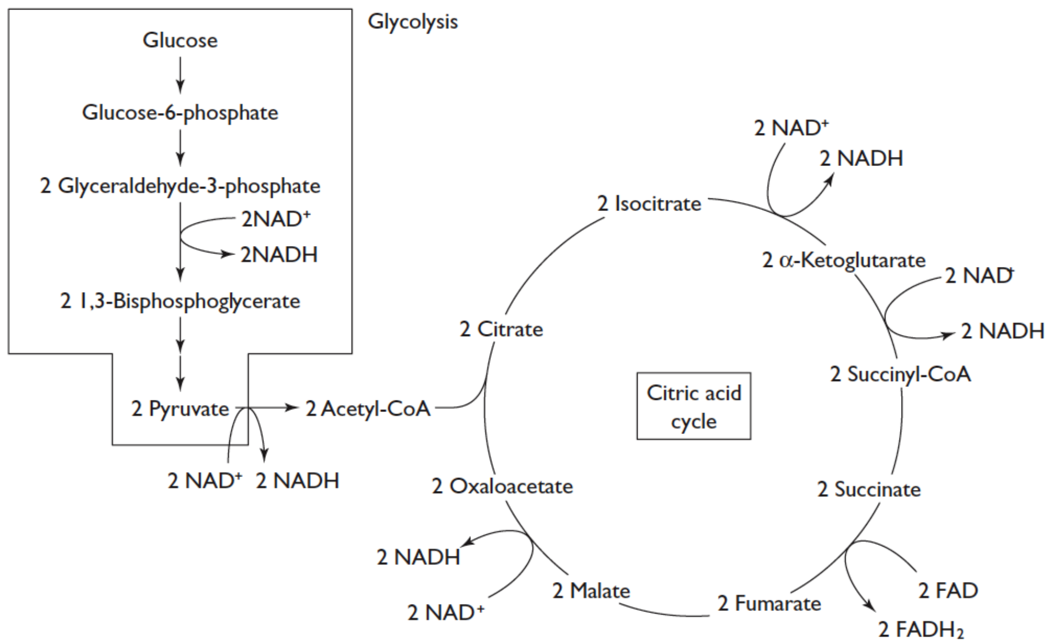
Bioenergetics





From D. T. Haynie “Biological Thermodynamics”

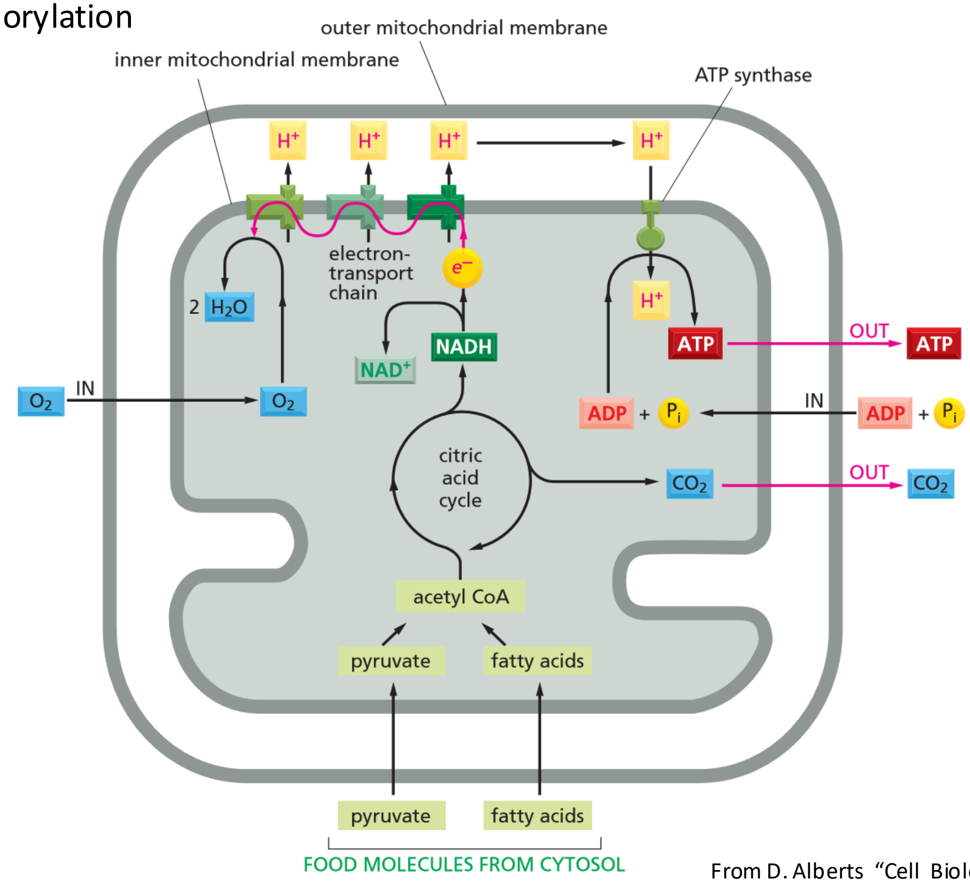
Bioenergetics



From D. T. Haynie "Biological Thermodynamics"

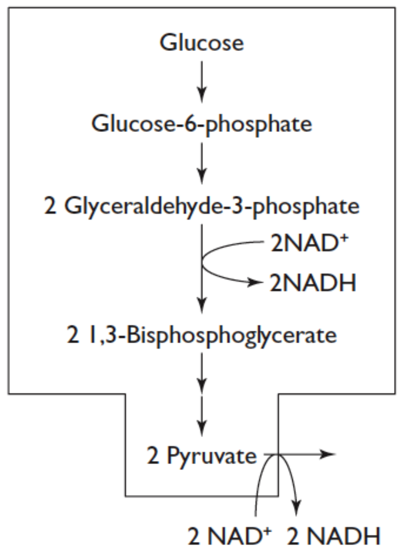
Bioenergetics

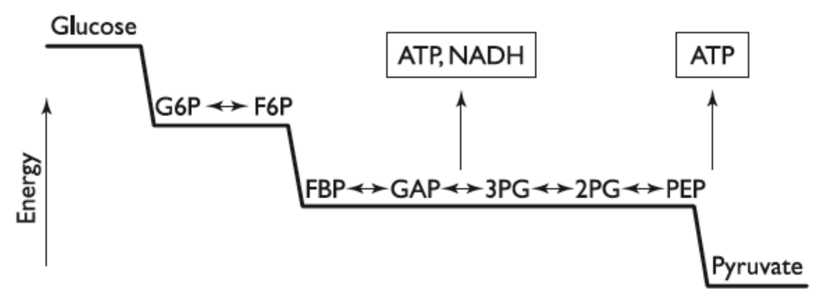
Oxidative phosphorylation



From D. Alberts "Cell Biology Essentials"

Bioenergetics





From D. T. Haynie “Biological Thermodynamics”

Thermodynamics

do you remember ΔG?

Thermodynamics

$$\Delta G^{\circ}$$

$$\Delta G'$$

$$\Delta G^{\circ\prime}$$

What is the difference?

Thermodynamics

$$A + B \xleftrightarrow{\Delta G' > 0} D + C$$

?

$$v_{A,B \rightarrow \ CD} > 0 \qquad v_{A,B \rightarrow \ CD} < 0$$

Thermodynamics

$$A + B \xleftrightarrow{\Delta_r G' > 0} D + C$$

$$\Delta_r G' = \Delta_r G'^{\circ} + RT\left(\ln x_C + \ln x_D - \ln x_A - \ln x_B\right)$$

Activity of C $\sim$ Concentration $[\boldsymbol{C}]$

If $\Delta_r G' < 0$:

$$v_{A,B \rightarrow CD} > 0$$

If $\Delta_r G' > 0$:

$$v_{A,B \rightarrow CD} < 0$$

Thermodynamic constraint

if $\Delta_r G_i^{'} \geq 0$, v_i (reaction rate i) ≤ 0

if $\Delta_r G_i^{'} \leq 0$, v_i (reaction rate i) ≥ 0

with

$$\Delta_r G_i^{'} = \Delta_r G_i^{'\circ} + RT \sum_{j=1}^{m} n_{i,j} \ln(x_j)$$

$\Delta_r G^{'\circ}$ = standard Gibbs free energy change of reaction at 298K, pH 7, and zero ionic strength

$n_{i,j}$ = stoichiometric coefficient of metabolite j in reaction i

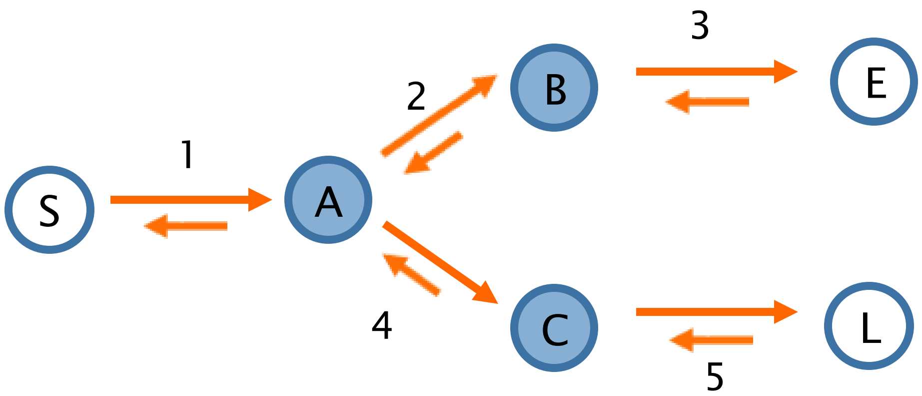
m = number of metabolites in the metabolic model

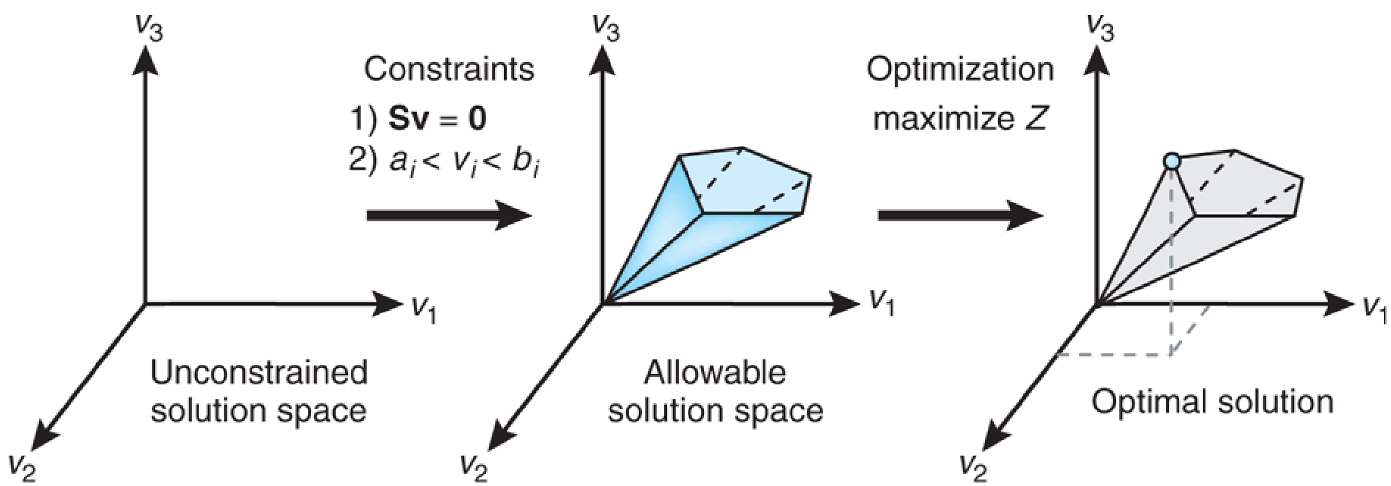
x_j = activity of metabolite j

Thermodynamics based Flux Balance Analysis

heading in the right direction

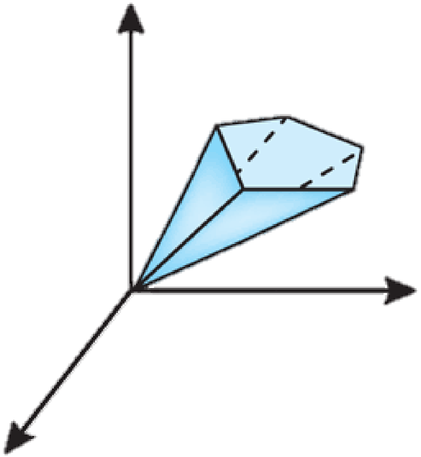
Flux Balance Analysis

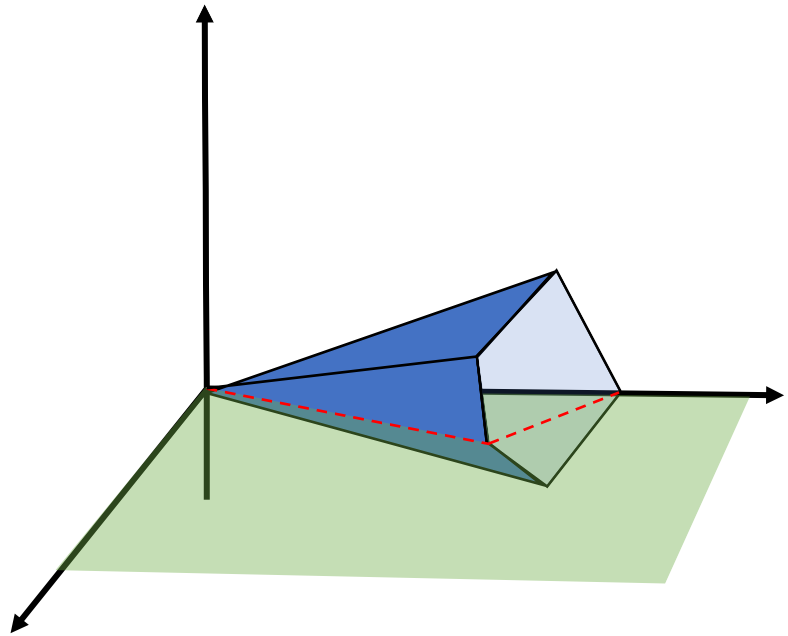




Orth, Jeffrey D., Ines Thiele, and Bernhard Ø. Palsson. "What is flux balance analysis?." *Nature biotechnology* 28.3 (2010): 245-248.

Flux Balance Analysis

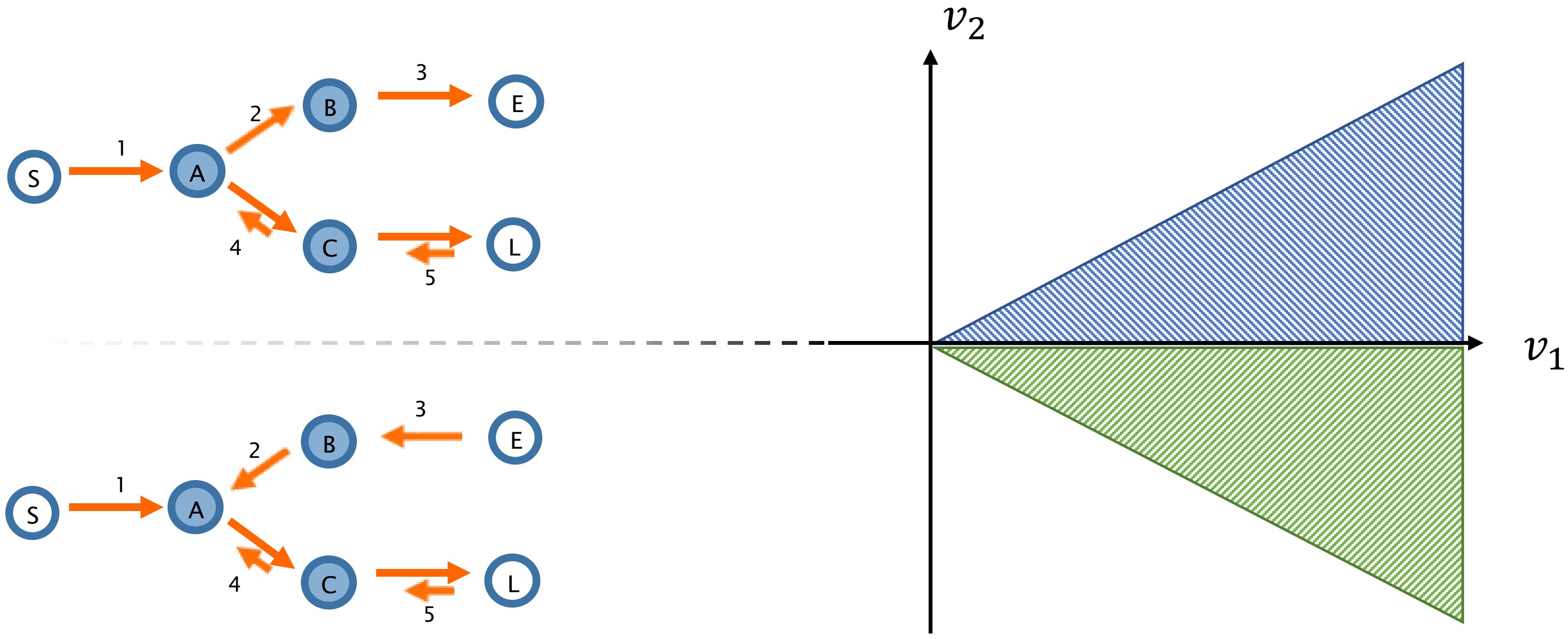




Orth, Jeffrey D., Ines Thiele, and Bernhard Ø. Palsson. "What is flux balance analysis?." *Nature biotechnology* 28.3 (2010): 245-248.

Thermodynamics based Flux Balance Analysis

- Directionality of the fluxes is physiology



Separating reversible reactions

Reversible Reaction:

$$A + B \longleftrightarrow C + D$$

$$-v_{Max} \leq v_{Reversible} \leq v_{Max}$$

Separated Reactions:

$$A + B \longrightarrow C + D$$

$$0 \leq v_F \leq v_{Max}$$

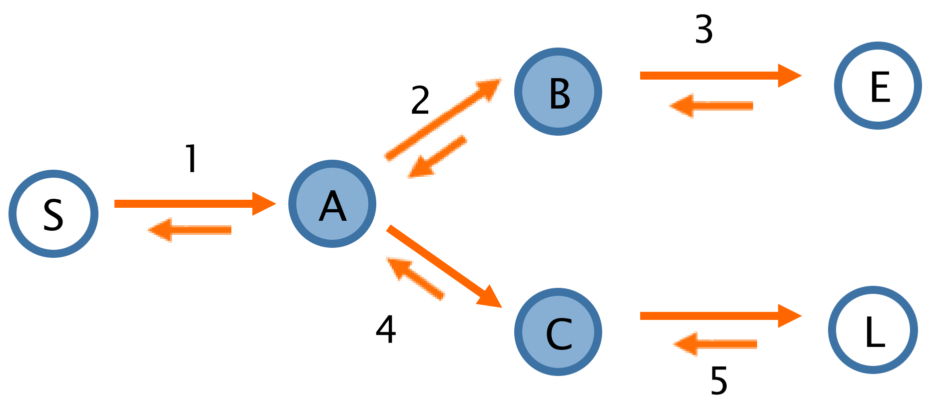
$$\Delta_r G_F' < 0 \text{ if } v_F > 0$$

$$A + B \longleftarrow C + D$$

$$0 \leq v_B \leq v_{Max}$$

$$\Delta_r G_B' < 0 \text{ if } v_B > 0$$

Thermodynamics based Flux Balance Analysis



Reaction 2 is forward

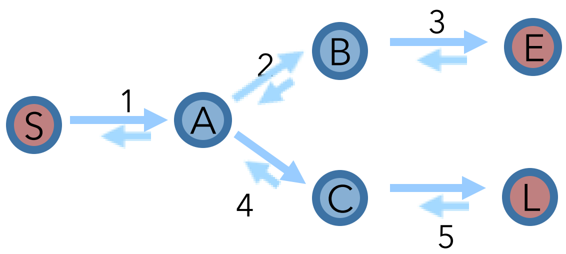
$$\Leftrightarrow \Delta G_2 \leq 0$$

$$\Leftrightarrow \Delta G_2^\circ + RT \ln \frac{B}{A} \leq 0$$

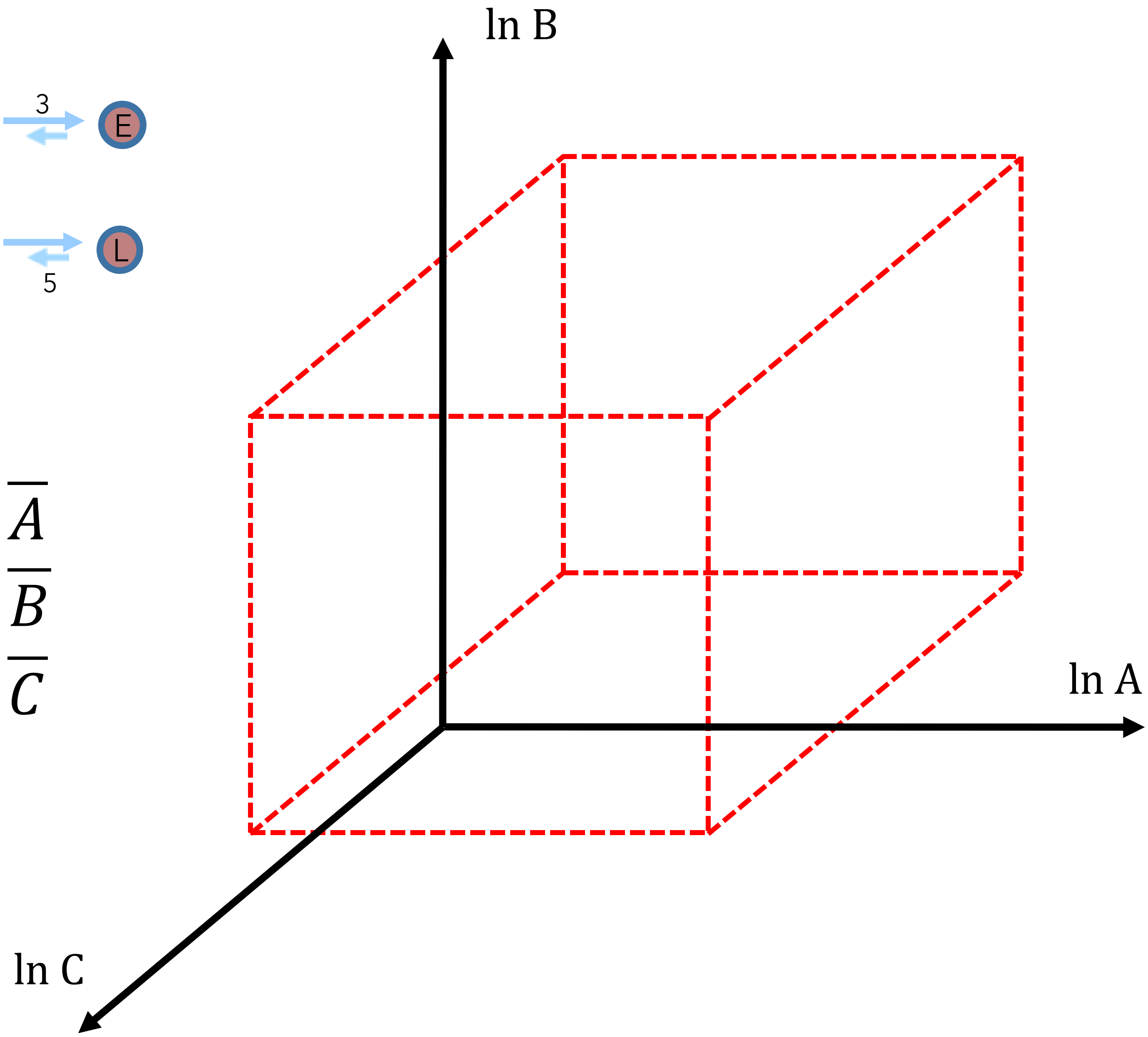
$$\Leftrightarrow \ln A \geq \ln B + \frac{\Delta G_2^\circ}{RT}$$

TFA adds
concentration of metabolites as variables, and **extra constraints** on directionalities, hence on **possible physiologies**

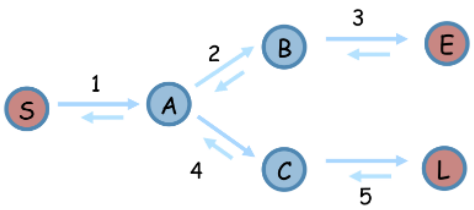
Thermodynamic constraints



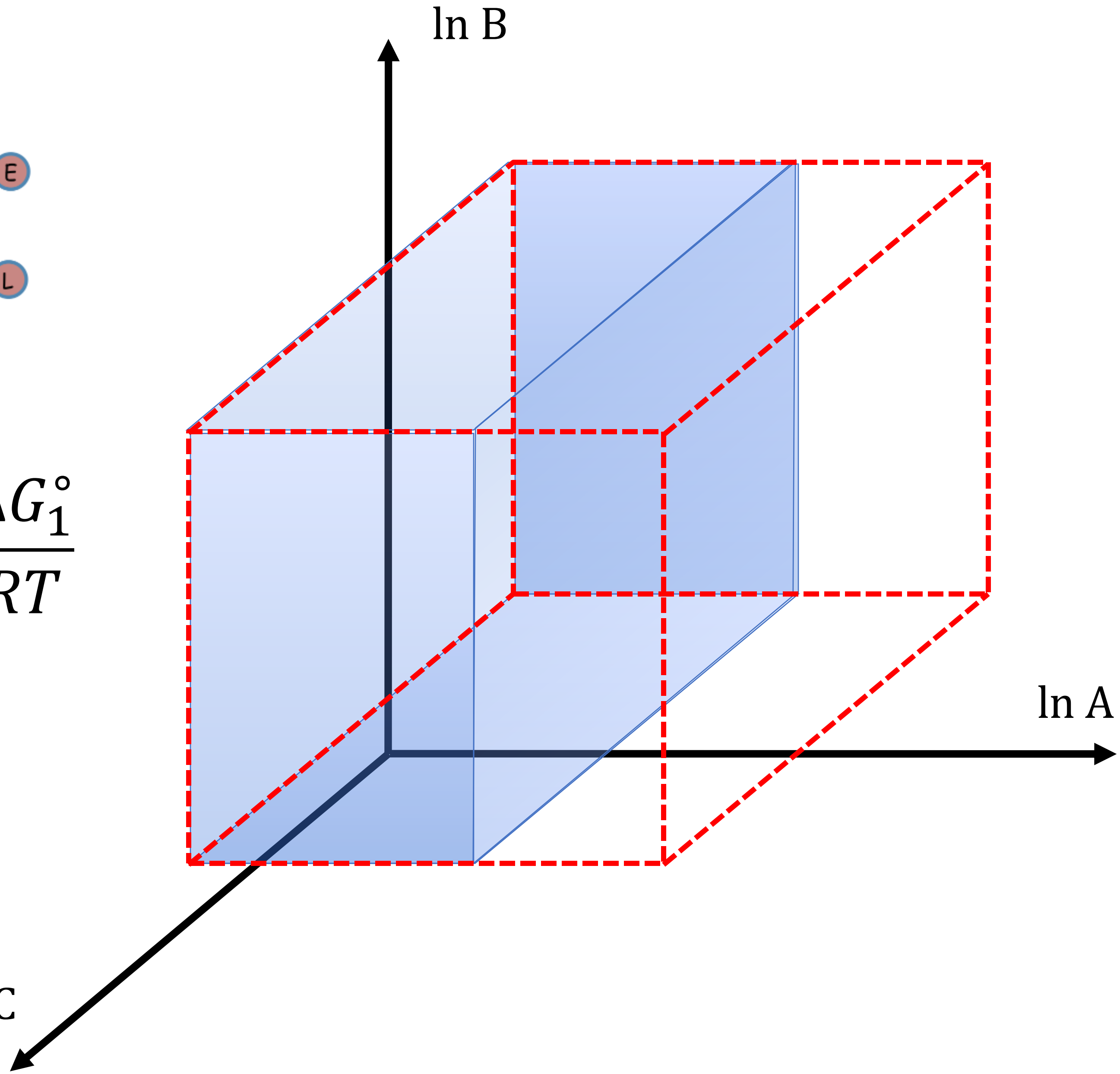
$$\underline{A} < \ln A < \overline{A}$$
$$\underline{B} < \ln B < \overline{B}$$
$$\underline{C} < \ln C < \overline{C}$$



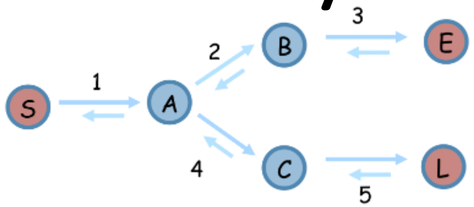
Thermodynamic constraints



$$\ln A < \ln S - \frac{\Delta G_1^\circ}{RT}$$

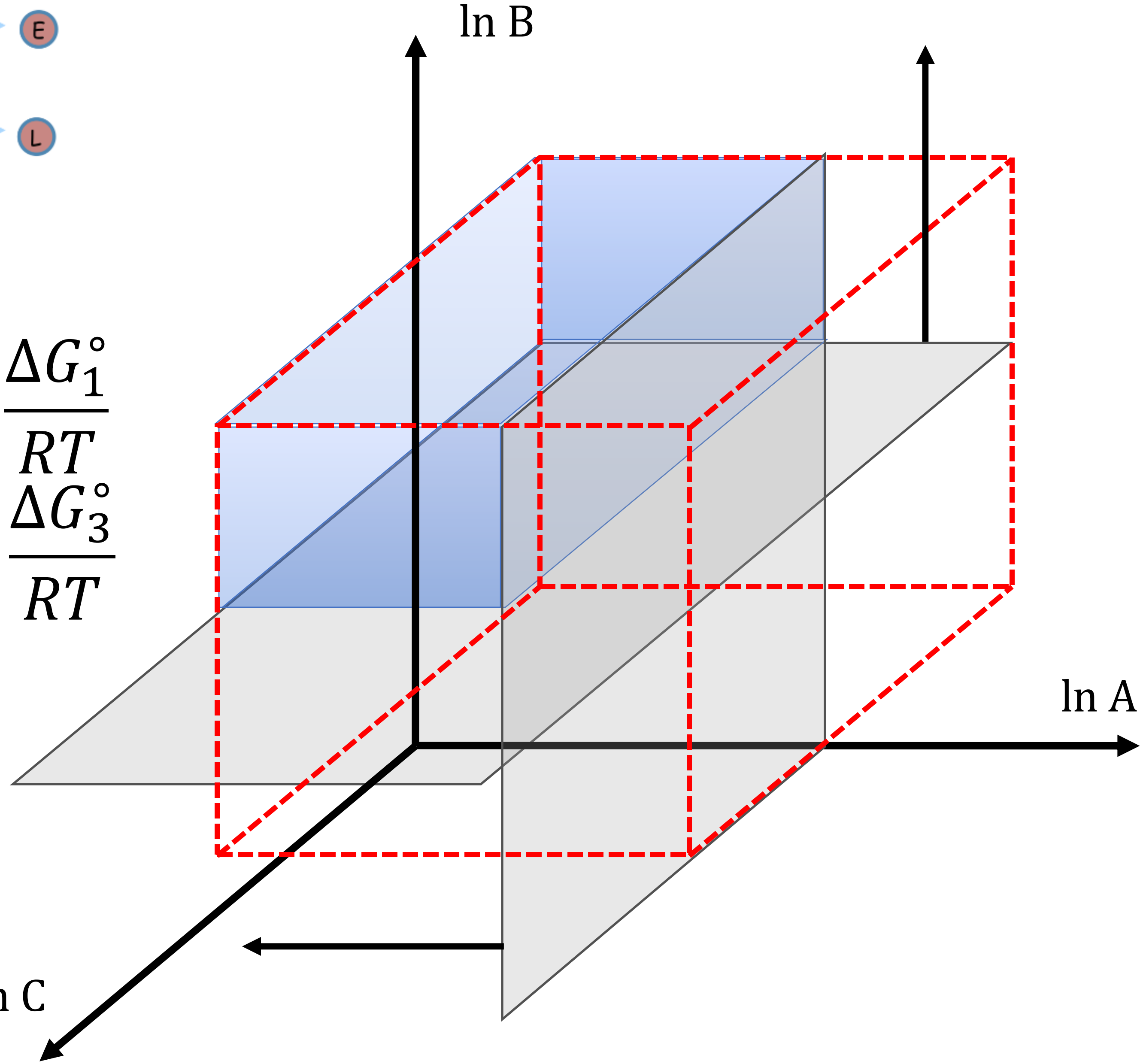


Thermodynamic constraints

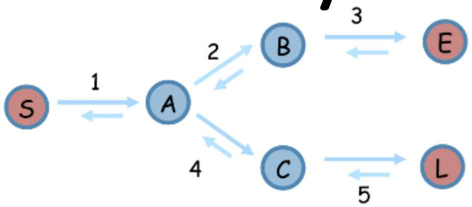


$$\ln A < \ln S - \frac{\Delta G_1^\circ}{RT}$$

$$\ln B > \ln E + \frac{\Delta G_3^\circ}{RT}$$



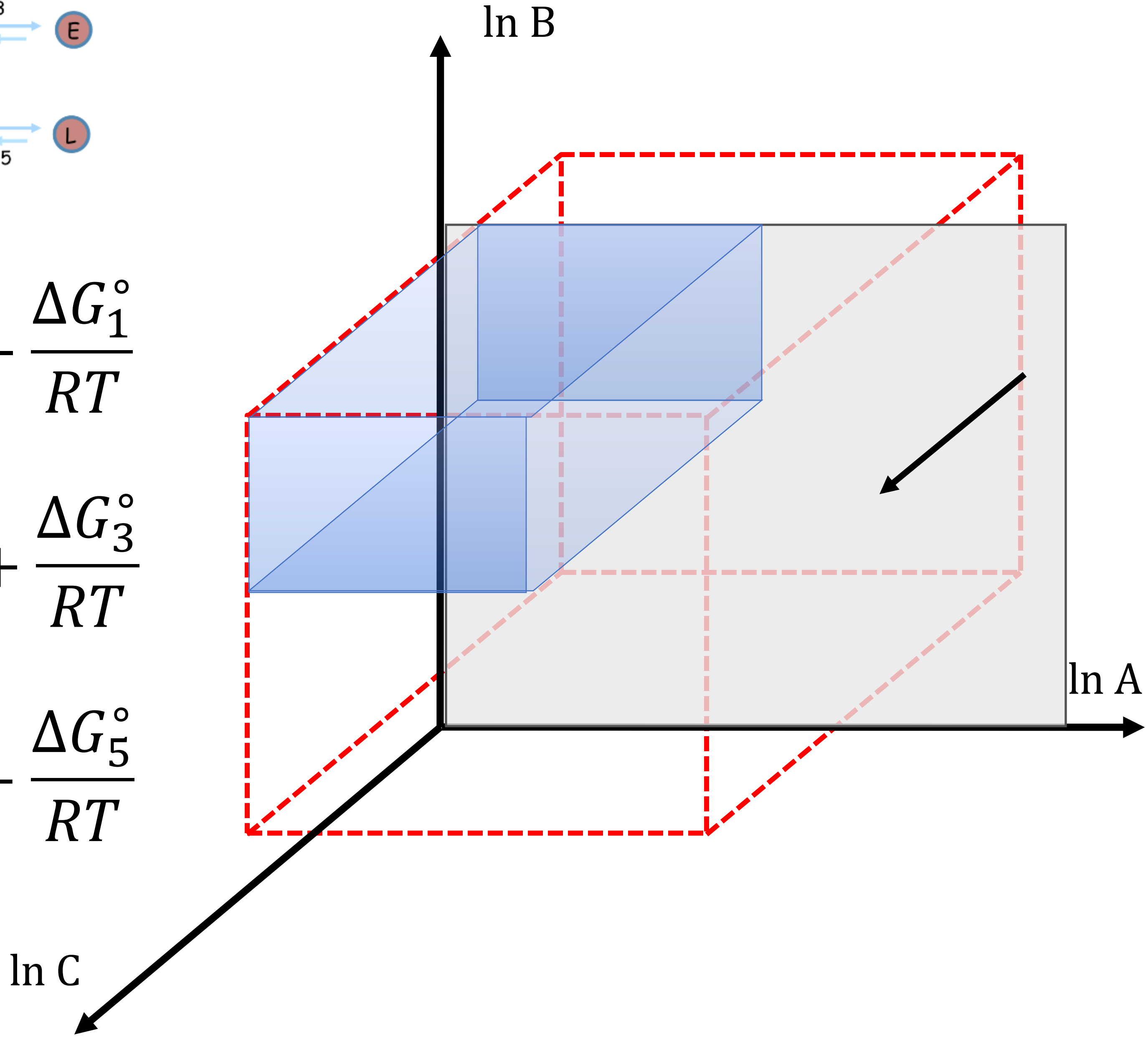
Thermodynamic constraints



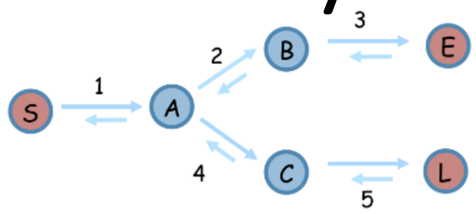
$$\ln A < \ln S - \frac{\Delta G_1^\circ}{RT}$$

$$\ln B > \ln E + \frac{\Delta G_3^\circ}{RT}$$

$$\ln C > \ln L + \frac{\Delta G_5^\circ}{RT}$$



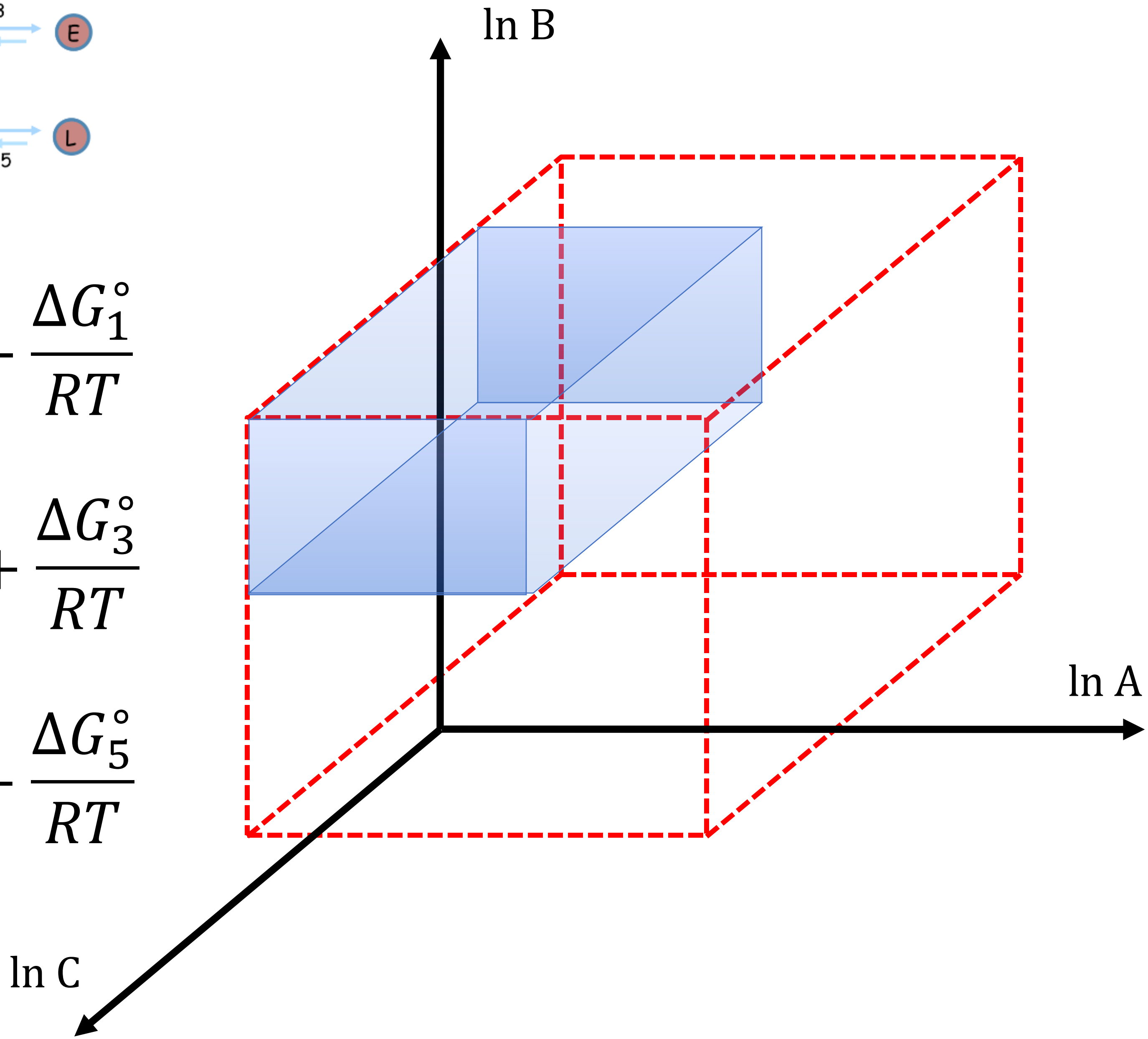
Thermodynamic constraints



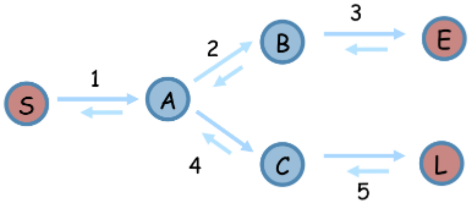
$$\ln A < \ln S - \frac{\Delta G_1^\circ}{RT}$$

$$\ln B > \ln E + \frac{\Delta G_3^\circ}{RT}$$

$$\ln C > \ln L + \frac{\Delta G_5^\circ}{RT}$$

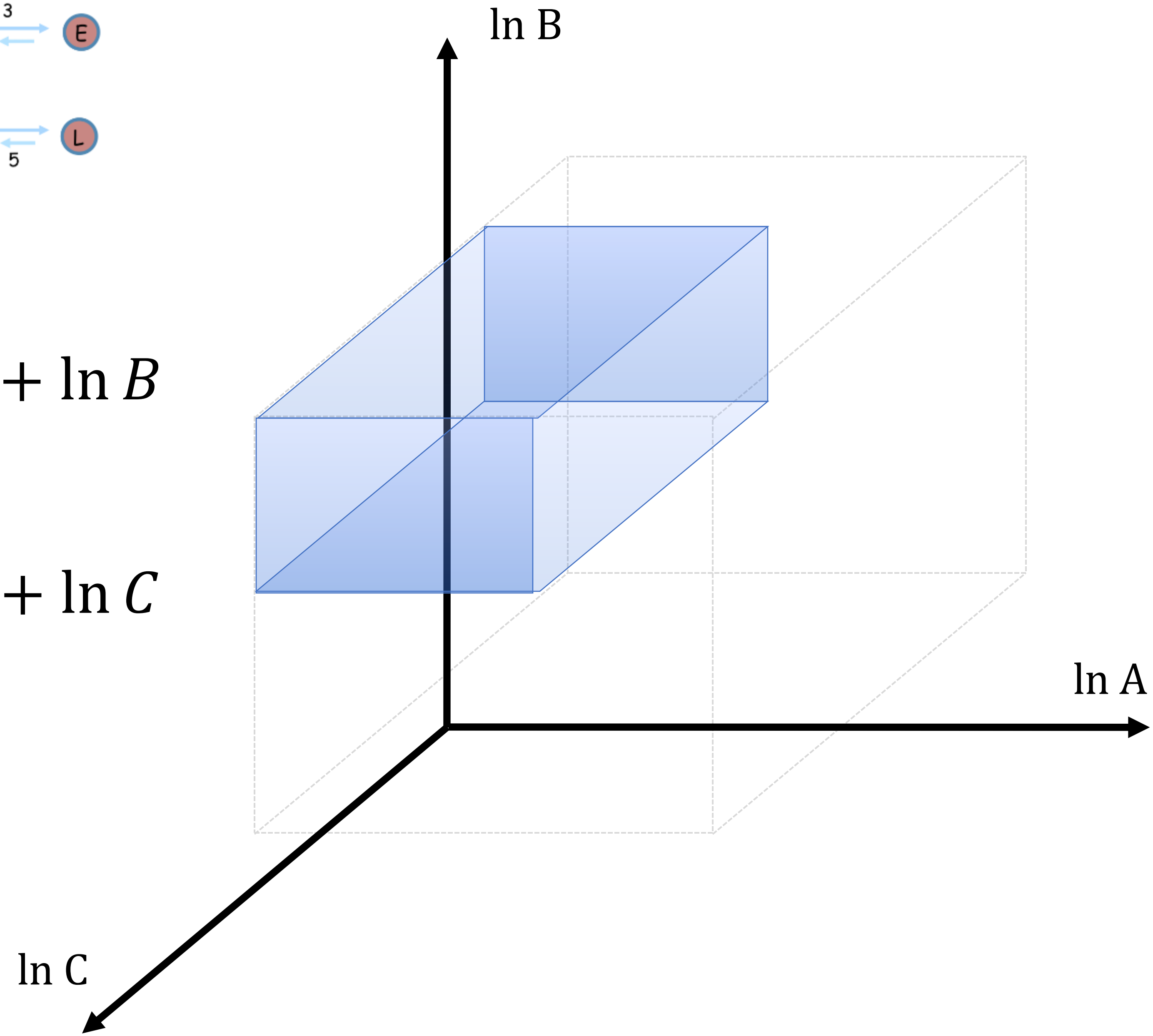


Thermodynamic constraints

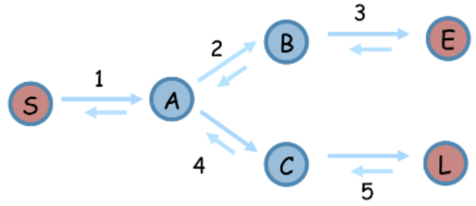


$$\ln A > \frac{\Delta G_2^\circ}{RT} + \ln B$$

$$\ln A > \frac{\Delta G_4^\circ}{RT} + \ln C$$

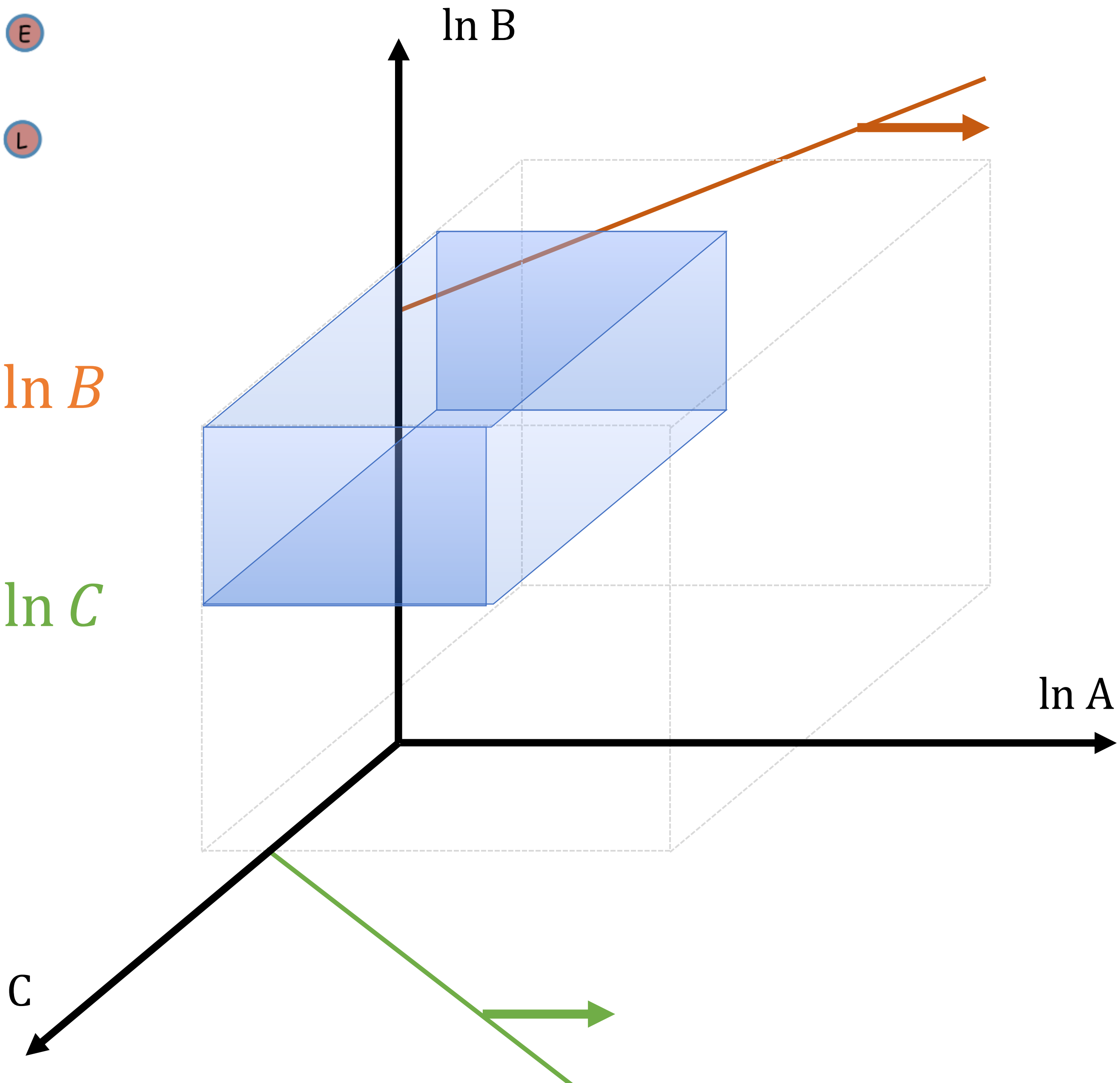


Thermodynamic constraints

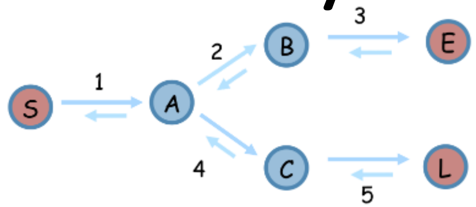


$$\ln A > \frac{\Delta G_2^\circ}{RT} + \ln B$$

$$\ln A > \frac{\Delta G_4^\circ}{RT} + \ln C$$

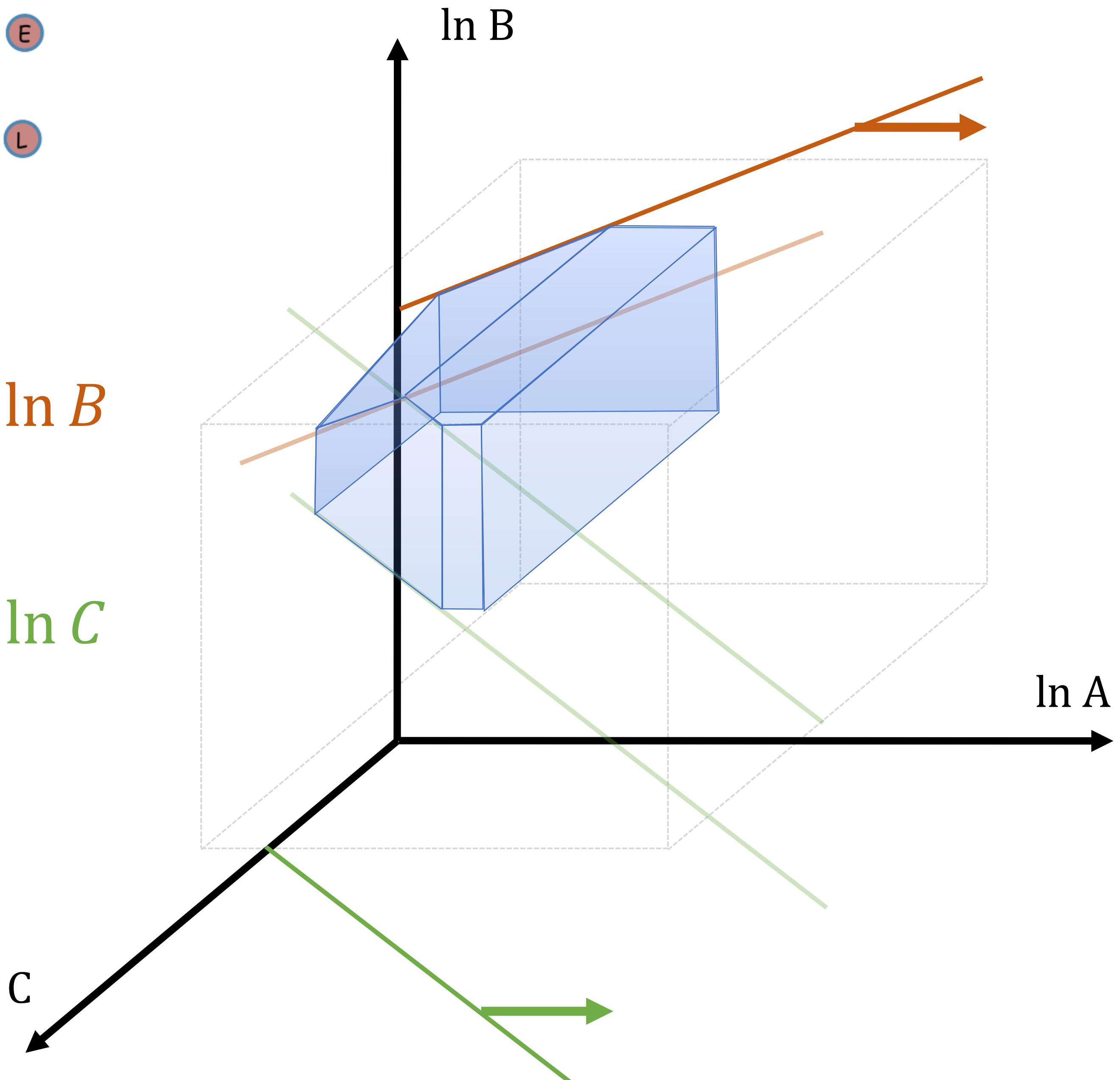


Thermodynamic constraints

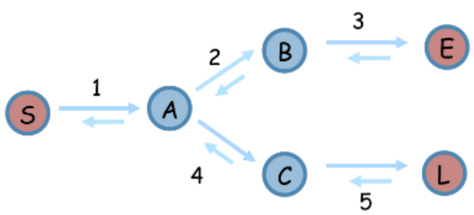


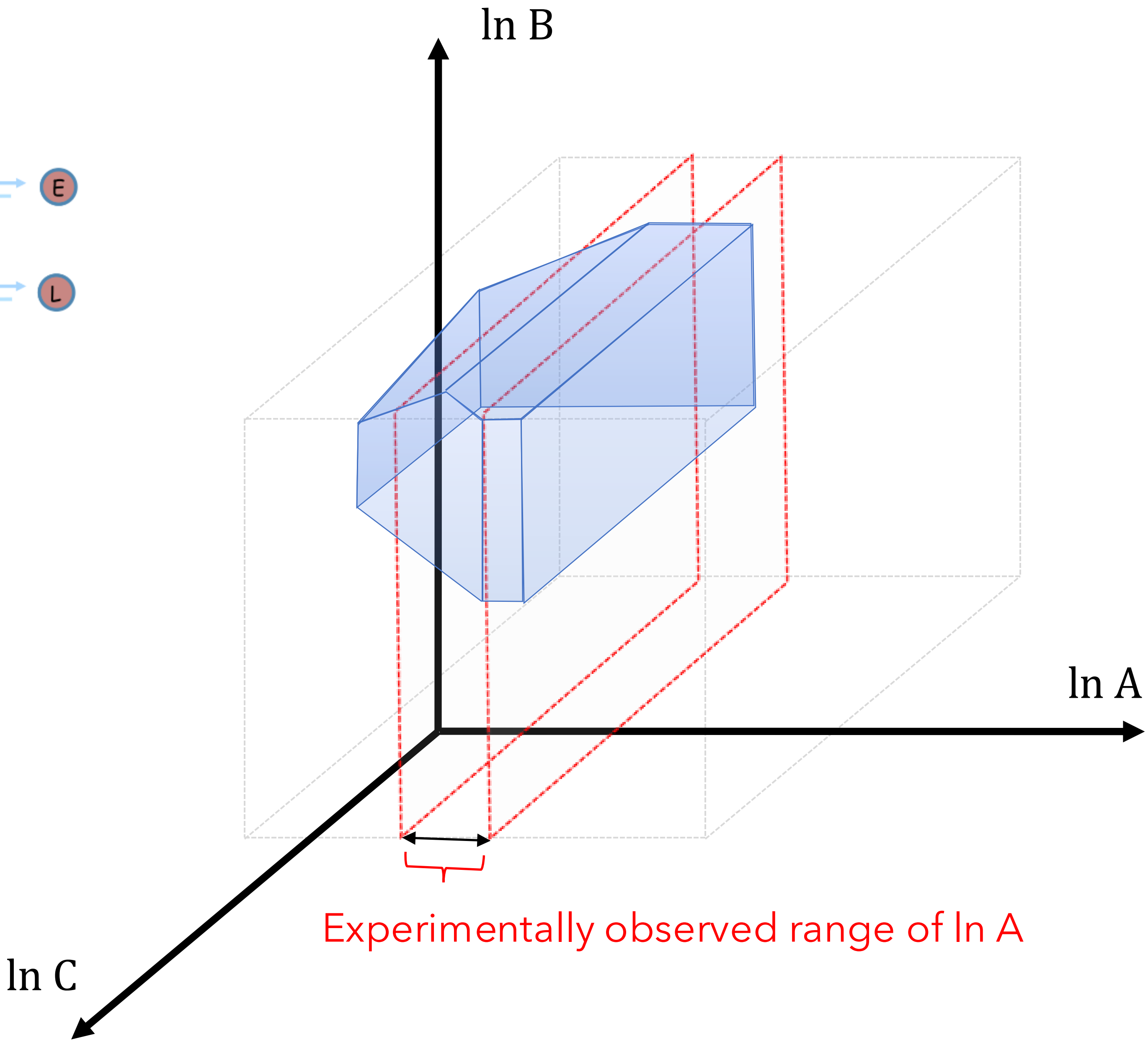
$$\ln A > \frac{\Delta G_2^\circ}{RT} + \ln B$$

$$\ln A > \frac{\Delta G_4^\circ}{RT} + \ln C$$

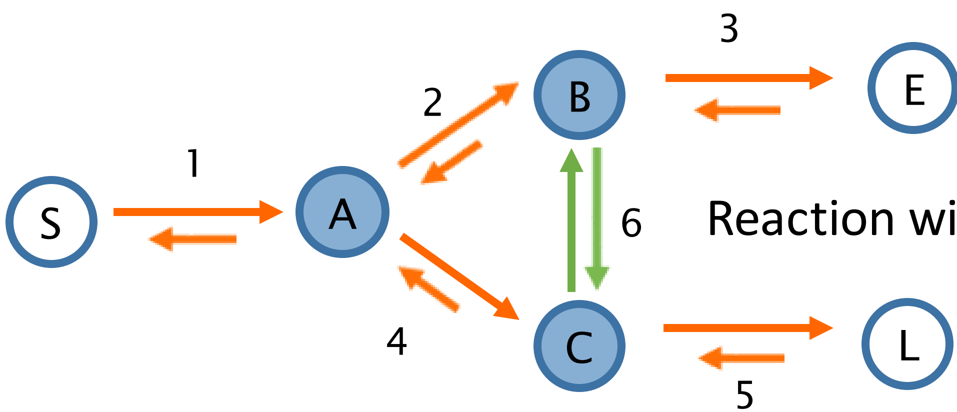


Metabolomics data





Bi-directional reactions

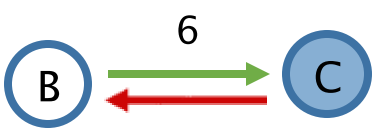


Reaction without known directionality

Thermodynamic feasibility:

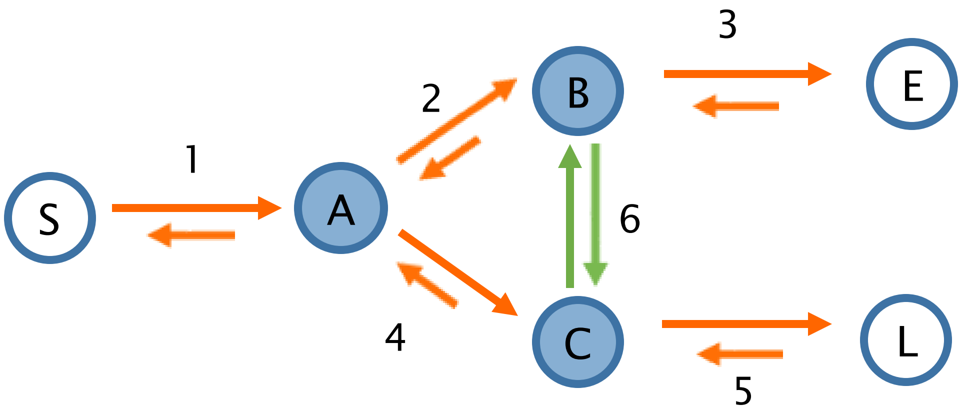
$$\Delta_r G'_6 = \Delta_r G'_6 + RT \ln C - RT \ln B$$

$\Delta_r G'_6 < 0$



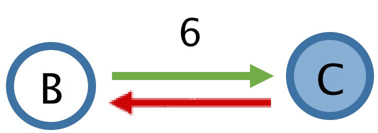
$\Delta_r G'_6 > 0$

Bi-directional reactions



$$\Delta_r G_6' = \Delta_r G_6'^{\circ} + RT \ln C - RT \ln B$$

$\Delta_r G_6' < 0$



$\Delta_r G_6' > 0$

$v_{1,FW}$, $v_{1,BW}$

$z_{1,FW}$, $z_{1,BW}$

How to formulate the problem?

Mixed integer linear programing

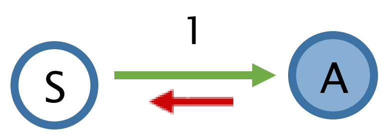
Mass balance

Variables: v_i reaction fluxes

z_i binary variables (either one or zero)

A reaction will yield therefore:

$\Delta_r G_i' < 0$



$$\begin{cases} v_{1,FW}, v_{1,BW} \\ z_{1,FW}, z_{1,BW} \end{cases}$$

$\Delta_r G_i' > 0$

Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **92** (2007), 1792-1805.

How to formulate the problem?

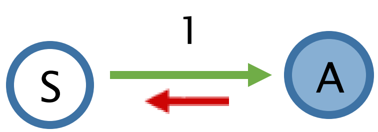
Mixed integer linear programing

Mass balance

Variables: v_i reaction fluxes
z_i binary variables (either one or zero)

A reaction will yield therefore:

$\Delta_r G_i' < 0$



$v_{1,FW}$, $z_{1,FW}$ | $v_{1,BW}$, $z_{1,BW}$

$\Delta_r G_i' > 0$

Constraints:

$$S\vec{v} = \vec{0}$$ Stoichiometry

$$z_{i,FW} + z_{i,BW} \leq 1$$ Either forward or backward or **none**

$$0 \leq v_{i,FW}, v_{i,BW} \leq v_{i,max}$$ Bound on the reaction rate

$$v_{i,FW} - z_{i,FW} v_{i,max} \leq 0$$ Binds the **flux** to the **binary variable**

Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **92** (2007), 1792-1805.

How to formulate the problem?

Mixed integer linear programing

Thermodynamics

Variables:

$\ln(x_i)$ log of the metabolite activity
$\Delta_r G_i'$ Free energy of the reaction

Constraints:

$$\ln(x_{i,min}) \leq \ln(x_i) \leq \ln(x_{i,max})$$

Bounds on the metabolite activities

$$\Delta_r G_i' - RT \sum_{j=1}^{m} n_{ij} \ln(x_j) = \Delta_r G_i'^{\circ}$$

Constraints on the $\Delta_r G_i'$ variables and the log metabolite activities

Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **92** (2007), 1792-1805.

How to formulate the problem?

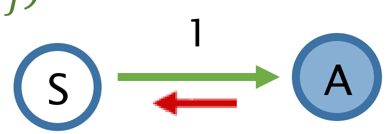
Mixed integer linear programing

Mass balance

Variables: v_i reaction fluxes

z_i binary variables (either one or zero)

$$\Delta_r G_i' = \Delta_r G_i'^{\circ} + RT \sum_{j=1}^{m} n_{ij} \ln(x_j) < 0$$



$$\begin{cases} v_{1,FW}, v_{1,BW} \\ z_{1,FW}, z_{1,BW} \end{cases}$$

$$\Delta_r G_i' = \Delta_r G_i'^{\circ} + RT \sum_{j=1}^{m} n_{ij} \ln(x_j) > 0$$

Constraints:

$S\vec{v} = \vec{0}$ Stoichiometry

$z_{i,FW} + z_{i,BW} \leq 1$ Either **forward** or **backward** or **none**

$0 \leq v_{i,FW}, v_{i,BW} \leq v_{i,max}$ Bound on the reaction rate

$v_{i,FW} - z_{i,FW} v_{i,max} \leq 0$ Binds the **flux** to the **binary variable**

Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **92** (2007), 1792-1805.

How to formulate the problem?

Mass balance | Thermodynamics

Variables:

Mass balance:

v_i reaction fluxes

z_i binary variables (either one or zero)

Thermodynamics:

$\ln(x_i)$ log of the metabolite activity

$\Delta_r G_i{}'$ Free energy of the reaction

Constraints:

Mass balance:

$$S\vec{v} = \vec{0}$$

$$z_{i,FW} + z_{i,BW} \leq 1$$

$$0 \leq v_{i,FW}, v_{i,BW} \leq v_{i,max}$$

$$v_{i,FW} - z_{i,FW} v_{i,max} \leq 0$$

$$v_{i,FW} - z_{i,BW} v_{i,max} \leq 0$$

Thermodynamics:

$$\ln(x_{i,min}) \leq \ln(x_i) \leq \ln(x_{i,max})$$

$$\Delta_r G_i' - RT \sum_{j=1}^{m} n_{ij} \ln(x_j) = \Delta_r G_i'^{\circ}$$

Combined:

$$\boldsymbol{\Delta_r G_i' - M + M z_{i,FW} \leq 0}$$

$$\boldsymbol{\Delta_r G_i' - M + M z_{i,BW} \geq 0}$$

M is a large constant selected such that: M > max($\Delta_r G_i'$)

Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **92** (2007), 1792-1805.

How to formulate the problem?

Mass balance | Thermodynamics

Variables:

Mass balance:

v_i reaction fluxes
z_i binary variables (either one or zero)

Thermodynamics:

$\ln(x_i)$ log of the metabolite activity
$\Delta_r G_i{}'$ Free energy of the reaction

Constraints:

Mass balance:

$$S\vec{v} = \vec{0}$$

$$z_{i,FW} + z_{i,BW} \leq 1$$

$$0 \leq v_{i,FW}, v_{i,BW} \leq v_{i,max}$$

$$v_{i,FW} - z_{i,FW} v_{i,max} \leq 0$$

$$v_{i,FW} - z_{i,BW} v_{i,max} \leq 0$$

Thermodynamics:

$$\ln(x_{i,min}) \leq \ln(x_i) \leq \ln(x_{i,max})$$

$$\Delta_r G_i' - RT \sum_{j=1}^{m} n_{ij} \ln(x_j) = \Delta_r G_i'^{\circ}$$

Combined:

$$\Delta_r G_i' - M + M z_{i,FW} \leq 0$$

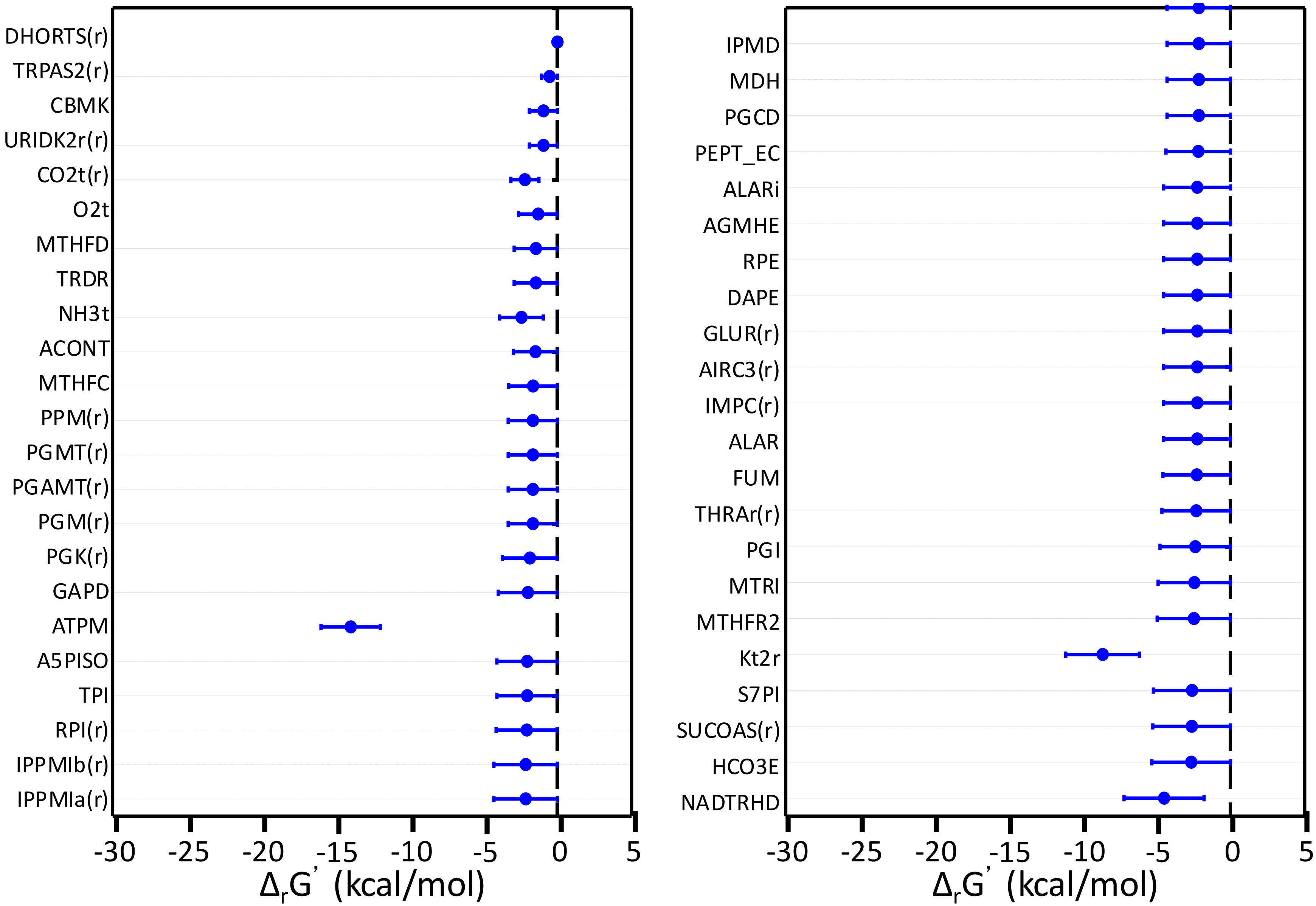
$$\Delta_r G_i' - M + M z_{i,BW} \geq 0$$

Objective function:

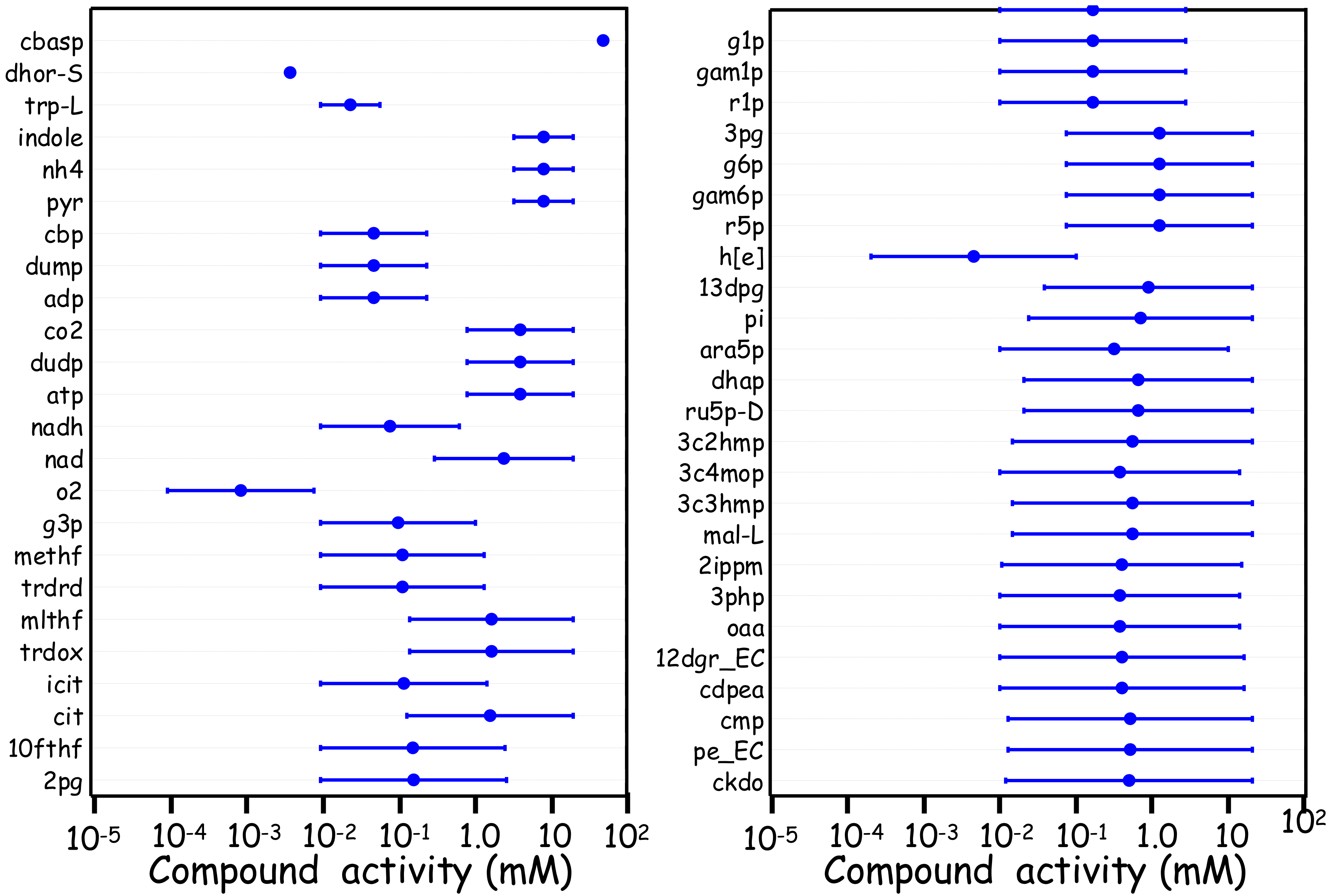
min/max(**objective**)

Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **92** (2007), 1792-1805.

Variability analysis $\Delta_r G_i'$



Variability analysis $\ln(x_i)$



Experimental error?

Mass balance

Thermodynamics

Variables:

v_i reaction fluxes
z_i binary variables (either one or zero)

$\ln(x_i)$ log of the metabolite activity
$\Delta_r G_i{}'$ Free energy of the reaction

Constraints:

$$S\vec{v} = \vec{0}$$

$$z_{i,FW} + z_{i,BW} \leq 1$$

$$0 \leq v_{i,FW}, v_{i,BW} \leq v_{i,max}$$

$$v_{i,FW} - z_{i,FW} v_{i,max} \leq 0$$

$$v_{i,FW} - z_{i,BW} v_{i,max} \leq 0$$

$$\ln(x_{i,min}) \leq \ln(x_i) \leq \ln(x_{i,max})$$

$$\Delta_r G_i' - RT \sum_{j=1}^{m} n_{ij} \ln(x_j) = \Delta_r G_i'^{\circ}$$

$$\Delta_r G_i' - M + M z_{i,FW} \leq 0$$

$$\Delta_r G_i' - M + M z_{i,BW} \geq 0$$

Objective function:

min/max(**objective**)

Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **92** (2007), 1792-1805.

Experimental error?

Mass balance

Variables:

v_i reaction fluxes
z_i binary variables (either one or zero)

$$S\vec{v} = \vec{0}$$

Constraints:

$$z_{i,FW} + z_{i,BW} \leq 1$$

$$0 \leq v_{i,FW}, v_{i,BW} \leq v_{i,max}$$

$$v_{i,FW} - z_{i,FW} v_{i,max} \leq 0$$

$$v_{i,FW} - z_{i,BW} v_{i,max} \leq 0$$

Thermodynamics

$\ln(x_i)$ log of the metabolite activity
$\Delta_r G_i{}'$ Free energy of the reaction
$\boldsymbol{\Delta_r G_i'^\circ}$ **Standard free energy of the reaction**

$$\ln(x_{i,min}) \leq \ln(x_i) \leq \ln(x_{i,max})$$

$$\Delta_r G'^\circ_{i,min} \leq \Delta_r G'^\circ_i \leq \Delta_r G'^\circ_{i,max}$$

$$\Delta_r G'_i - RT \sum_{j=1}^{m} n_{ij} \ln(x_j) = \Delta_r G'^\circ_i$$

$$\Delta_r G'_i - M + M z_{i,FW} \leq 0$$

$$\Delta_r G'_i - M + M z_{i,BW} \geq 0$$

Objective function:

min/max(**objective**)

$$\Delta_r G'^\circ_{i,min/max} = \Delta_r G'^\circ_{i,estimate} \pm \Delta_r G'^\circ_{i,error}$$

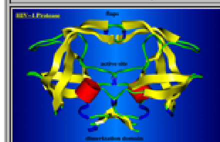
Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **92** (2007), 1792-1805.

Thermodynamic information

can we look up the free energy?

Thermodynamic information

Experimental information about
the free energy of reactions

Enzyme	Thermodynamics of Enzyme-Catalyzed Reactions	NIST
	Thermodynamics of Enzyme-Catalyzed Reactions	NIST

NIST Standard Reference Database 74

Biochemical Science Division, National Institute of Standards and Technology Gaithersburg, MD 20899 U.S.A.

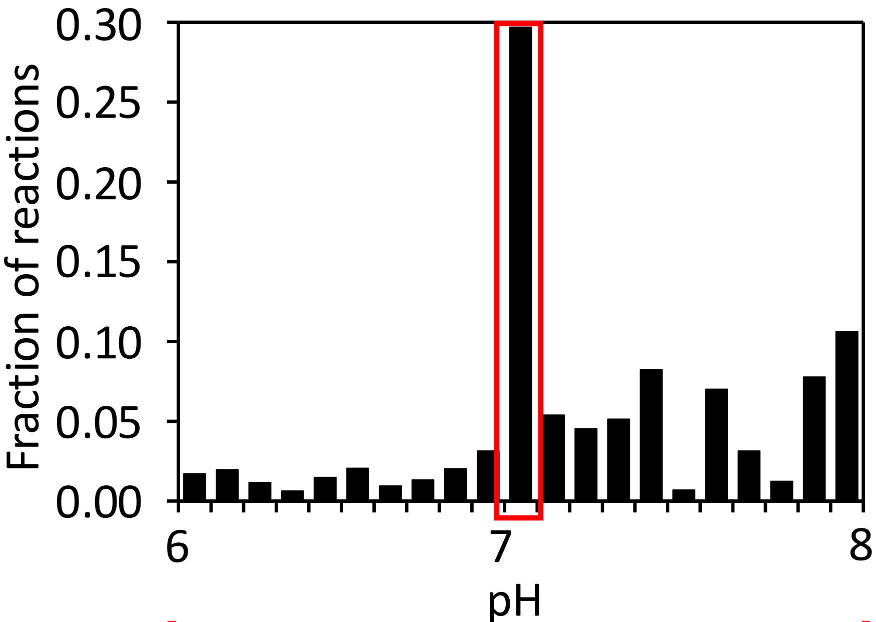
In citing this work please use:
Goldberg RN, Tewari YB, Bhat TN, "Thermodynamics of Enzyme-Catalyzed Reactions -a Database for Quantitative Biochemistry", Bioinformatics 2004;20(16):2874-2877.

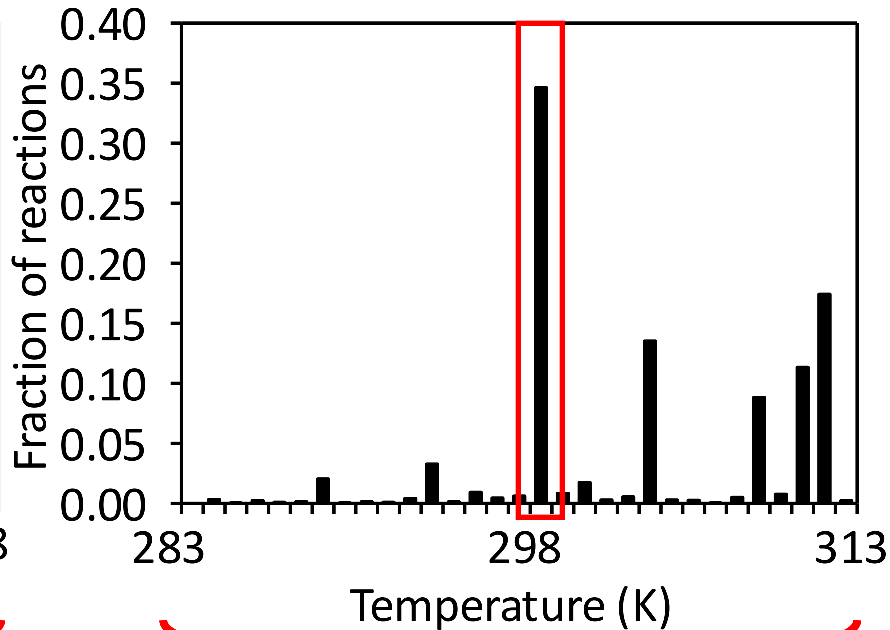
Information is available on the following topics:

- Thermodynamic Background
- How to Use This Database (Help)
- Enzyme Catalyzed Reactions: Search using predefined values |?
- Enzyme Catalyzed Reactions: Search with user defined values |?
- Enzyme Catalyzed Reactions: Show distinct values |?
- Enzyme Catalyzed Reactions: Show all values |?
- References: Search references |?
- Abbreviations
- Downloads
- Prior Version: Search prior version

Experimental $\Delta_r G'$

- Standard state for new method: pH 7, 298K, zero ionic strength





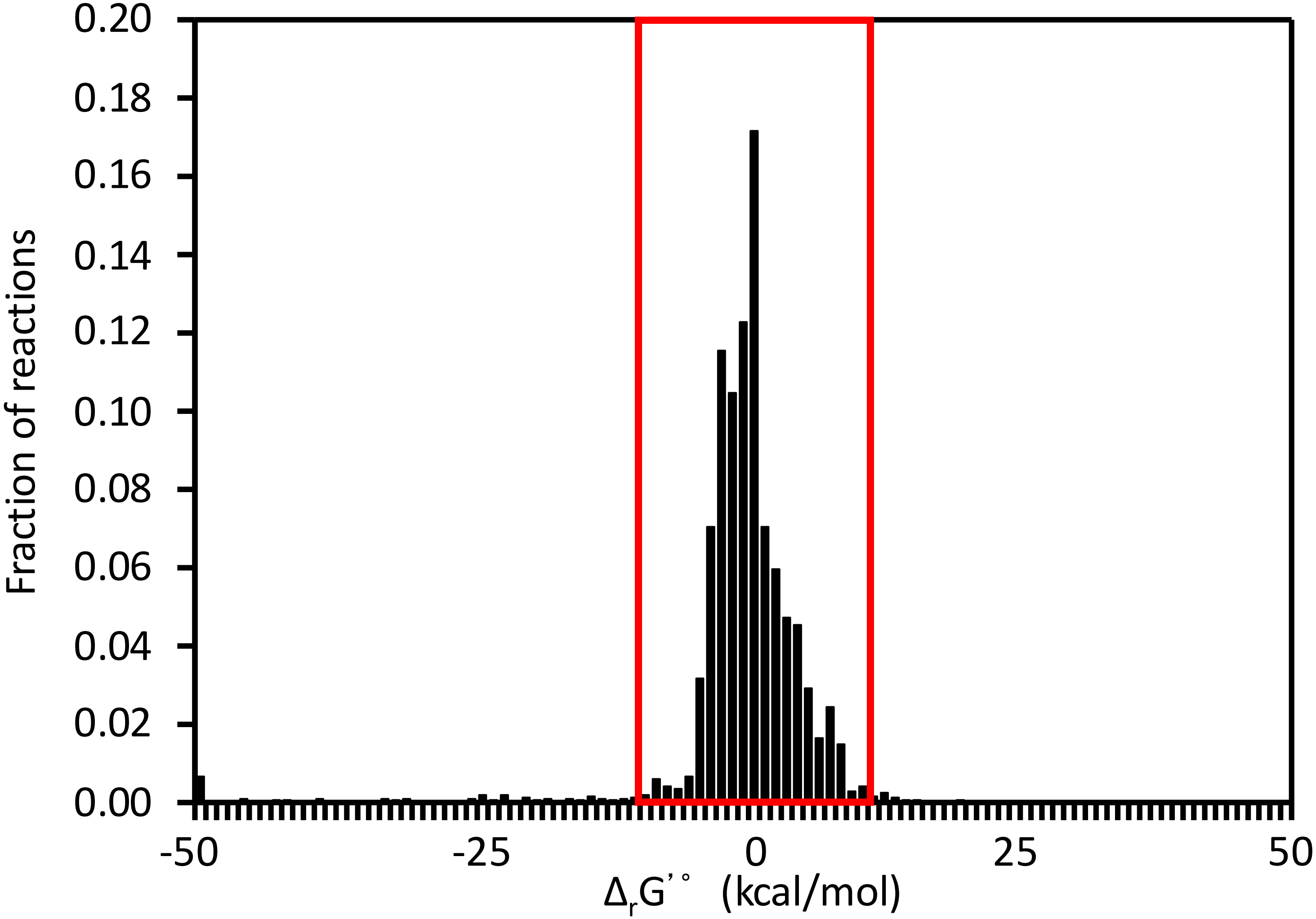
- Only $\Delta_r G'^{\circ}$ observations measured within 1 pH unit of pH 7 were accepted

- Only $\Delta_r G'^{\circ}$ observations measured within 15 K of 298 K were accepted

- The most prevalent condition for the $\Delta_r G'^{\circ}$ observations in the training set is **pH 7** and **298 K**

Jankowski, M. D., Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **95** (2008), 1487-1499.

Experimental $\Delta_r G'$



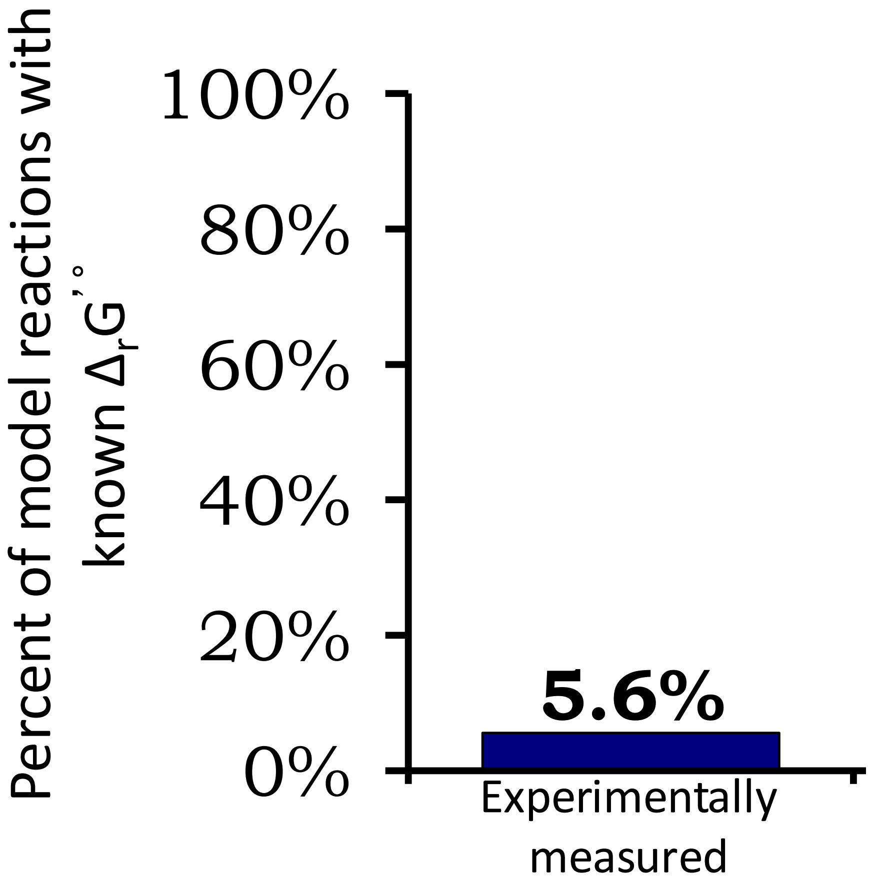
- Most of the $\Delta_r G'^{\circ}$ values less than 10 kcal/mol in magnitude

Jankowski, M. D., Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **95** (2008), 1487-1499.

Thermodynamic information

$\Delta_r G'^\circ$ data for biochemical reactions

Identifying every entry in the NIST database that **matches an E. coli reaction** and the selected **conditions: 298 ±10 K** and **pH 7 ±1** results in:



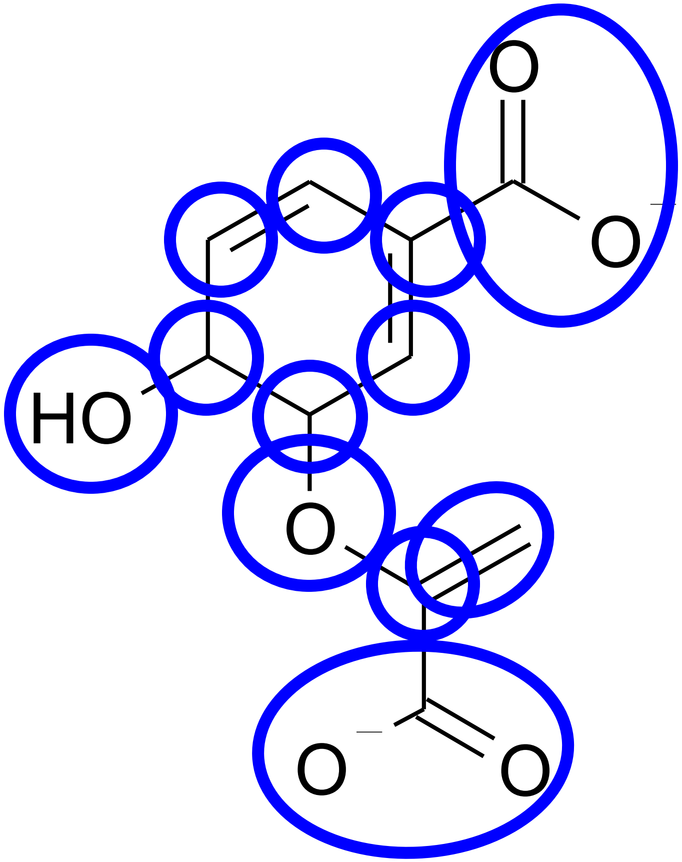
Goldberg, R. N., Y. B. Tewari and T. N. Bhat. *Bioinformatics.* **20** (2004), 2874-7.

Group contribution method

estimating Gibbs

Group contribution method

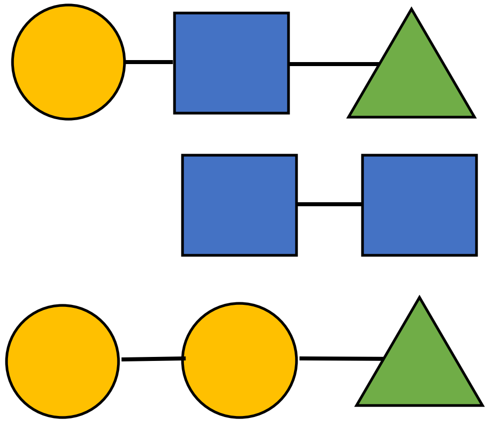
In the original group contribution method, the estimated $\Delta_f G'^\circ$ of a molecular is equal to the sum of the estimated energies for the molecular substructures that make up the molecule



Group	$\Delta_f G'^\circ_{group}$	Count
[-COO$^-$]	-72.0	2
[-OH]$_{sec}$	-32.0	1
[-CH=]$_{ring}$	9.6	3
[-O-]	-22.5	1
[>C=]	5.0	1
[>CH-]$_{ring}$	-2.2	2
[-CH$_2$-]	1.7	1
[>C=]$_{ring}$	8.2	1
[=CH$_2$]	18.4	1
Origin	-23.6	1
Chorismate:	-181.4 ± 5 kcal/mol	

Mavrovouniotis, M.L. *Biotechnology and Bioengineering*. **36** (1990), 1070-1082.

Linear regression

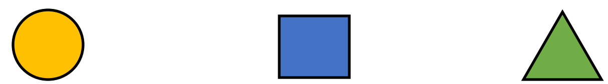


Compound 1 — $\Delta_f G'^\circ = 10 \text{ kcal /mol}$

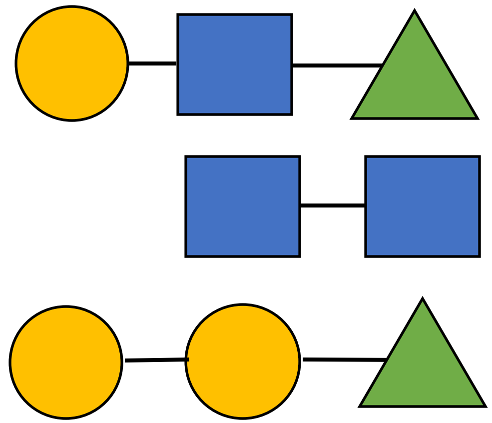
Compound 2 — $\Delta_f G'^\circ = 50 \text{ kcal /mol}$

Compound 3 — $\Delta_f G'^\circ = 30 \text{ kcal /mol}$

Group 1 Group 2 Group 3



Linear regression

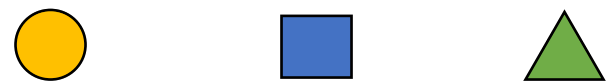


Compound 1 — $\Delta_f G'^{\circ} = 10 \text{ kcal /mol}$

Compound 2 — $\Delta_f G'^{\circ} = 50 \text{ kcal /mol}$

Compound 3 — $\Delta_f G'^{\circ} = 30 \text{ kcal /mol}$

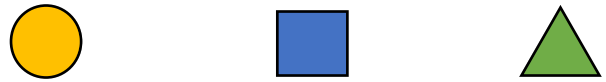
Group 1 Group 2 Group 3



$$\begin{pmatrix} \Delta_f G'^{\circ}_{C1,est} \\ \Delta_f G'^{\circ}_{C2,est} \\ \Delta_f G'^{\circ}_{C2,est} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta_{gr1} G \\ \Delta_{gr2} G \\ \Delta_{gr3} G \end{pmatrix}$$

Linear regression

Group 1 **Group 2** **Group 3**



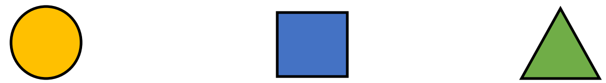
$$\begin{pmatrix} \Delta_f G'^{\circ}_{C1,est} \\ \Delta_f G'^{\circ}_{C2,est} \\ \Delta_f G'^{\circ}_{C2,est} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta_{gr1} G \\ \Delta_{gr2} G \\ \Delta_{gr3} G \end{pmatrix}$$

Minimize the difference between estimated and measured values:

$$\sum_i \left(\Delta_f G'^{\circ}_{i,exp} - \Delta_f G'^{\circ}_{i,est} \right)$$

Linear regression

Group 1 **Group 2** **Group 3**



$$\begin{pmatrix} \Delta_f G'^{\circ}_{C1,est} \\ \Delta_f G'^{\circ}_{C2,est} \\ \Delta_f G'^{\circ}_{C2,est} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta_{gr1} G \\ \Delta_{gr2} G \\ \Delta_{gr3} G \end{pmatrix}$$

$$\boldsymbol{\Delta_f G'^{\circ}_{est}} = X \cdot \boldsymbol{\Delta_{gr} G'^{\circ}}$$

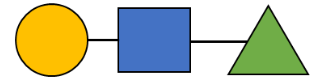
Minimize the difference between estimated and measured values:

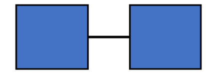
$$\sum_i \left(\Delta_f G'^{\circ}_{i,exp} - \Delta_f G'^{\circ}_{i,est} \right)$$

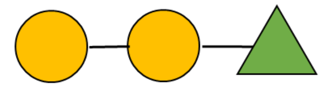
Can be solved analytically as:

$$\boldsymbol{\Delta_{gr} G'^{\circ}} = (X^T X)^{-1} X \boldsymbol{\Delta_f G'^{\circ}_{exp}}$$

Linear regression

Compound 1  $\Delta_f G'^{\circ} = 10 \text{ kcal/mol}$

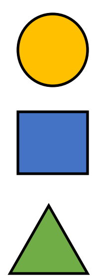
Compound 2  $\Delta_f G'^{\circ} = 50 \text{ kcal/mol}$

Compound 3  $\Delta_f G'^{\circ} = 30 \text{ kcal/mol}$

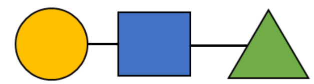
$$\boldsymbol{\Delta_{gr} G'^{\circ}} = (X^T X)^{-1} X \boldsymbol{\Delta_f G'^{\circ}_{exp}}$$

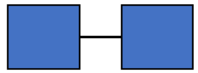
$$\boldsymbol{\Delta_{gr} G'^{\circ}} = \left(\begin{pmatrix} 1 & 0 & 2 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 10 \\ 50 \\ 30 \end{pmatrix}$$

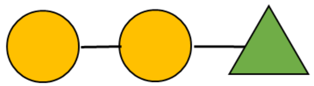
$$\boldsymbol{\Delta_{gr} G'^{\circ}} = \begin{pmatrix} 52.5 \\ 22.5 \\ -65.0 \end{pmatrix}$$



Linear regression

Compound 1  $\Delta_f G'^\circ = 10 \text{ kcal /mol}$

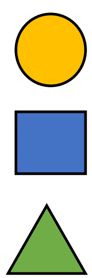
Compound 2  $\Delta_f G'^\circ = 50 \text{ kcal /mol}$

Compound 3  $\Delta_f G'^\circ = 30 \text{ kcal /mol}$

$$\boldsymbol{\Delta_{gr} G'^\circ} = (X^T X)^{-1} X \boldsymbol{\Delta_f G'^\circ_{exp}}$$

$$\boldsymbol{\Delta_{gr} G'^\circ} = \left(\begin{pmatrix} 1 & 0 & 2 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 10 \\ 50 \\ 30 \end{pmatrix}$$

$$\boldsymbol{\Delta_{gr} G'^\circ} = \begin{pmatrix} 52.5 \\ 22.5 \\ -65.0 \end{pmatrix}$$



$$=> \boldsymbol{\Delta_{est} G'^\circ} = \begin{pmatrix} 10 \\ 45 \\ 40 \end{pmatrix}$$

Linear regression

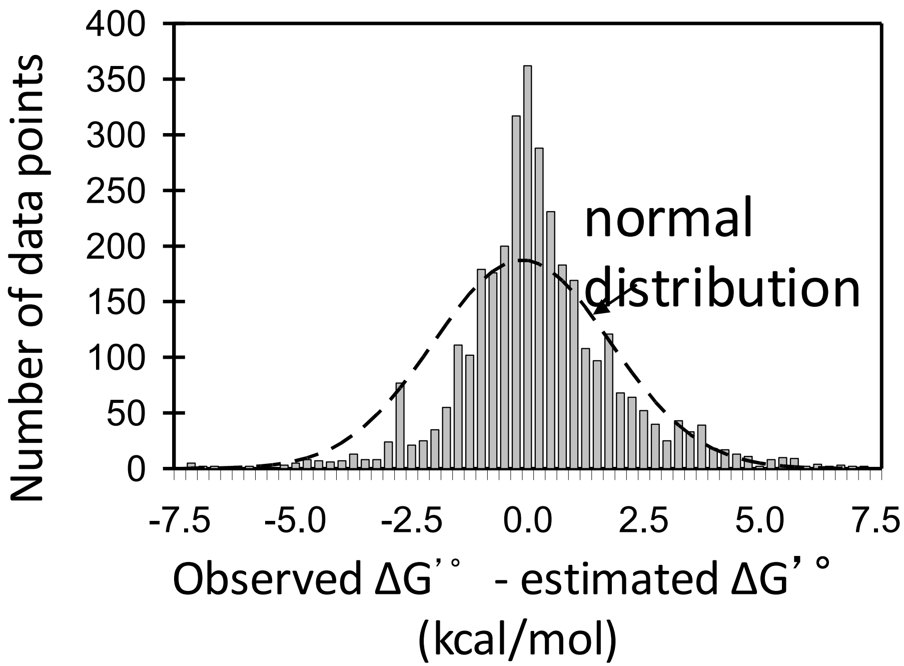
$$\Delta_{gr}G'^{\circ} = \left(X'X\right)^{-1}\left(X\,\Delta G'^{\circ}_{obs}\right)$$

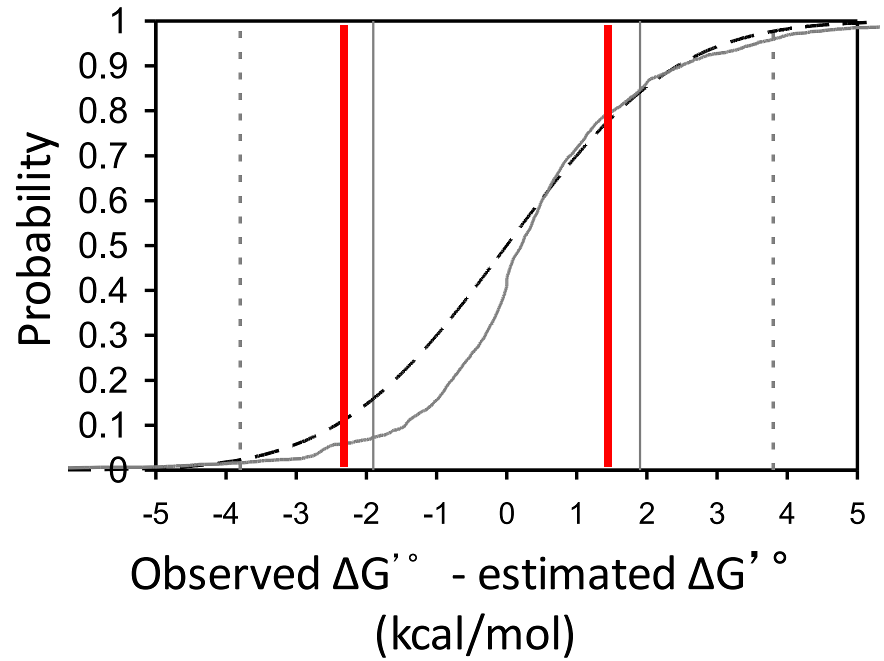
matrix containing the number of each group involved in each compound and reaction in the training set

vector of observed $\Delta_f G'^{\circ}$ and $\Delta_r G'^{\circ}$ values in the training set

vector of group contribution energies for the method

Standard error for the fitting: 1.90 kcal/mol





Uncertainty from the linear regression

- Covariance matrix from the fitting was used to estimate the uncertainty in $\Delta_{gr}G'^{\circ}$ values

$$SE_{gr,i} = \sqrt{\left(SE_{MLR}^{2}\left(X'X\right)^{-1}\right)_{i,i}}$$

- matrix indicating the number of each group involved in each compound and reaction in the training set

- scalar equal to the standard deviation of the residuals from the least squares fitting

- uncertainty in group contribution energy value *i*

- $\Delta_{gr}G'^{\circ}$ values used to determine the uncertainty in the $\Delta_{f}G'^{\circ}$ and $\Delta_{r}G'^{\circ}$ estimates:

$$SE_{\Delta G'^{\circ}_{est},j} = \sqrt{\sum_{i=1}^{N_{gr}}\left(X_{i,j}SE_{gr,i}\right)^{2}}$$

- Uncertainty in the $\Delta_{f}G'^{\circ}$ and $\Delta_{r}G'^{\circ}$ estimates

Jankowski, M. D., Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **95** (2008), 1487-1499.

Molecular substructures (Mavrovouniotis)

Group name	New $\Delta_{gr}G'^{\circ}$ kcal/mol	SE_{gr} kcal/mol	*f*
-SH	-0.740	0.636	260
-S-	8.77	0.740	190
-S-S-	5.69	1.20	16
$-O-PO_3^{2-}$	-254	0.159	380
$-O-PO_2^{2-}$	-205	0.440	149
$-O-PO_2^{1-}-$	-208	0.122	490
$-O-PO_2^{1-}$(ring)	-190	0.957	11
$-O-PO_2^{1-}-O-$	-234	0.438	48
$-COOPO_3^{2-}-$	-298	0.239	97
$-O-PO_3^{2-}$	-254	0.159	380
$-NH_3^{+}$	-6.25	0.196	236
$-NH_2$	2.04	0.331	223
>N-	24.4	1.14	9
$>NH_2^{+}$	5.95	0.900	5
>N- (two fused rings)	12.4	1.10	18
>NH	10.5	0.515	250
>NH (ring)	6.18	0.532	108
$>NH^{+}-$	15.5	1.17	3
>N- (ring)	22.1	0.617	777
=NH	-21.7	1.52	6
$=NH_2^{+}$	-22.7	1.34	9
=N- (ring)	4.17	0.572	41
$=N^{+}<$ (ring)	13.5	0.672	721
≡N	-32.1	4.34	4
-OH	-41.5	0.126	1117
-O- (ring)	-36.6	0.902	195
-O-	-23.2	0.408	39

Jankowski, M. D., Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **95** (2008), 1487-1499.

Molecular substructures (Mavrovouniotis)

Group name	New $\Delta_{gr}G'^{\circ}$ kcal/mol	SE_{gr} kcal/mol	f
>C=O	-28.4	0.180	734
>C=O (ring)	-30.1	0.292	88
-CH=O	-30.4	0.164	204
$-COO^{1-}$	-83.1	0.111	455
-O-CO-	-75.3	0.422	26
-O-CO- (ring)	-71.0	0.787	18
-CH3	-3.65	0.109	332
>CH2	1.62	0.0880	916
>CH2 (ring)	3.18	0.247	781
>CH-	5.08	0.153	981
>CH- (ring)	4.84	0.216	409
>C<	7.12	0.298	148
>C< (ring)	7.17	0.420	153

Group name	New $\Delta_{gr}G'^{\circ}$ kcal/mol	SE_{gr} kcal/mol	f
=CH-	12.8	0.242	198
>C=	15.7	0.394	135
$=CH_2$	6.87	0.312	110
≡CH	60.7	4.74	1
≡C-	41.6	2.32	3
>CH- (two fused rings)	2.60	0.779	30
>C= (aromatic ring)	6.95	0.313	66
>C= (dbl-sgl ring)	32.1	2.14	3
>C= (non-aromatic ring)	11.7	0.362	58
=CH- (non-aromatic ring)	8.46	0.293	755

Jankowski, M. D., Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **95** (2008), 1487-1499.

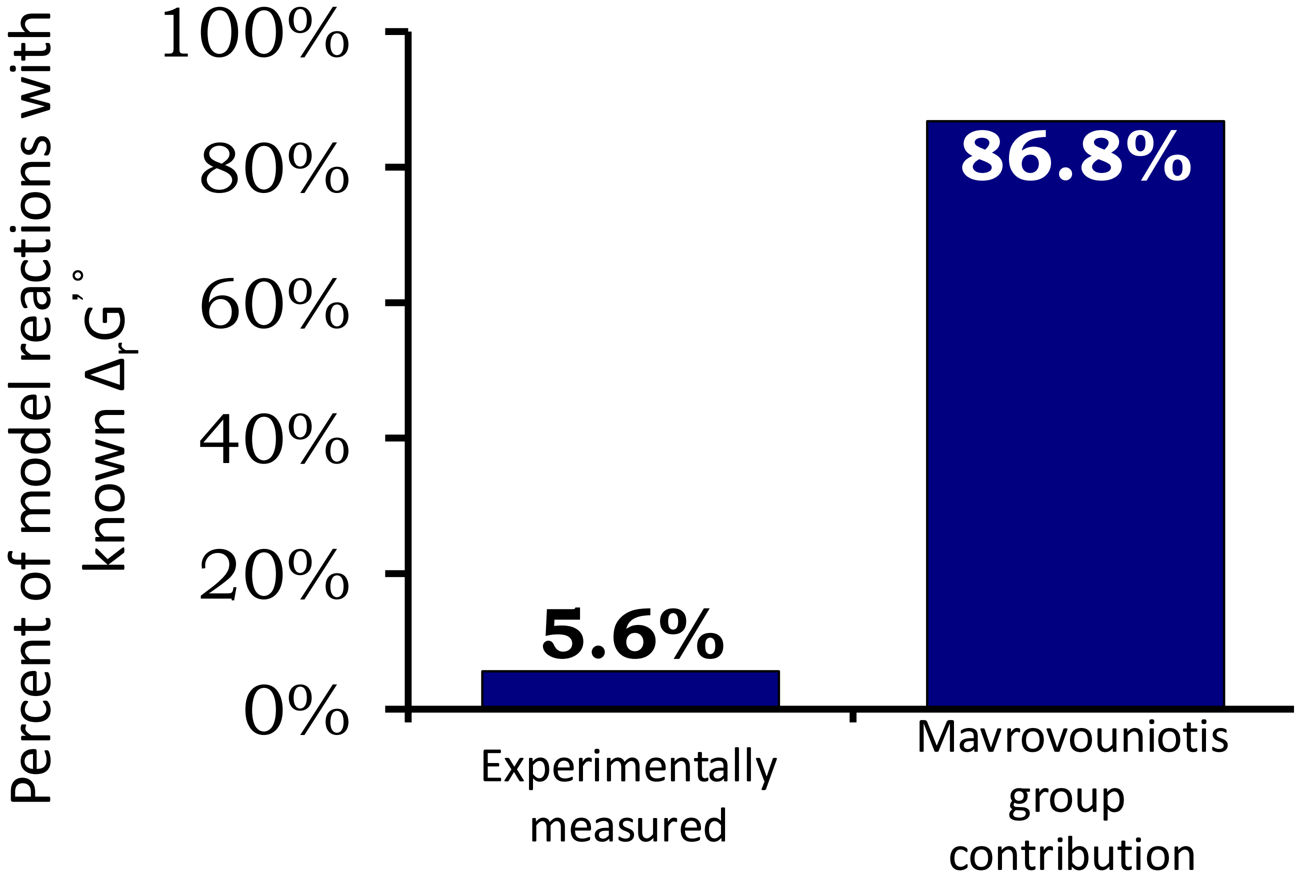
Molecular substructures (Mavrovouniotis)

Group name	New $\Delta_{gr}G'^{\circ}$ kcal/mol	SE_{gr} kcal/mol	f
>C= (two fused non-aromatic rings)	16.7	0.891	10
>C= (two fused rings: aromatic/ non-aromatic)	6.77	0.607	9
=CH- (aromatic ring)	4.93	0.142	64
>C= (two fused aromatic rings)	-0.0245	0.927	4

Jankowski, M. D., Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **95** (2008), 1487-1499.

Group contribution method

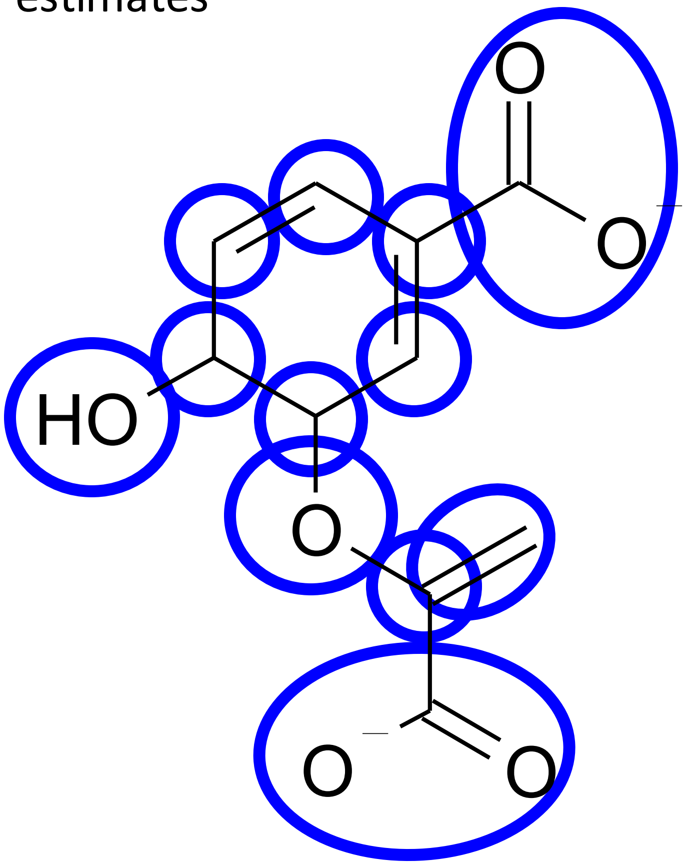
• Use of the group contribution method developed by Mavrovouniotis for biochemical reactions allows for the estimation of $\Delta_r G'$ for the majority of the E. coli reactions at 298K, pH 7, and zero ionic strength.



Henry, C. S., Jankowski, M. D., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal.*, **90** (2006), 1453-1461.

Group contribution method 1.1

- The uncertainty associated with each group energy value is known in the upgraded method allowing for **a better quantification of the uncertainty** in the estimates



Group	$\Delta_f G'^{\circ}_{group}$	Count	Uncertainty
[-COO$^-$]	-82.2	2	0.3
$[\text{-OH}]_{sec}$	-42.2	1	0.4
$[\text{-CH=}]_{ring}$	4.6	3	0.5
[-O-]	-24.9	1	0.9
[>C=]	14.5	1	0.7
$[\text{>CH-}]_{ring}$	5.5	2	0.6
[-CH_2-]	0.9	1	0.1
$[\text{>C=}]_{ring}$	9.7	1	1.0
[=CH_2]	4.8	1	1.1

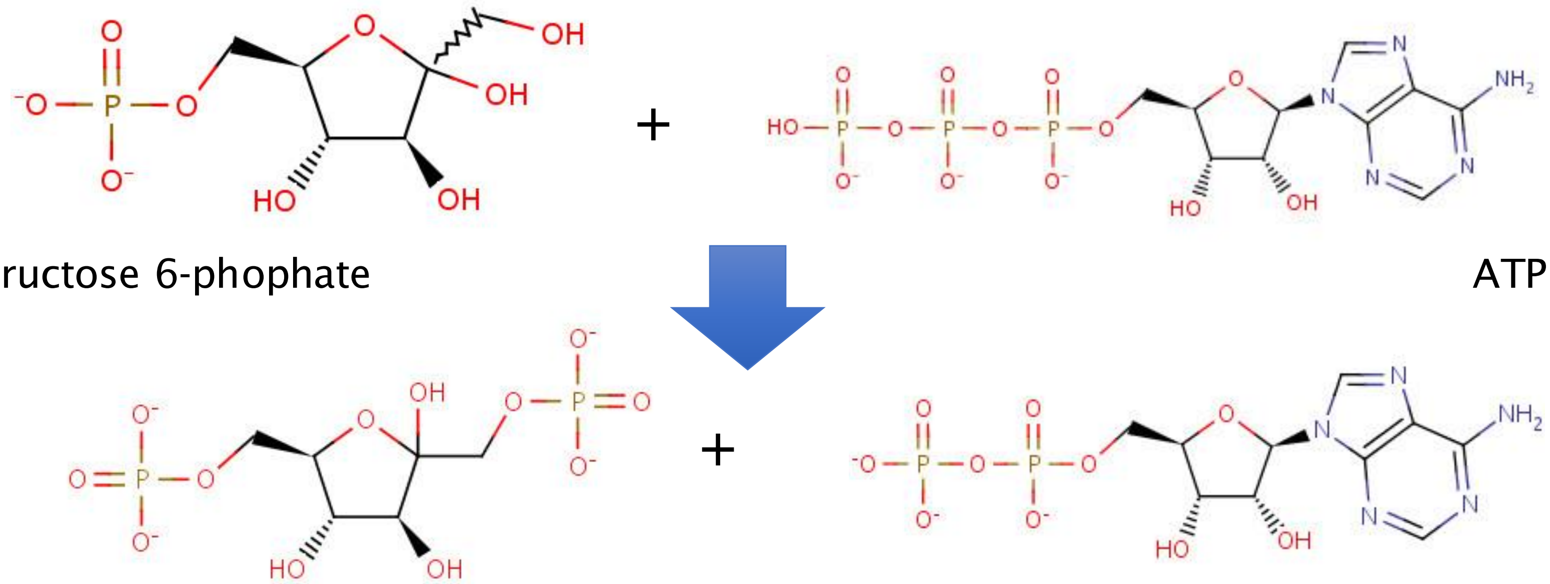
Chorismate: -176.8 ± 2.3 kcal/mol

$$\text{Uncertainty}_{\Delta_r G'^{\circ}} = \sqrt{\sum_{i}^{\text{Number of groups}} \left(n_i \varepsilon_{\Delta_g G_i'^{\circ}}\right)^2}$$

Jankowski, M. D., Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **95** (2008), 1487-1499.

Group contribution method

Phosphofructokinase



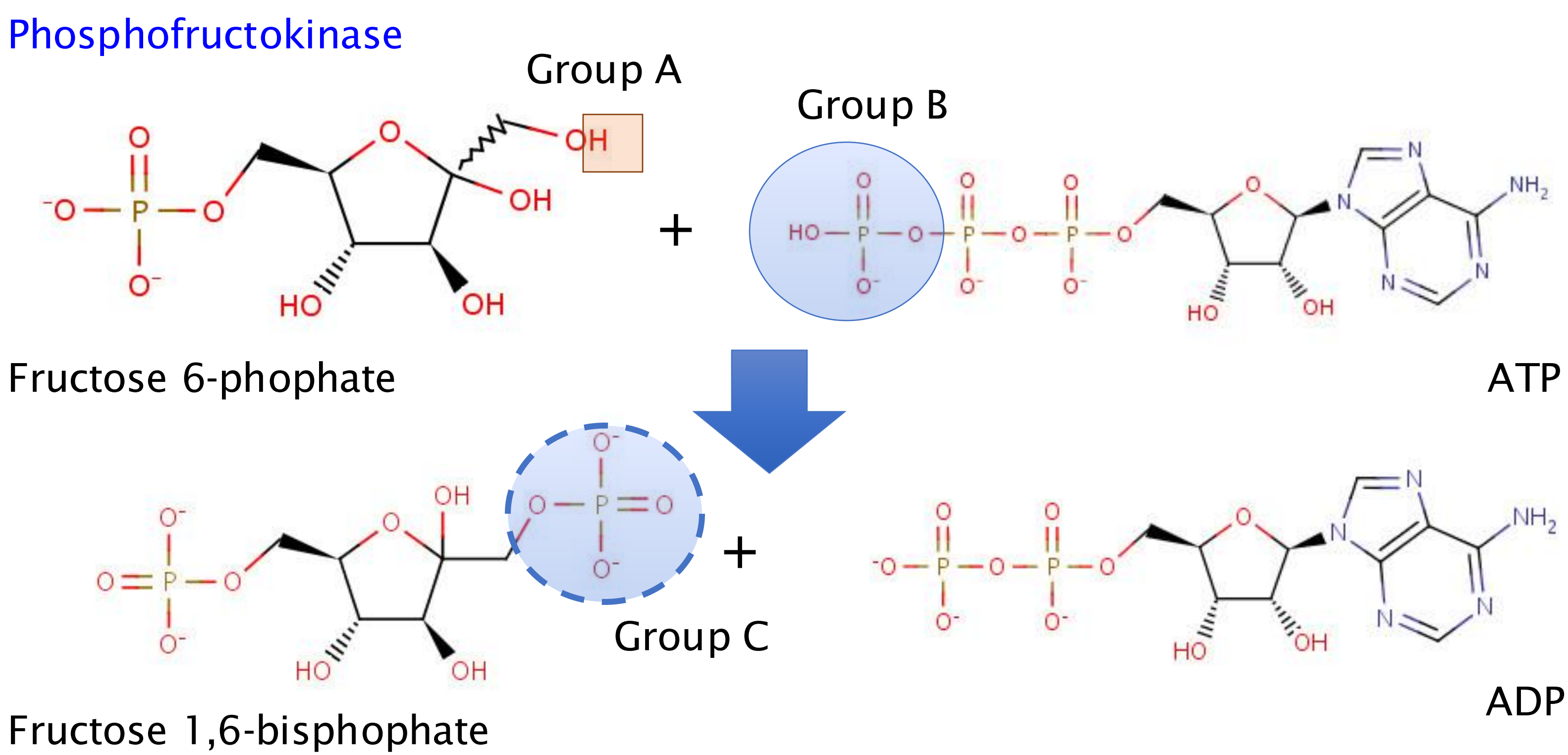
Fructose 6-phophate

ATP

Fructose 1,6-bisphophate

ADP

Group contribution method

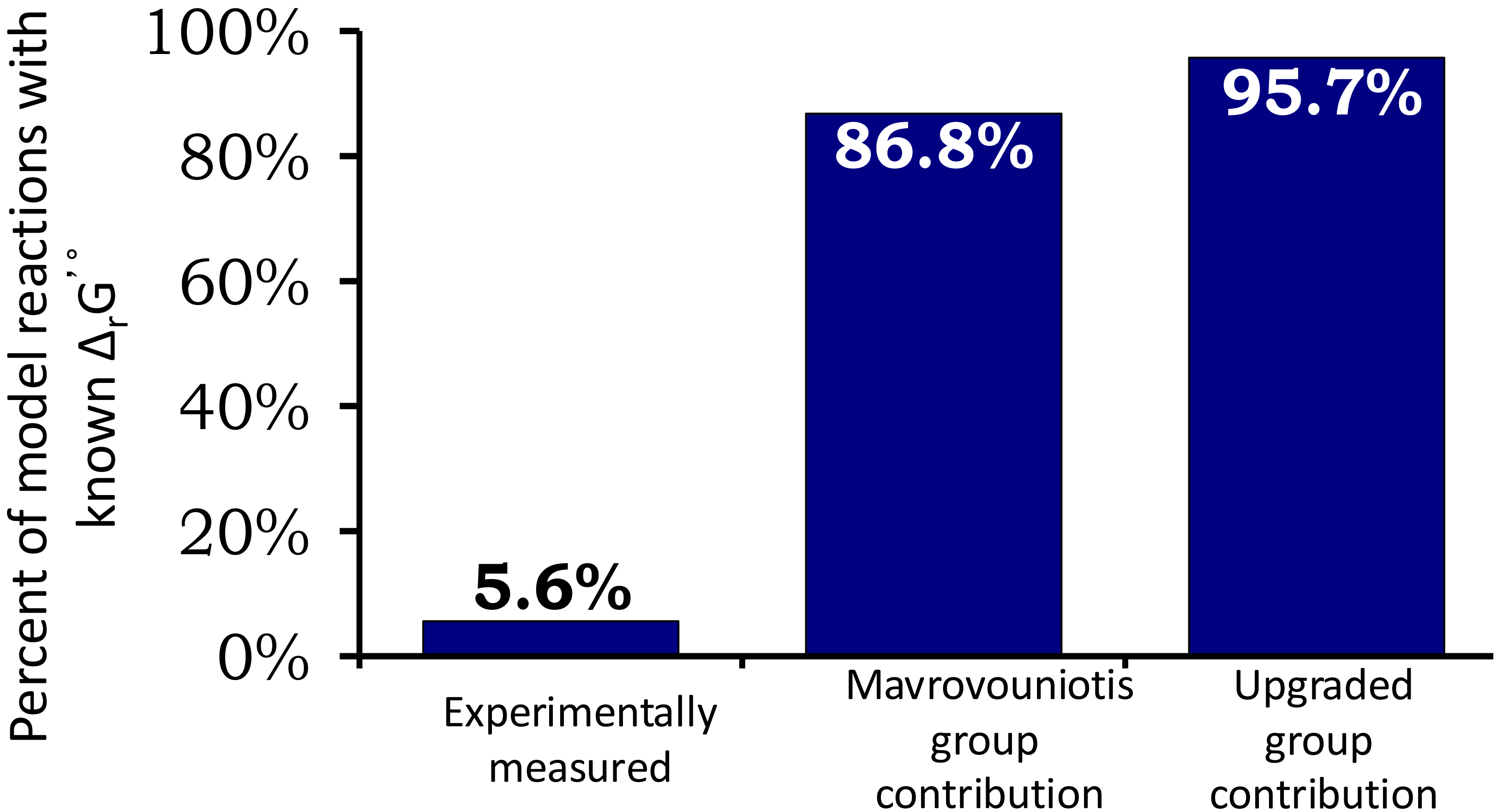


Gibbs Energy of Reaction:

$$\Delta G_r = - \Delta G_{g,grpA} - \Delta G_{g,grpB} + \Delta G_{g,grpC}$$

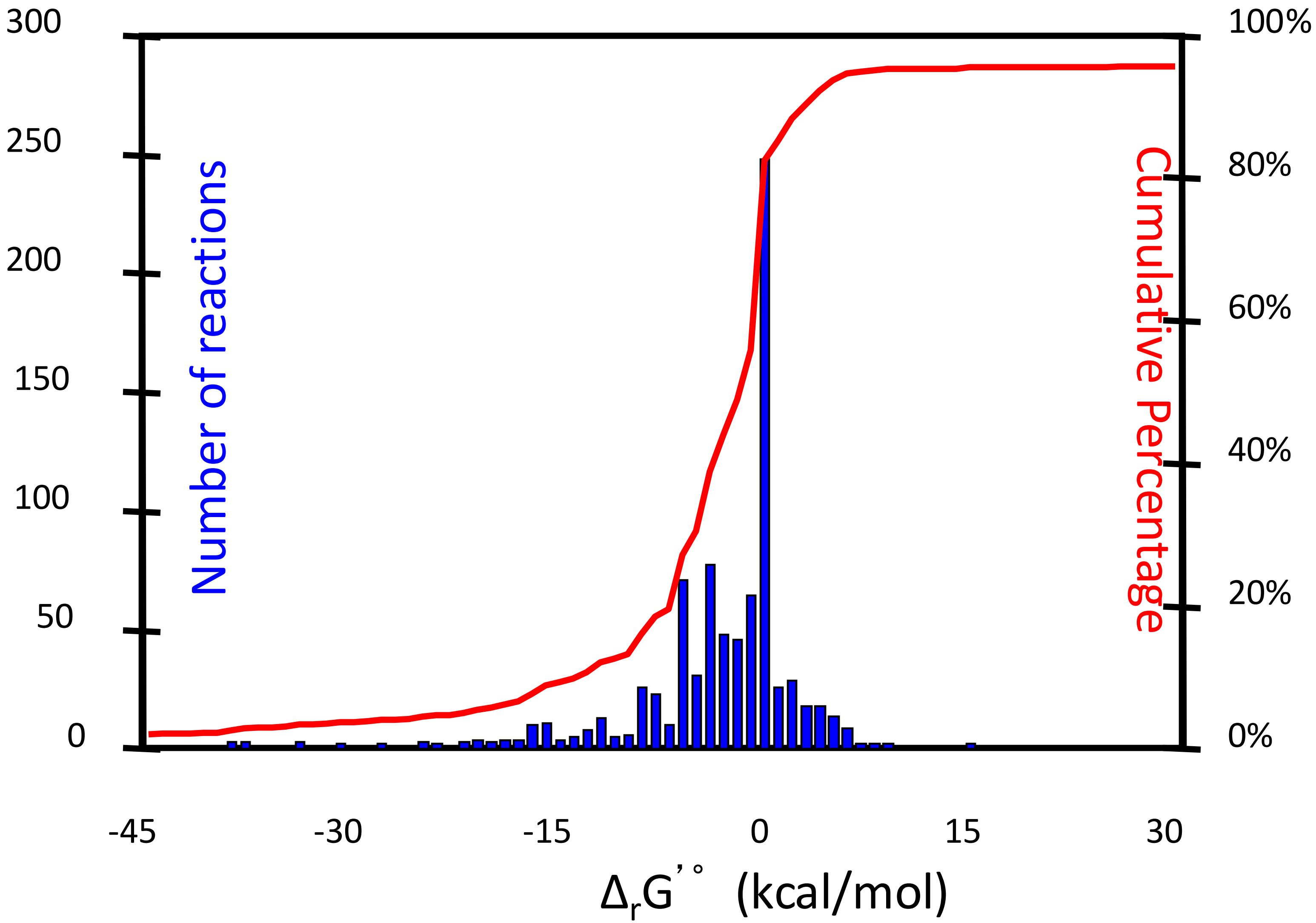
Group contribution method 1.1

- Expanded and Upgraded group contribution for the estimation of $\Delta_r G'^\circ$ for an even larger portion of the model reactions (891 of 931 reactions)

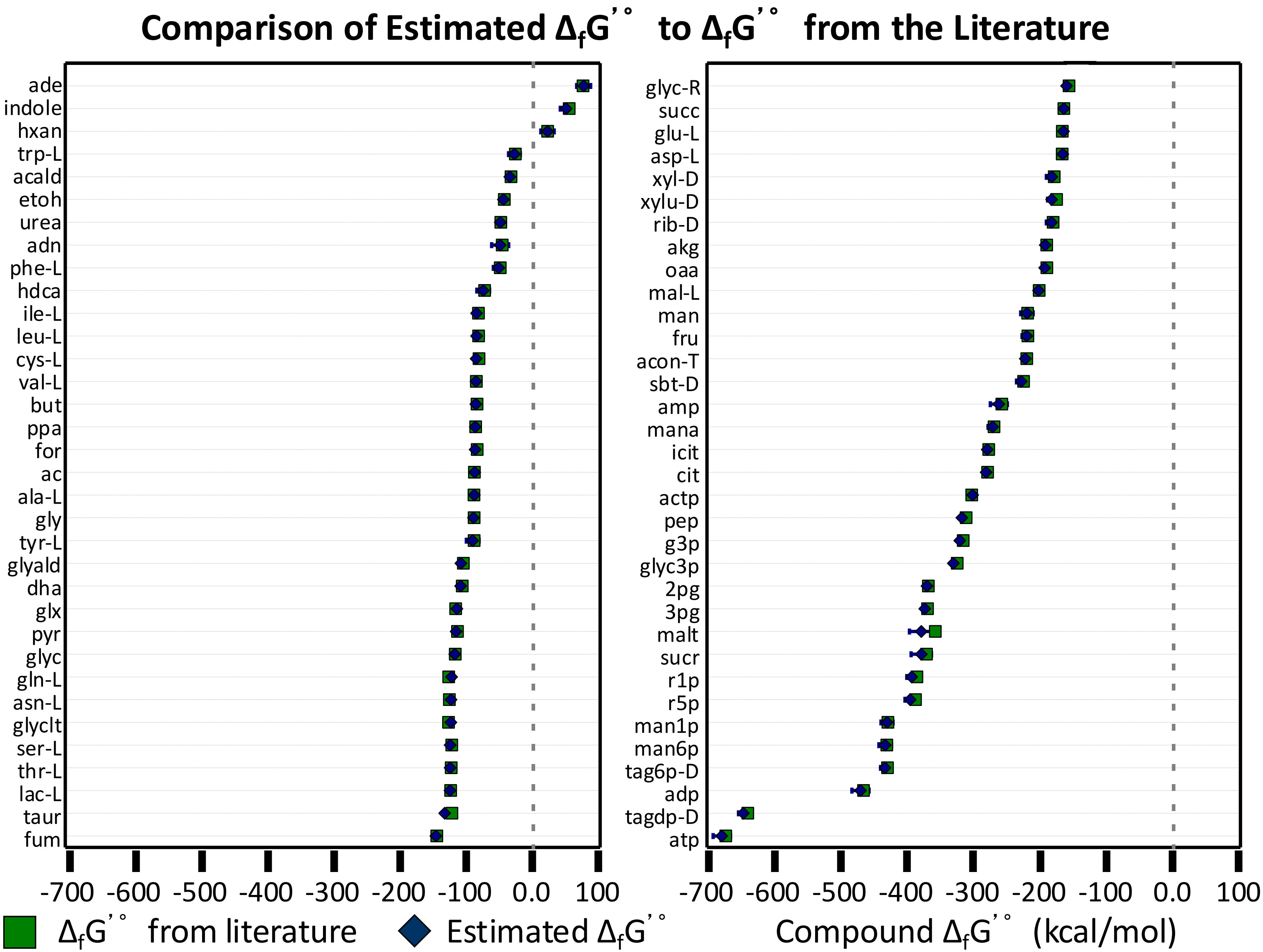


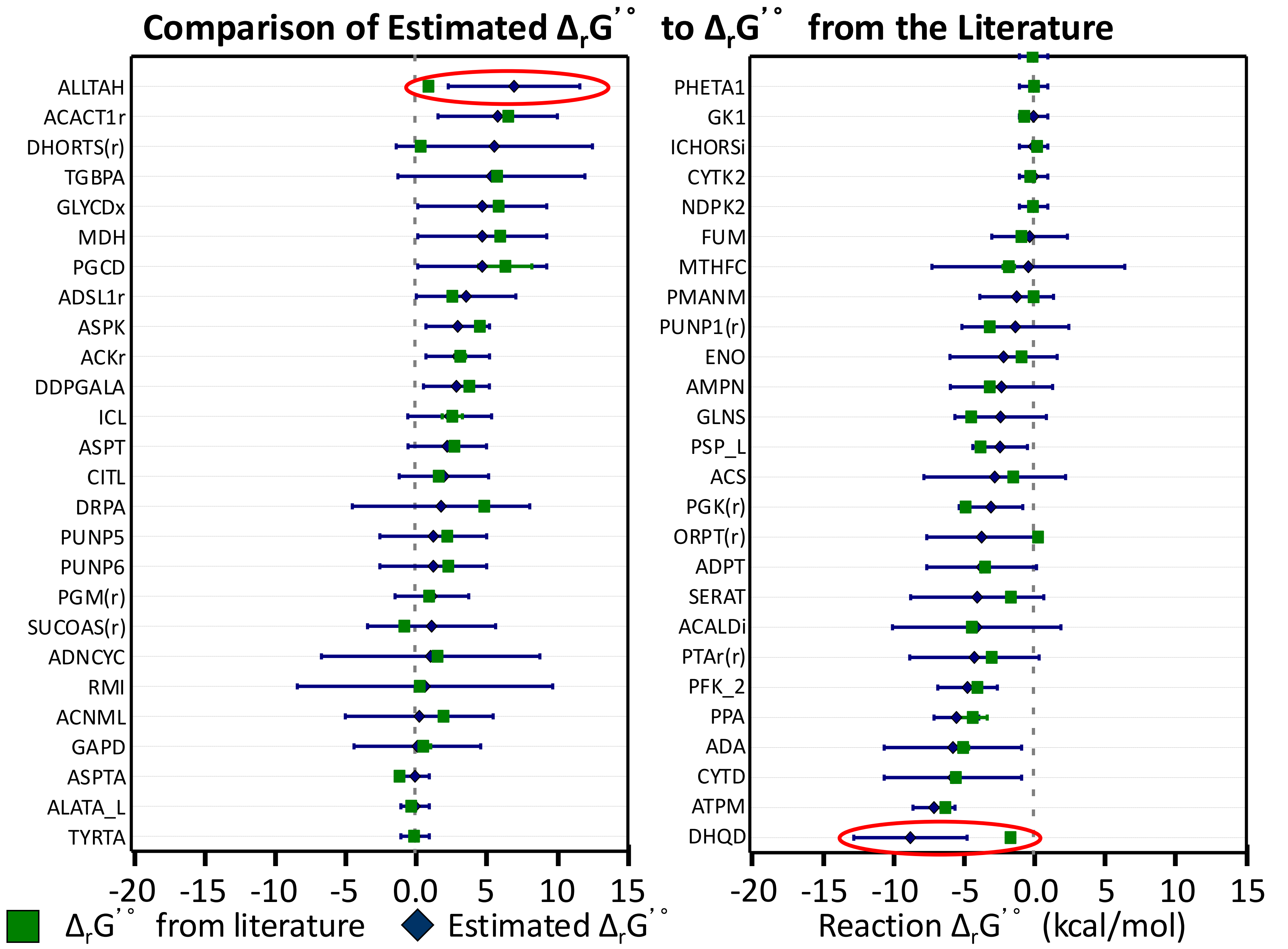
Jankowski, M. D., Henry, C. S., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. **95** (2008), 1487-1499.

Estimated ΔG distribution in E. coli



Henry, C. S., Jankowski, M. D., Broadbelt, L. J., and Vassily Hatzimanikatis. *Biophysical Journal*. (2006), 1453-1461.


Comparison of Estimated $\Delta_f G'^\circ$ to $\Delta_f G'^\circ$ from the Literature
ade
indole
hxan
trp-L
acald
etoh
urea
adn
phe-L
hdca
ile-L
leu-L
cys-L
val-L
but
ppa
for
ac
ala-L
gly
tyr-L
glyald
dha
glx
pyr
glyc
gln-L
asn-L
glyclt
ser-L
thr-L
lac-L
taur
fum
glyc-R
succ
glu-L
asp-L
xyl-D
xylu-D
rib-D
akg
oaa
mal-L
man
fru
acon-T
sbt-D
amp
mana
icit
cit
actp
pep
g3p
glyc3p
2pg
3pg
malt
sucr
r1p
r5p
man1p
man6p
tag6p-D
adp
tagdp-D
atp
-700
-600
-500
-400
-300
-200
-100
0.0
100
$\Delta_f G'^\circ$ from literature
Estimated $\Delta_f G'^\circ$
Compound $\Delta_f G'^\circ$ (kcal/mol)



Comparison of Estimated $\Delta_r G'^\circ$ to $\Delta_r G'^\circ$ from the Literature
ALLTAH
ACACT1r
DHORTS(r)
TGBPA
GLYCDx
MDH
PGCD
ADSL1r
ASPK
ACKr
DDPGALA
ICL
ASPT
CITL
DRPA
PUNP5
PUNP6
PGM(r)
SUCOAS(r)
ADNCYC
RMI
ACNML
GAPD
ASPTA
ALATA_L
TYRTA
PHETA1
GK1
ICHORSi
CYTK2
NDPK2
FUM
MTHFC
PMANM
PUNP1(r)
ENO
AMPN
GLNS
PSP_L
ACS
PGK(r)
ORPT(r)
ADPT
SERAT
ACALDi
PTAr(r)
PFK_2
PPA
ADA
CYTD
ATPM
DHQD
-20 -15 -10 -5 0.0 5 10 15
-20 -15 -10 -5 0.0 5 10 15
$\Delta_r G'^\circ$ from literature
Estimated $\Delta_r G'^\circ$
Reaction $\Delta_r G'^\circ$ (kcal/mol)

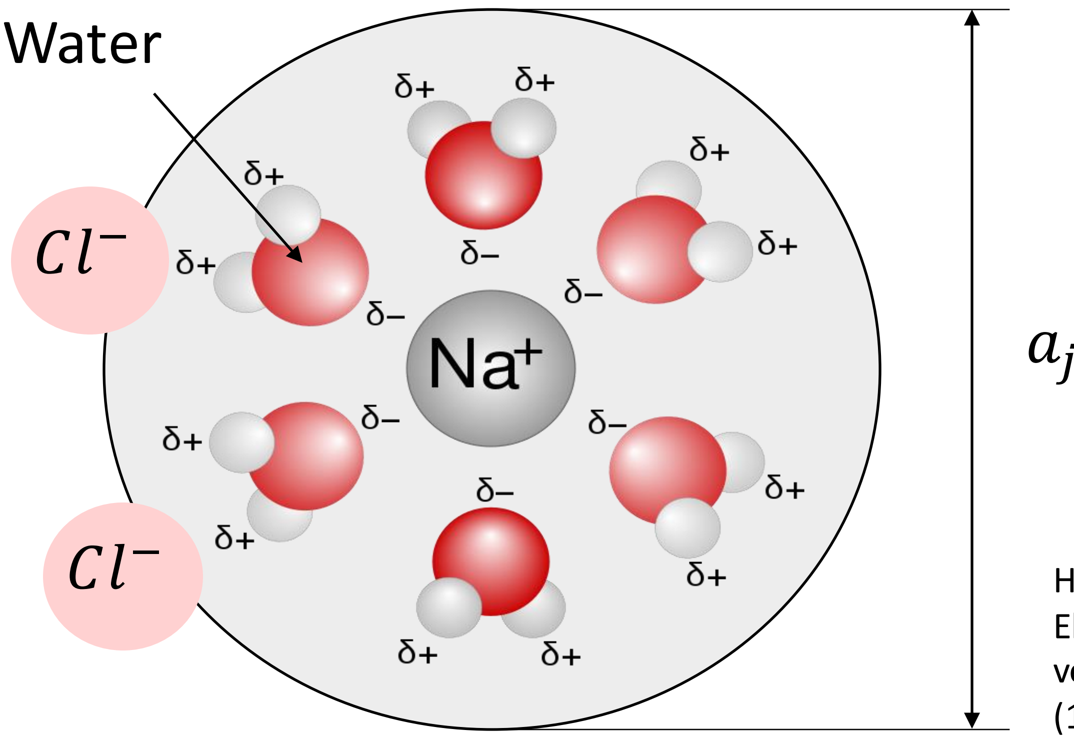
Correction to the free energy

from ideal to real

$\Delta_f G'^\circ$ in solution with ionic strength

Excess free energy due interaction with ions

$$\Delta_f G_j'^\circ(I) = \Delta_f G_j'^\circ(I = 0) - \Delta\Delta_f G_j'^\circ(I)$$



Huckel, E., and P. Debye. "Zur Theorie der Elektrolyte. I. Gefrierpunktserniedrigung und verwandte Erscheinungen." *Phys. Zeitschrift* 24 (1923): 185.

$\Delta_f G'^\circ$ with solution ionic strength

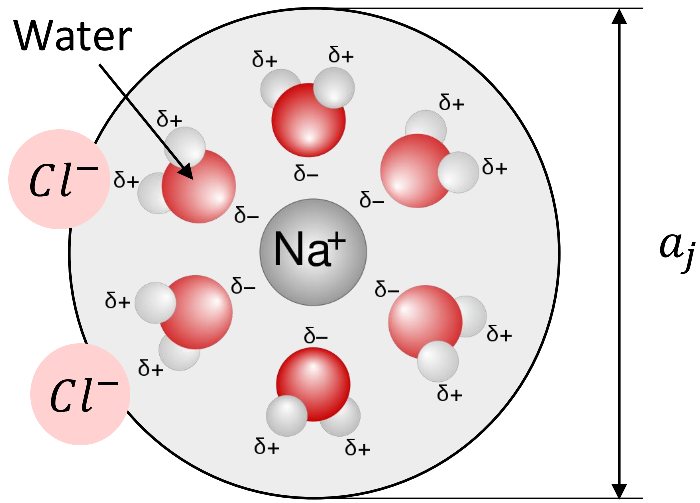
Debye-Hückel equation for the activity of ions:

$$\Delta_f G_j'^\circ(I) = \Delta_f G_j'^\circ(I = 0) - RT \frac{A\left(z_j^2 - N_{H,j}\right)\sqrt{I}}{1 + a_j B\sqrt{I}} \ln 10$$

R	ideal gas constant	$N_{H,j}$	number of hydrogen of compound j
T	temperature	I	ionic strength of the solution
z_j	charge of compound j	a_j	distance of the closest approach of ions to compound j

Where the Debye-Hückel parameters A, B can be calculated as a function of temperature:

Temp in °C	A in $\mathrm{mol}^{-1/2}\ \mathrm{L}^{1/2}$	B in $\mathrm{mol}^{-1/2}\ \mathrm{L}^{1/2}\ \mathrm{Å}^{-1}$
0	0,4883	0,3241
15	0,5002	0,3267
20	0,5046	0,3276
25	0,5092	0,3286
30	0,5141	0,3297
40	0,5241	0,3318
50	0,5351	0,3341
60	0,5471	0,3366
80	0,5739	0,3420



Huckel, E., and P. Debye. "Zur Theorie der Elektrolyte. I. Gefrierpunktserniedrigung und verwandte Erscheinungen." *Phys. Zeitschrift* 24 (1923): 185.

$\Delta_f G'^\circ$ with pH $\neq$ 7

$$\Delta_f G_j'^\circ(pH) = \Delta_f G_j'^\circ(pH = 7) + N_{H,j} RT \ln 10 \, \mathrm{pH}$$

R ideal gas constant
T temperature
N_j number of hydrogen of compound j

Other factors effecting $\Delta_r G'^{\circ}$

2394 Biophysical Journal Volume 109 December 2015 2394–2405

Article

The Influence of Crowding Conditions on the Thermodynamic Feasibility of Metabolic Pathways

Liliana Angeles-Martinez[1] and Constantinos Theodoropoulos[1,*]
[1]School of Chemical Engineering and Analytical Science, University of Manchester, Manchester, UK

ABSTRACT Intracellular reactions are carried out in a crowded medium where the macromolecules occupy ~40% of the total volume. This decrease in the available volume affects the activity of the reactants. Scaled particle theory is used for the estimation of the activity coefficients of the metabolites, and thereby for the assessment of the impact of the presence of background molecules, on the estimation of the Gibbs free energy change ($\Delta_r G$) of the reactions. The lactic acid pathway and the central carbon metabolism of *Actinobacillus succinogenes* for the production of succinic acid from glycerol have been used as illustrative case studies. Results suggest the importance of maintaining intracellular crowded regions to favor the feasibility of a pathway that in other circumstances would be infeasible. Moreover, the crowding conditions may change the directionality of reactions and can modify the feasible range of fluxes estimated for a metabolic system compared with those obtained at standard biological conditions.