

The laws of thermodynamics: applications in models of metabolism

“Advanced School on Quantitative Principles in Microbial Physiology: from Single Cells to Cell Communities”
ICTP - Trieste, Italy

Elad Noor

October 3-13, 2023

Contents

1	Introduction	2
1.1	The laws of thermodynamics	2
1.2	Biochemical reactions	2
1.3	The Gibbs free energy of reaction	2
1.4	Enzyme kinetics and the connection to thermodynamics	3
1.5	The flux-force relationship	4
1.6	Linear Programming (LP)	4
1.7	Quadratic Programming (QP)	5
1.8	Mixed-Integer Linear Programming (MILP)	5
2	Convex analysis of a metabolic flux cone	5
2.1	The flux cone	6
2.2	Elementary Flux Modes	6
2.3	Classification of EFMs	6
3	Thermodynamics in constraint-based models	8
3.1	Irreversibility is an approximation	8
3.2	Thermodynamic Feasibility of EFMs	8
3.3	Max-min Driving Force	8
3.4	Thermodynamic Flux Balance Analysis (TFBA)	9
3.5	Loopless Flux Balance Analysis (ll-FBA)	10
4	Gibbs energy dissipation	11
4.1	The Gibbs energy dissipation rate of a reaction	11
4.2	The Total Gibbs energy dissipation rate	11
4.2.1	OptKnock and its derivatives	12
4.3	A few relevant algorithms	13
4.3.1	Phenotypic Phase Plane (PPP)	13
4.3.2	Derivatives of OptKnock	13
4.3.3	Probabilistic Thermodynamic Analysis (PTA)	14

Abstract

We will start with the basic laws of thermodynamics, and see a few fundamental implications in enzyme catalysis and metabolism (including flux direction constraints, the Haldane relationship, and the flux-force relationship). We will also learn how to apply these laws to constraint-based metabolic models, and when selecting parameters for kinetic models. Finally, we will see how thermodynamics can also be used in optimization problems, such as the Max-min Driving Force (MDF) algorithm. Exercises will be carried out in Python. A good MILP solver (CPLEX or GUROBI) is recommended but not necessary.

1 Introduction

1.1 The laws of thermodynamics

If you can read French, you might be interested in Sadi Carnot's original book about thermodynamics: "Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance" [4].

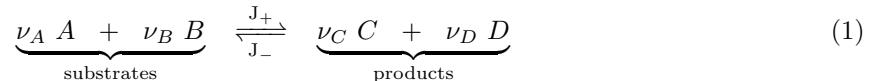
Over the years, thermodynamic theory developed further and is now typically stands upon 4 laws (or axioms):

- **0th law** – thermodynamic equilibrium is a transitive property.
- **1st law** – energy is conserved.
- **2nd law** – the entropy of an isolated system (such as the entire universe) can only increase over time.
- **3rd law** – a perfect crystal structure at absolute zero will have no entropy.

In metabolic cell models, we assume that no work is done by the system and that it is isobaric and isothermal. Another fundamental principle in enzymatic catalysis is the notion that enzymes can accelerate the rate of a reaction, but not change its thermodynamic equilibrium, therefore they have no effect on the *direction* of flux. Instead, reaction directions are entirely determined by the Gibbs free energy.

1.2 Biochemical reactions

Metabolism consists of individual biochemical reactions of the form:



where the uppercase letters represent the substrates and product as chemical entities, ν_* are the respective stoichiometric coefficients, and J^* are the forward (+) or backward (−) rates. Biochemical reactions occur in aqueous environments where each metabolite has a certain concentration. We will represent these concentrations by c_* , and define the reaction quotient as

$$Q' = \frac{c_C^{\nu_C} \cdot c_D^{\nu_D}}{c_A^{\nu_A} \cdot c_B^{\nu_B}}, \quad (2)$$

or more generally (with an arbitrary number of reacting substrates and products):

$$Q' = \prod_{i \in \text{products}} c_i^{\nu_i} \cdot \prod_{j \in \text{substrates}} c_j^{-\nu_j}. \quad (3)$$

Now, we can define an thermodynamic equilibrium as a state where the forward and backward fluxes are equal (i.e., $J^+ = J^-$). For all possible equilibrium states, the reaction quotient will always have the same value, which is denoted K^{eq} – the *equilibrium constant*. Note that there can be many different states of equilibrium, meaning that the individual metabolite concentrations don't have to be always the same. However, Q' will always be equal to K^{eq} at equilibrium.

1.3 The Gibbs free energy of reaction

The second law of thermodynamics states that entropy can only increase over time in an isolated system. In living cells, which are open systems in constant pressure, temperature, and pH, entropy can be replaced by the concept of transformed Gibbs free energy of reaction [1], denoted $\Delta_r G'$. Then, for a chemical reaction to be feasible, it must impose a negative change in $\Delta_r G'$:

$$\Delta_r G' < 0 \quad (4)$$

When the reaction reaches the minimum energy state, it will be at equilibrium and $\Delta_r G'$ will be equal to 0. On the other hand, if the reaction passes the equilibrium (e.g. someone adds a high concentration

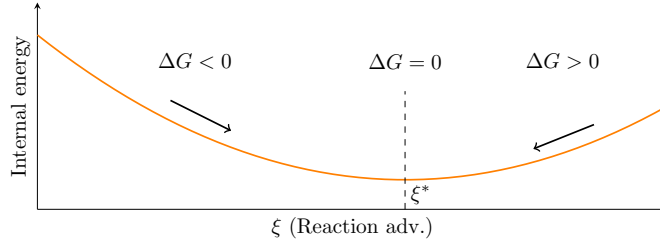


FIGURE 1: *Cartoon representation of Gibbs free energy of reaction and the thermodynamic equilibrium. As a chemical reaction proceeds, the concentrations of substrates and products change, which in turn affect the ‘energy in the chemical system’. We can, thus, capture the reaction advancement in a graph, where the x-axis represents the reaction advancement (i.e. the concentrations of substrates and products at different times in the reaction course) and the y-axis the internal energy of the system. The Gibbs free energy of reaction, in a way, indicates the position of the system in this graphical representation, where the thermodynamic equilibrium would be the energy minimum. At equilibrium, reaction Gibbs free energy would be zero, allowing us to derive the relation between substrate and product concentrations at that point and their free energy of formation. This relation is known as the equilibrium constant of the reaction.*

of the product to the solution), then $\Delta_r G'$ can be positive, and the reaction will spontaneously go in the reverse direction (see Figure 1).

We already said that the presence of enzymes does not affect the $\Delta_r G'$ directly, so what does actually determine its value? We typically use the assumption that biochemical reactions occur in dilute aqueous solutions. In this case, the transformed Gibbs free energy of reaction j is given by:

$$\Delta_r G' = \Delta_r G'^{\circ} + R \cdot T \cdot \ln(Q') \quad (5)$$

where R is the gas constant, T is the temperature (in Kelvin). $\Delta_r G'^{\circ}$ represents the $\Delta_r G'$ in standard conditions, which are typically defined as 1 M for each one of the reactants. One way to measure $\Delta_r G'^{\circ}$, is to let the reaction run until it reaches equilibrium, and measure the concentrations of all reactants. Also, since we know that at equilibrium $\Delta_r G' = 0$, we can solve for $\Delta_r G_j'^{\circ}$ and get:

$$\Delta_r G'^{\circ} = -R \cdot T \cdot \ln(K^{\text{eq}}) = -R \cdot T \cdot \left(\sum_{i \in \text{products}} \nu_i \ln(c_i^{\text{eq}}) - \sum_{j \in \text{substrates}} \nu_j \ln(c_j^{\text{eq}}) \right) \quad (6)$$

where c_i^{eq} are the measured molar concentrations of the reactants at an equilibrium state.

You can find measured values for K^{eq} (based on equation 6) on randr.nist.gov. Unfortunately, only a small fraction of known biochemical reactions have been measured. eQuilibrator is an alternative that covers almost all enzyme-catalyzed reactions. Its predictions are based on the component contribution method [18].

For one of the best introductions to thermodynamics of biochemical reactions, read the book by Alberty [1].

1.4 Enzyme kinetics and the connection to thermodynamics

It is a well cited fact, that enzymes can catalyze the rate of a given reaction, but cannot alter the direction of net flux. In the following section (1.5), we will see that an enzyme cannot even change the ratio between the forward and backward rates, only increase both of them proportionally.

But, is the converse true as well, or can information about the thermodynamics of a reaction tell us something about an enzyme catalyzing it? The answer is: yes, it can.

Mellors [12] explained how the relationship first described by JS Haldane in 1930 connects between an enzyme’s kinetic parameters and the reaction’s equilibrium constant. Imagine an enzyme that catalyzes a reaction $S \rightleftharpoons P$, whose flux is given by the following rate law:

$$v = E \frac{k_{\text{cat}}^+ \cdot s/K_s - k_{\text{cat}}^- \cdot p/K_p}{1 + s/K_s + p/K_p} \quad (7)$$

where E is the enzyme concentration, k_{cat}^+ and k_{cat}^- are the forward and backward turnover numbers, K_s and K_p are the affinity constants, and s and p are the concentrations of the substrate and product. This

rate law is sometimes mistakenly called reversible Michaelis-Menten kinetics, but is actually attributed to Haldane. He proved that given the standard assumptions about the catalytic energy profile, the following equation must be satisfied:

$$K'_{\text{eq}} = \frac{k_{\text{cat}}^+ \cdot K_p}{k_{\text{cat}}^- \cdot K_s}. \quad (8)$$

This equation is commonly known as the *Haldane relationship*.

1.5 The flux-force relationship

From the work of Beard and Qian [2], we know that there is a relationship between the thermodynamic driving force and the ratio of forward and backward reaction rates:

$$\frac{J^+}{J^-} = e^{-\Delta_r G' / RT} \quad (9)$$

where the forward rate (J^+) and the backward rate (J^-) are the two opposing rates of a reversible reaction. The net flux is given by:

$$J = J^+ - J^- \quad (10)$$

When a reaction is close to equilibrium, e.g. when $\Delta_r G' > -10$ kJ/mol, the backward rate can become quite large and significantly reduce the net rate. At equilibrium ($\Delta_r G' = 0$), both forward and backward rates become equal and the net rate is zero.

The effect of the flux-force relationship can be explicitly quantified in a reformulation of the reversible Haldane rate law (equation 7), which we call the separable form [16]:

$$\begin{aligned} v &= E \cdot k_{\text{cat}}^+ \cdot \eta^{\text{rev}} \cdot \eta^{\text{sat}} \\ \eta^{\text{rev}} &= 1 - e^{\Delta_r G' / RT} \\ \eta^{\text{sat}} &= \frac{s/K_s}{1 + s/K_s + p/K_p}. \end{aligned} \quad (11)$$

E is the enzyme concentration and k_{cat}^+ is the turnover rate (measured in the forward direction). The saturation term η^{sat} is similar to an (irreversible) Michaelis-Menten rate law, except that there we usually ignore the product term (p/K_p) in the denominator. The reversibility term η^{rev} is exactly the part that reflects how reversibility decreases the net flux. Both terms are always between 0 and 1.

For further reading on the flux-force relationship and how it can be used to restructure kinetic, see [17].

1.6 Linear Programming (LP)

Every Linear Programming problem (see Figure 2) can be converted to a canonical form that is:

Primal

$$\begin{aligned} &\text{maximize} && \mathbf{c}^\top \mathbf{x} \\ &\text{subject to} && \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ &&& \text{and} \quad \mathbf{x} \geq \mathbf{0}. \end{aligned} \quad (12)$$

The dual is a symmetrical linear problem described by the following constraints

Dual

$$\begin{aligned} &\text{minimize} && \mathbf{b}^\top \mathbf{y} \\ &\text{subject to} && \mathbf{A}^\top \mathbf{y} \geq \mathbf{c} \\ &&& \text{and} \quad \mathbf{y} \geq \mathbf{0}. \end{aligned} \quad (13)$$

If the primal has an optimal solution (\mathbf{x}^*), then the dual will also have an optimal solution (\mathbf{y}^*) and their values will be equal:

Strong duality theorem

$$\mathbf{c}^\top \mathbf{x}^* = \mathbf{b}^\top \mathbf{y}^* \quad (14)$$

Linear Programming problems can be solved very efficiently using established algorithms such as Simplex. Many good solvers exist for LP, including open source options such as glpk, CLP, and the internal R-project solver.

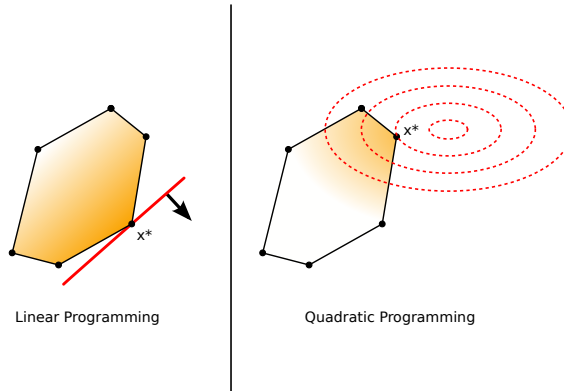


FIGURE 2: A graphical example for a problem with linear constraints, where the objective is either a linear function (Linear Programming) or a quadratic function (Quadratic Programming).

1.7 Quadratic Programming (QP)

Quadratic Programming is essentially the same as LP, except that the objective has an extra quadratic term:

$$\begin{aligned}
 &\mathbf{QP} \\
 &\text{minimize} && \mathbf{q}^\top \mathbf{x} + \frac{1}{2} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\
 &\text{subject to} && \mathbf{A} \mathbf{x} \leq \mathbf{b} \\
 &&& \text{and} \quad \mathbf{x} \geq \mathbf{0}.
 \end{aligned} \tag{15}$$

It can essentially be solved using the same algorithms as LP and is therefore easily scalable. We recommend the CVXPY package for Python users, but the Matlab function QUADPROG is an alternative.

1.8 Mixed-Integer Linear Programming (MILP)

If an LP contains variables that can only have integer values (or in a more specific case, boolean values), the problem cannot be solved using the standard methods. In fact, MILP problems in general are NP-complete, i.e. computationally hard. The free solvers are typically not very efficient in solving large MILPs. Nevertheless, advanced commercial solvers such as CPLEX and Gurobi are much more capable, and provide free licenses for academics.

2 Convex analysis of a metabolic flux cone

The basic idea underlying most metabolic constraint-based models, is mass-balance combined with the steady-state assumption. Formally, given a stoichiometric matrix \mathbf{S} , where rows represent metabolites and column represent reactions, and given a flux vector \mathbf{v} , the rate of change in metabolite levels (\mathbf{c}) is given by the equation.

$$\text{Dynamic mass-balance} \quad \frac{d\mathbf{c}}{dt} = \mathbf{S} \mathbf{v} \tag{16}$$

All metabolites are separated into two groups, internal and external, where the boundary is typically defined by the cell outer membrane. However, mathematically, external metabolites are those that have such large pools (infinite) that we can essentially ignore their change over time even if it is not 0. These include things like atmospheric gases (O_2 , CO_2) or the carbon and nitrogen sources in the medium surrounding the cells. According to this division, we can represent the stoichiometric matrix as a concatenation of two sub-matrices:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{\text{int}} \\ \mathbf{S}_{\text{ext}} \end{bmatrix} \tag{17}$$

For the subset of internal metabolites, the pseudo steady-state assumption states that their concentrations do not change over time.

Pseudo steady-state

$$\mathbf{S}_{\text{int}}\mathbf{v} = 0 \quad (18)$$

Therefore, pseudo steady-state flux solutions are contained within the null-space (or kernel) of \mathbf{S}_{int} , i.e. $\mathbf{v} \in \ker(\mathbf{S}_{\text{int}})$.

2.1 The flux cone

Typically, many of the reactions in our model will be considered irreversible, due to thermodynamic constraints (and we will elaborate on this topic in section 3). Therefore, a subset of the fluxes will be constrained to be positive:

$$\forall i \in I_{\text{irr}} \quad v_i \geq 0 \quad (19)$$

Therefore, the set of possible flux solutions would be all fluxes that satisfy both equation 18 and 19. Explicitly, it would be the intersection of all the hyperplanes that represent rows in \mathbf{S}_{int} described by $S_{j*} \cdot \mathbf{v} = 0$ and all half-spaces corresponding to irreversible reactions $\mathbf{e}_i \cdot \mathbf{v} \geq 0$ (where \mathbf{e}_i is the unit vector corresponding to reaction i). This intersection is an unbound convex polyhedron denoted the steady-state *flux cone* or \mathcal{C} [10].

It is important to note, that the flux cone is unbounded in some directions, i.e. it can extend to infinity. Therefore, we typically ignore the absolute values of the fluxes (represented, for example, by $\|\mathbf{v}\|$) and consider only the relative values – i.e. the direction towards which the vector is pointing. It is often easier to only consider normalized vectors, such as ones whose biomass rate is set to 1.

Flux cone

$$\mathcal{C} = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{S}_{\text{int}}\mathbf{v} = 0 \wedge v_i \geq 0 \quad \forall i \in I_{\text{irr}}\} \quad (20)$$

2.2 Elementary Flux Modes

Elementary Flux Modes (EFMs) are defined as the set of all non-decomposable vectors in the flux cone. A non-decomposable vector is one whose *support* (the set of reactions with non-zero flux) is minimal – i.e. the flux cone does not contain any vectors whose support is a proper subset of it. EFMs also form a convex basis for the flux cone [22]:

$$\forall \mathbf{v} \in \mathcal{C} : \exists \lambda_j \geq 0 \quad \text{s.t.} \quad \mathbf{v} = \sum_j \lambda_j \mathbf{e}^j \quad (21)$$

where $\{\mathbf{e}^j\}_j$ is the set of all EFMs. Figure 3 illustrates a model with 10 reactions and 9 EFMs.

EFMs are closely related to Extreme Pathways (EPs) and distinguishing them requires more subtle definitions. For further reading you can refer to [8].

2.3 Classification of EFMs

Even in our small toy example in Figure 3, we already encounter a common problem that arises when dealing with EFMs – i.e. cycles. In this example, it’s quite obvious that the two EFMs that consist of combining two opposing reactions is futile and should be ignored. Larger networks, however, might have much larger cycles that are harder to identify and get rid of.

First, we need to distinguish between *primary exchange* and *currency exchange* reactions. The former are the standard reactions that exchange nutrients between the cell/compartments and its environment (sugar import, CO₂ export, etc.). *Currency exchange* reactions do not, in fact, exchange material between compartments, but are rather abstract representations of energy dissipation. A canonical example would be the ATP maintenance reaction that is present in most metabolic models: $\text{ATP} + \text{H}_2\text{O} \rightarrow \text{ADP} + \text{P}_i$.

Now, we can define the three types of EFMs ([20], see Figure 4):

I – *Primary systemic EFMs* have at least one active primary exchange flux.

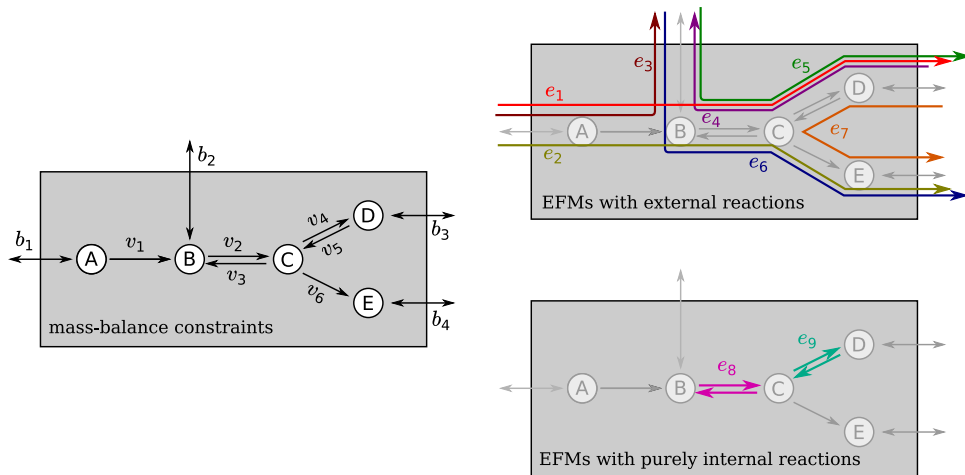


FIGURE 3: A toy example with 6 internal reactions and 4 exchange reactions. The flux cone is spanned by 9 EFMs.

II – *Futile cycles* have no active primary exchange fluxes, and at least one active currency exchange flux.

III – *Internal cycles* have no active exchange fluxes (neither primary nor currency).

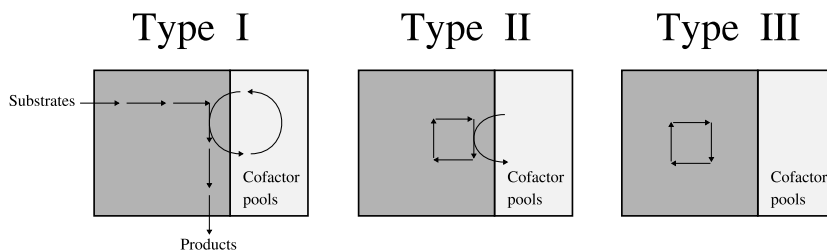


FIGURE 4: The three types of Elementary Flux Modes.

To illustrate the three types of pathways, we present another toy model with one currency exchange reaction and two primary exchange reactions (Figure 5). This model contains all the three types of EFMs.

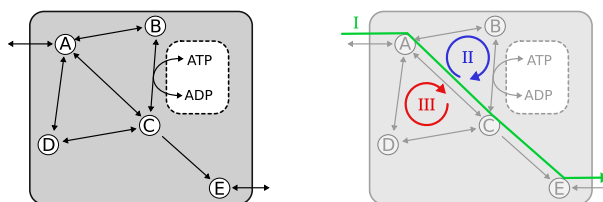


FIGURE 5: A small toy model with 8 internal reactions, 2 primary exchange reactions, and one currency exchange reaction. In this simple network, one can identify all three types of extreme pathways. The Type I pathway (green) is typically the type of solution that most constrain-based models are seeking. The Type II pathway ($A \rightarrow B \rightarrow C \rightarrow A$, blue) is a typical futile cycle, since it does not involve any primary exchange reactions, but does waste ATP. The Type III pathway ($A \rightarrow C \rightarrow D \rightarrow A$, red) is called internal since none of its reactions are exchange reactions. An internal cycle will never be thermodynamically feasible.

In the next section, we will see how this three type classification is useful to separate thermodynamically feasible pathways from infeasible ones.

3 Thermodynamics in constraint-based models

We would like to rewrite equation 4 in a more convenient way that corresponds well with our linear algebra notation. First, one can notice that taking the log from Q' makes it a linear function of the log-concentration values: $\ln(Q'_j) = \sum_i S_{ij} \cdot \ln(c_i)$. In addition, from now on we will use a matrix notation for calculating the vector of all reaction Gibbs energies ($\Delta_r \mathbf{G}'$):

$$\Delta_r \mathbf{G}' = \Delta_r \mathbf{G}'^o + RT \cdot \mathbf{S}^\top \mathbf{x} \quad (22)$$

where we define \mathbf{x} as the vector of log-concentrations, i.e. $x_i = \ln(c_i)$.

3.1 Irreversibility is an approximation

In pure physics terms, every chemical reaction should be reversible. There are cases, where the equilibrium is so far from unity, which makes it impossible for the reaction to reach equilibrium in practice. In these cases, it is sometimes convenient to assume that the reaction is irreversible and completely ignore the reverse direction. In defining the flux cone, this assumption can significantly reduce the solution space and number of EFMs.

Most thermodynamics-aware models, however, cannot make this assumption. Every reaction must have a defined equilibrium constant, no matter how extreme it might be. Therefore, from now on, we assume $\Delta_r G'_j$ has a finite value, even for reactions that are defined as irreversible in our model. There is no universal threshold for $\Delta_r G'_j$ which make a reaction irreversible. It is context dependent and a topic of much debate [11, 7, 14].

Beyond $\Delta_r G'_j$, probably the most important parameters for determining the reversibility are the range of concentrations allowed for each of the reactants. Basically, a reversible reaction should have both positive and negative values to its $\Delta_r G'_j$, depending on the chosen reactant concentrations within their predefined ranges. If we denote the vectors of lower and upper bounds as \mathbf{b}^L and \mathbf{b}^U , respectively, the constraint on \mathbf{x} would be:

$$\ln(\mathbf{b}^L) \leq \mathbf{x} \leq \ln(\mathbf{b}^U) \quad (23)$$

It is easy to see, that the lowest value for $\Delta_r G'_j$ is achieved when all substrate concentrations are set to their value in \mathbf{b}^U , and all product concentrations to \mathbf{b}^L . The highest value of $\Delta_r G'_j$ occurs in the opposite extreme. Then, it becomes straightforward to check if a reaction is, by itself, reversible or not.

3.2 Thermodynamic Feasibility of EFMs

Given a specific elementary flux mode \mathbf{e}^j (or any flux vector \mathbf{v} , for that matter), we would like to know whether it is thermodynamically feasible according to the second law. First, every one of the support reactions must be feasible in the direction defined by \mathbf{e}^j . However, the same metabolite can be a substrate for one reaction, and a product for another and therefore there is inter-dependence between the $\Delta_r G'_j$ values of these reactions. Sometimes, even though a series of reactions can be individually feasible, their combination is infeasible. This has been termed a *distributed bottleneck* by Mavrouniotis [11] in 1993.

Using Linear Programming, it is relatively simple to test thermodynamic feasibility:

Feasibility

$\exists \mathbf{x}$

such that

$$\Delta_r \mathbf{G}' \equiv \Delta_r \mathbf{G}'^o + RT \cdot \mathbf{S}^\top \mathbf{x} \quad (24)$$

$$\Delta_r \mathbf{G}' < 0$$

$$\ln(\mathbf{b}^L) \leq \mathbf{x} \leq \ln(\mathbf{b}^U)$$

3.3 Max-min Driving Force

In order to provide a more quantitative measure for the thermodynamic feasibility of a given pathway, the Max-min Driving Force (MDF) method provides a relatively simple measure. We adjust the Linear

Program in Equation 24, by adding a margin variable (B) and maximizing its value:

MDF

$$B^* = \max_{B, \mathbf{x}} B$$

such that

$$\Delta_r \mathbf{G}' \equiv \Delta_r \mathbf{G}'^o + RT \cdot \mathbf{S}^\top \mathbf{x} \quad (25)$$

$$\Delta_r \mathbf{G}' < -B$$

$$\ln(\mathbf{b}^L) \leq \mathbf{x} \leq \ln(\mathbf{b}^U)$$

If the MDF is positive, then there exists a set of concentrations \mathbf{x} (inside the allowed range) such that $\Delta_r \mathbf{G}' \leq -B < 0$, i.e. this pathway is thermodynamically feasible. The larger B is, the “more feasible” it is, since we can keep all reactions farther away from equilibrium. According to the flux-force relationship (section 1.5), the farther one is from equilibrium, the lower the backward flux is relative to the forward flux and therefore the reaction is more efficient. An example for MDF calculation is given in Figure 6.

An online interface for running MDF analysis can be found here: <https://equilibrator.weizmann.ac.il/pathway/>.

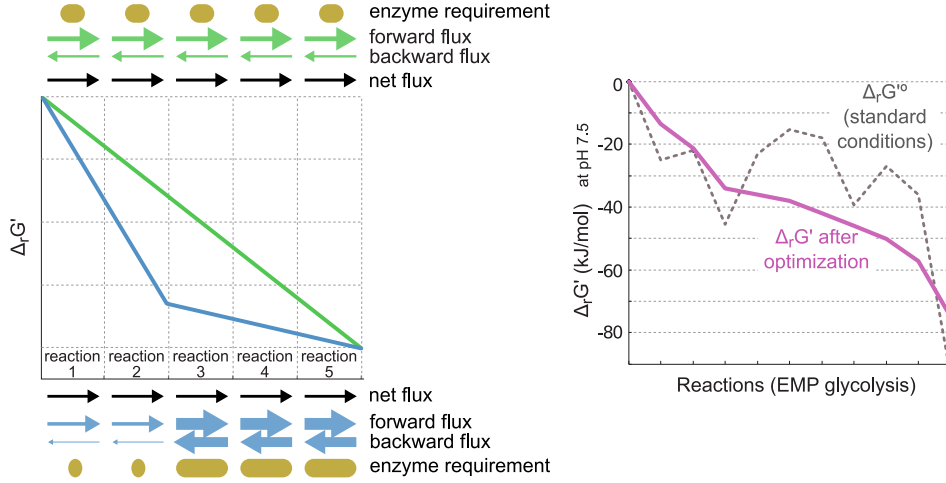


FIGURE 6: (left) Schematic comparison between two pathways. Each pathway starts and ends with the same compounds, employs five enzymes and carries the same net flux. The kinetic parameters of all enzymes in both pathways, as well as enzyme and metabolite concentrations, are assumed to be identical. (right) Energetic profile of Embden-Meyerhof-Parnas glycolysis. Dashed black line corresponds to $\Delta_r G'^o$ values (metabolite concentrations of 1M) of pathway reactions at pH 7.5. Red line corresponds to $\Delta_r G'$ values of pathway reactions after an optimization procedure that maximizes the driving force of the thermodynamic bottleneck reactions. Figure is from Noor et al. [15].

3.4 Thermodynamic Flux Balance Analysis (TFBA)

Thermodynamic FBA (also known as Thermodynamic-based Metabolic Flux Analysis [7]) was designed to deal with thermodynamically infeasible flux solutions within the framework of FBA:

FBA

$$\mathbf{v}^* = \arg \max_{\mathbf{v}} \mathbf{c}^\top \mathbf{v} \quad (26)$$

such that

$$\mathbf{S}_{\text{int}} \mathbf{v} = \mathbf{0}$$

$$\mathbf{v}_{LB} \leq \mathbf{v} \leq \mathbf{v}_{UB}$$

where $\mathbf{c} \in \mathbb{R}^r$ is the objective function, and the constants are the internal stoichiometric matrix $\mathbf{S}_{\text{int}} \in \mathbb{R}^{m \times r}$. \mathbf{v}_{LB} and \mathbf{v}_{UB} are the lower and upper bounds on the fluxes, typically relevant only for exchange fluxes.

TFBA adds another two sets of variables – the boolean flux indicators ($\mathbf{y} \in \{0, 1\}^r$), and the log-concentrations ($\mathbf{x} \in \mathbb{R}^r$). Then, the following constraints are added to the Linear Problem (which is now actually a Mixed-Integer Linear Problem – MILP):

TFBA

$$\mathbf{v}^* = \arg \max_{\mathbf{v}} \mathbf{c}^\top \mathbf{v}$$

such that

$$\begin{aligned} \mathbf{S}_{\text{int}} \mathbf{v} &= \mathbf{0} \\ \mathbf{v}_{LB} &\leq \mathbf{v} \leq \mathbf{v}_{UB} \\ 0 &\leq M\mathbf{y} - \mathbf{v} \leq M \end{aligned} \quad (27)$$

$$0 < M\mathbf{y} + \Delta_r \mathbf{G}' < M \quad (28)$$

$$\Delta_r \mathbf{G}' = \Delta_r \mathbf{G}'^o + RT \cdot \mathbf{S}^\top \mathbf{x} \quad (29)$$

$$\ln(\mathbf{b}^L) \leq \mathbf{x} \leq \ln(\mathbf{b}^U) \quad (30)$$

The new constants are the vector of standard Gibbs energies of reaction $\Delta_r \mathbf{G}'^o$ (in units of kJ/mol), the gas constant $R = 8.31$ J/mol/K and temperature $T = 300$ K. M which is a very large number (larger than any of the possible flux and Gibbs free energy). Note also the difference between the internal stoichiometric matrix (\mathbf{S}_{int}) used for the pseudo steady-state assumption, versus the full stoichiometric matrix (\mathbf{S}) used in the Gibbs free energy calculation.

Equations 29-30 should be familiar, and are exactly the same as in previous sections. To understand how the other two constraints (27-28) enforce thermodynamic feasibility, we consider three possible cases for each reaction j separately:

1. $v_j > 0$: the only possible value that y_j can have is 1, otherwise, $M y_j - v_j$ would be negative and violate constraint 27. Therefore, constraint 28 becomes $0 < M + \Delta_r G'_j < M$, which means that $\Delta_r G'_j < 0$.
2. $v_j < 0$: the only possible value that y_j can have is 0, otherwise, $M y_j - v_j$ would be larger than M and violate constraint 27. Therefore, constraint 28 becomes $0 < \Delta_r G'_j < M$, which means that $\Delta_r G'_j > 0$.
3. $v_j = 0$: both 0 and 1 are possible solutions for y_j . Therefore, there are no constraints on $\Delta_r G'_j$.

Summarizing these 3 cases, one can concisely write:

$$\forall j : v_j = 0 \vee \text{sign}(v_j) = -\text{sign}(\Delta_r G'_j), \quad (31)$$

which is exactly the second law of thermodynamics.

3.5 Loopless Flux Balance Analysis (ll-FBA)

The loopless algorithm [21] is very similar to TFBA, except that there are no actual thermodynamic values. This way, thermodynamically infeasible internal (Type III) cycles are eliminated, while all other pathways are kept [19]. The set of equations describing ll-FBA are:

ll-FBA

$$\mathbf{v}^* = \arg \max_{\mathbf{v}} \mathbf{c}^\top \mathbf{v}$$

such that

$$\mathbf{S}_{\text{int}} \mathbf{v} = \mathbf{0} \quad (32)$$

$$\mathbf{v}_{LB} \leq \mathbf{v} \leq \mathbf{v}_{UB} \quad (33)$$

$$0 \leq M\mathbf{y} - \mathbf{v} \leq M \quad (34)$$

$$0 < M\mathbf{y} + \Delta_r \mathbf{G}' < M \quad (35)$$

$$\Delta_r \mathbf{G}' \in \ker(\mathbf{S})^\perp \quad (36)$$

here, $\Delta_r \mathbf{G}'$ is not constrained by the $\Delta_r \mathbf{G}'^o$ and the metabolite concentrations, but is only required to be orthogonal to the null-space of \mathbf{S} (or, equivalently, to be in $\text{image}(\mathbf{S}^\top)$).

Why does ll-FBA not eliminate type II cycles from the set of flux solutions? These futile cycles “waste” resources such as ATP, but are thermodynamically feasible. They are only considered “cycles” because we chose to give co-factors a special status (namely, they are external metabolites that do not need to be kept at steady-state). In other words, type II EFMs are in the null-space of \mathbf{S}_{int} , but not in the null-space of \mathbf{S} (just like type I EFMs). Therefore, it is possible to assign negative $\Delta_r G'$ values to all reactions in a type II EFMs.

Interestingly, this argument can be applied also to the reverse of such futile cycle. Running a futile cycle in reverse is sometimes called an Energy Generating Cycle (EGC), and is obviously unrealistic. Furthermore, unlike type III cycles that do not affect the biomass rate in standard FBA, EGCs have the potential to increase the maximal biomass yield and pose a more serious problem in FBA models [5]. This is one of the cases where TFBA differs from ll-FBA, as it prevents the use of EGCs completely, while ll-FBA doesn't.

4 Gibbs energy dissipation

4.1 The Gibbs energy dissipation rate of a reaction

For a single reaction, the Gibbs energy dissipation rate of a reaction i is simply the negation of its $\Delta_r G'$ multiplied by the net flux:

$$g_i = -\Delta_r G'_i \cdot v_i \quad (37)$$

Note that if $v_i \neq 0$, then the second law of thermodynamics states that $\Delta_r G'_i$ and v_i must have opposite signs – therefore $g_i \geq 0$ is always true.

4.2 The Total Gibbs energy dissipation rate

The total Gibbs energy dissipation rate of an entire cell is defined as $g^{\text{diss}} \equiv \sum_i g_i$. If we represent fluxes in units per gram of dry biomass (e.g. $\mu\text{mol}/\text{min}/\text{gr}$), then g^{diss} can be converted to units of Watts per gram (J/s/gr). Niebel, Leupold, and Heinemann [13] hypothesized that the total dissipation rate per unit of biomass might be capped, and showed how adding this constraint to an existing metabolic model can help predictions. To calculate g^{diss} , they started with a standard thermodynamic model (such as in TFBA – see section 3.4) and added the extra constraints needed for the upper bound on total dissipation:

TFBA[†]

$$\mathbf{v}^* = \arg \max_{\mathbf{v}} \mathbf{c}^\top \mathbf{v}$$

such that

$$\mathbf{S}_{\text{int}} \mathbf{v} = \mathbf{0}$$

$$\mathbf{v}_{LB} \leq \mathbf{v} \leq \mathbf{v}_{UB}$$

$$0 \leq M\mathbf{y} - \mathbf{v} \leq M \quad (38)$$

$$0 < M\mathbf{y} + \Delta_r \mathbf{G}' < M$$

$$\Delta_r \mathbf{G}' = \Delta_r \mathbf{G}'^o + RT \cdot \mathbf{S}^\top \mathbf{x}$$

$$\ln(\mathbf{b}^L) \leq \mathbf{x} \leq \ln(\mathbf{b}^U)$$

$$g^{\text{diss}} = -\Delta_r \mathbf{G}'^\top \mathbf{v}$$

$$g^{\text{diss}} \leq g_{\text{max}}^{\text{diss}}$$

[†] – with an upper bound on Gibbs energy dissipation rate.

Sometimes, it might be useful to calculate g^{diss} without committing to a full TFBA model (which requires collecting all reaction energies and solving it using an MILP solver). Here, we'll see how the system in 38 can be converted to a simple LP if we are fine with ignoring the second law constraints of internal reactions ($\Delta_r G' \leq 0$).

First, we recall that any set of reaction Gibbs energies can be represented by formation energies multiplied by the stoichiometric matrix:

$$\Delta_r \mathbf{G}' = \mathbf{S}^\top \cdot \Delta_f \mathbf{G}' \quad (39)$$

Then, the total dissipation rate can be rewritten as:

$$g^{\text{diss}} = -\Delta_r \mathbf{G}'^\top \mathbf{v} = -(\mathbf{S}^\top \cdot \Delta_f \mathbf{G}')^\top \mathbf{v} = -\Delta_f \mathbf{G}'^\top \mathbf{S} \mathbf{v}. \quad (40)$$

The full stoichiometric matrix is a concatenation of the external (\mathbf{S}_{ext}) and internal matrix (\mathbf{S}_{int}), so if we consider the steady-state assumption ($\mathbf{S}_{\text{int}} \mathbf{v} = \mathbf{0}$), we can continue to simplify the expression such that:

$$g^{\text{diss}} = -\Delta_f \mathbf{G}'^\top \begin{bmatrix} \mathbf{S}_{\text{int}} \\ \mathbf{S}_{\text{ext}} \end{bmatrix} \mathbf{v} = -\Delta_f \mathbf{G}'^\top \begin{bmatrix} \mathbf{S}_{\text{int}} \mathbf{v} \\ \mathbf{S}_{\text{ext}} \mathbf{v} \end{bmatrix} = -\Delta_f \mathbf{G}'^\top \begin{bmatrix} \mathbf{0} \\ \mathbf{S}_{\text{ext}} \mathbf{v} \end{bmatrix} = -\Delta_f \mathbf{G}'_{\text{ext}}^\top \mathbf{S}_{\text{ext}} \mathbf{v}, \quad (41)$$

where $\Delta_f \mathbf{G}'_{\text{ext}}$ is the vector of formation energies for the external metabolites (a sub-vector of $\Delta_f \mathbf{G}'$). Since the concentration of external metabolites is fixed by definition, $\Delta_f \mathbf{G}'_{\text{ext}}$ will have fixed values that can usually be computed directly. Using this result, we can formulate the standard FBA problem with an upper bound on the total Gibbs energy dissipation rate:

FBA

$$\mathbf{v}^* = \arg \max_{\mathbf{v}} \mathbf{c}^\top \mathbf{v}$$

such that

$$\begin{aligned} \mathbf{S}_{\text{int}} \mathbf{v} &= \mathbf{0} \\ \mathbf{v}_{LB} &\leq \mathbf{v} \leq \mathbf{v}_{UB} \\ g_{\text{max}}^{\text{diss}} &\geq -\Delta_f \mathbf{G}'_{\text{ext}}^\top \mathbf{S}_{\text{ext}} \mathbf{v} \end{aligned} \quad (42)$$

4.2.1 OptKnock and its derivatives

Burgard, Pharkya, and Maranas [3] phrased the OptKnock optimization problem first as a bi-level mixed-integer linear optimization problem. The inner problem is standard FBA, maximizing an objective function $\mathbf{c}_{\text{in}}^\top \mathbf{v}$. This encodes the assumption that each knockout strain we consider for the outer problem, evolves to maximize the objective (which is typically set to be the biomass function). Then, we choose the one knockout that maximizes the outer objective $\mathbf{c}_{\text{out}}^\top \mathbf{v}$ – typically the secretion rate of the desired product (e.g. succinate).

OptKnock

$$\max_{\mathbf{y}} \mathbf{c}_{\text{out}}^\top \mathbf{v}$$

subject to

$$\begin{aligned} \max_{\mathbf{v}} \mathbf{c}_{\text{in}}^\top \mathbf{v} \\ \text{subject to} \end{aligned}$$

$$\mathbf{S}_{\text{int}} \mathbf{v} = \mathbf{0}$$

$$v_{\text{LB},j} \cdot y_j \leq v_j \leq y_j \cdot v_{\text{UB},j}$$

$$\mathbf{y} \in \{0, 1\}^r$$

$$\sum_j (1 - y_j) \leq K \quad (43)$$

where, K is the maximum number of allowed knockouts. The inner LP is a standard FBA optimization problem, except that a few of the reactions are “knockout out” – when a reaction indicator is set to 0 by the external LP it forces the corresponding flux to zero, i.e. $y_j = 0 \rightarrow v_j = 0$. In addition, we set $v_{\text{LB}, \text{biomass}} > 0$ and $y_{\text{biomass}} = 1$, since we only want to consider mutants that show at least some minimal growth. Finally, the external LP places an upper bound on the total number of knockouts, which is useful for two reasons: (1) making sure we don’t have too many knockouts that would make the genetic manipulation impractical, and (2) to reduce the computational load required to solve the MILP. Nevertheless, even for small K , solving the bi-level MILP directly is not tractable for large networks (such as the genome-scale *E. coli* model). Fortunately, Burgard, Pharkya, and Maranas [3] provided a solution, by converting into a single-level MILP using duality theory.

The dual of the inner optimization problem in 43 is:

$$\min_{\lambda, \mu, \omega} \sum_j y_j (v_{\text{UB},j} \cdot \mu_j - v_{\text{LB},j} \cdot \omega_j)$$

subject to

$$\mathbf{S}_{\text{int}}^\top \lambda + \mu - \omega \geq \mathbf{c}_{\text{in}}$$

$$\mu, \omega \geq 0. \quad (44)$$

Since duality theory states if the optimum solutions of the primal and dual problems are bounded, their objective function values must be equal to one another, so ensuring optimality can be achieved by equating the two objectives: $\mathbf{c}_{\text{in}}^\top \mathbf{v} = \sum_j y_j (v_{\text{UB},j} \cdot \mu_j - v_{\text{LB},j} \cdot \omega_j)$. However, using this equation as a constraint in the outer optimization problem would make it non-linear, since it contains products of two variables ($y_j \cdot \mu_j$ and $y_j \cdot \omega_j$). We can, however, use the fact that y_j are binary in order to cast this expression into a linear form.

First, we introduce a new set of auxiliary variables $\mathbf{m} \in \mathbb{R}^r$, and constrain them to be $m_j = y_j \cdot \mu_j$ by adding the following:

$$\begin{aligned} 0 &\leq m_j \leq y_j \cdot M \\ \mu_j - (1 - y_j) \cdot M &\leq m_j \leq \mu_j. \end{aligned} \quad (45)$$

As before, M is a constant larger than any possible value of μ_j . Considering the two cases for y_j , one can easily see that $y_j = 0 \rightarrow m_j = 0$ and $y_j = 1 \rightarrow m_j = \mu_j$, which is what we wanted to achieve. We do exactly the same for ω and another set of variables $\mathbf{u} \in \mathbb{R}^r$. Then we can rewrite the dual objective as: $\sum_j y_j (v_{\text{UB},j} \mu_j - v_{\text{LB},j} \omega_j) = \mathbf{v}_{\text{UB}}^\top \mathbf{m} - \mathbf{v}_{\text{LB}}^\top \mathbf{u}$.

Using this notion, we get a single-level MILP problem for OptKnock:

$$\begin{aligned} &\mathbf{OptKnock} \\ &\max_{\mathbf{y}, \mathbf{v}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{m}, \mathbf{u}} \quad \mathbf{c}_{\text{out}}^\top \mathbf{v} \\ &\text{subject to} \\ &\quad \mathbf{S}_{\text{int}} \mathbf{v} = \mathbf{0} \\ &\quad v_{\text{LB},j} \cdot y_j \leq v_j \leq y_j \cdot v_{\text{UB},j} \\ &\quad \mathbf{c}_{\text{in}}^\top \mathbf{v} = \mathbf{v}_{\text{UB}}^\top \mathbf{m} - \mathbf{v}_{\text{LB}}^\top \mathbf{u} \\ &\quad \mathbf{c}_{\text{in}} \leq \mathbf{S}_{\text{int}}^\top \boldsymbol{\lambda} + \boldsymbol{\mu} - \boldsymbol{\omega} \\ &\quad 0 \leq \boldsymbol{\mu}, \boldsymbol{\omega} \\ &\quad 0 \leq \mathbf{m} \leq \mathbf{y} \cdot M \\ &\quad \boldsymbol{\mu} - (1 - \mathbf{y}) \cdot M \leq \mathbf{m} \leq \boldsymbol{\mu} \\ &\quad 0 \leq \mathbf{u} \leq \mathbf{y} \cdot M \\ &\quad \boldsymbol{\omega} - (1 - \mathbf{y}) \cdot M \leq \mathbf{u} \leq \boldsymbol{\omega} \\ &\quad \sum_j (1 - y_j) \leq K \end{aligned} \quad (46)$$

4.3 A few relevant algorithms

4.3.1 Phenotypic Phase Plane (PPP)

This useful concept derived from FBA, is a term for 2D projections on the flux polytope. Typically, one chooses specific reaction (e.g. exchange of succinate) for the y-axis and the FBA objective (i.e. biomass rate) for the x-axis. The projection of the flux polytope onto this plane creates a convex polygon shape and helps to visualize how important a certain flux is for creating biomass. Often for cases where the selected flux is a fermentation product, the PPP is referred to as a *production envelope*. It is standard practice to present strain-design results by overlaying two or more PPPs, e.g. a knockout strain versus the wild-type (Figure 7). Using such comparisons, it is easy to see if the knockout is useful for forcing a cell to produce the desired chemical byproduct, by coupling its production to the biomass rate.

4.3.2 Derivatives of OptKnock

A comprehensive review by Machado and Herrgård [9] lists the many published algorithms for *in silico* strain design that followed after OptKnock (see Figure 8).

RobustKnock is an improvement for OptKnock, introduced by Tepper and Shlomi [23] in 2010, which accounts for competing pathways that might reduce the chemical production rates. One common issue with standard OptKnock, is that it only looks at the maximal rate of the chemical production rate: $\max \mathbf{c}_{\text{out}}^\top \mathbf{v}$. However, as visualized by the production envelope in Figure 7a, there might be redundancy in the solution space, where the same biomass rate can be achieved without any production of the desired chemical. RobustKnock addresses this problem by performing a max-min optimization, i.e. maximizing the lowest value of $\mathbf{c}_{\text{out}}^\top \mathbf{v}$ over all \mathbf{v} that satisfy the inner LP. This method yields solutions like in Figure 7b-c.

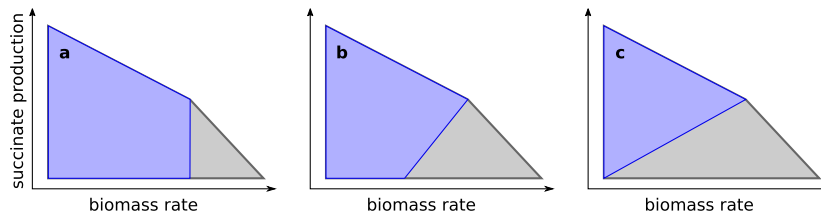


FIGURE 7: Phenotypic phase planes for anaerobic succinate production from glucose based on the *E. coli* core model. (a) wild-type strain (light gray) vs. triple-deletion mutant (*ACKr*, *ATPS4r*, *FUM*) resulting in a design without growth-coupling (purple). (b) wild-type strain vs. triple-deletion mutant (*ACALD*, *PYK*, *ME2*) resulting in a partially growth-coupled design. (c) wild-type strain vs. double-deletion mutant (*ACALD*, *LDH-D*) resulting in a fully growth-coupled design. Figure is from Machado and Herrgård [9].

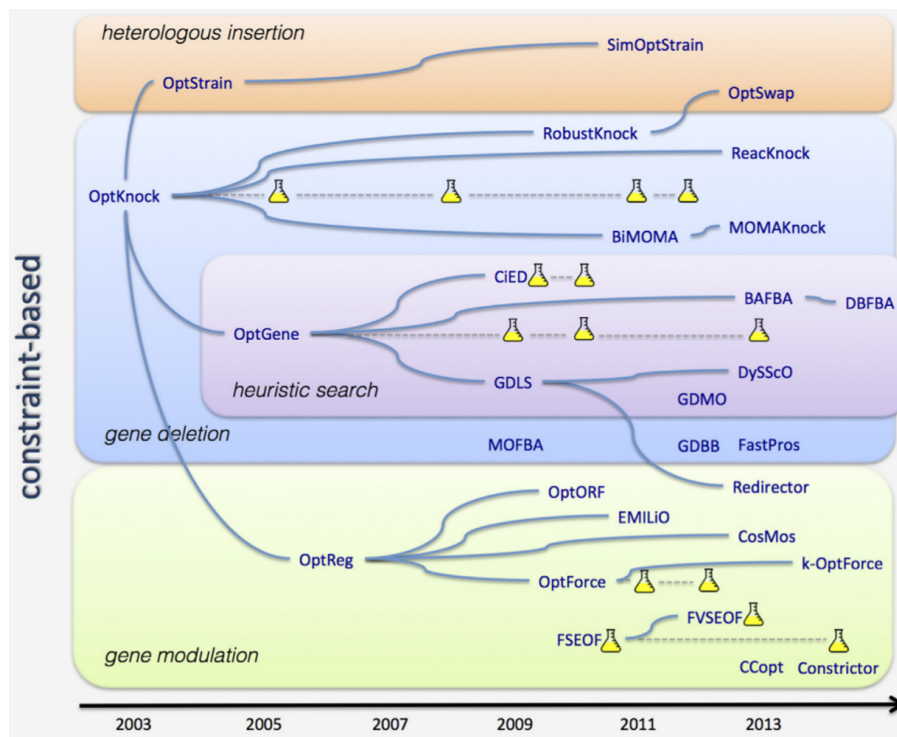


FIGURE 8: A timeline of all constraint-based strain design methods that followed OptKnock (figure is from [9]).

4.3.3 Probabilistic Thermodynamic Analysis (PTA)

Abstract: Random sampling of metabolic fluxes can provide a comprehensive description of the capabilities of a metabolic network. However, current sampling approaches do not model thermodynamics explicitly, leading to inaccurate predictions of an organism’s potential or actual metabolic operations. We present a probabilistic framework combining thermodynamic quantities with steady-state flux constraints to analyze the properties of a metabolic network. It includes methods for probabilistic metabolic optimization and for joint sampling of thermodynamic and flux spaces. Applied to a model of *Escherichia coli*, we use the methods to reveal known and novel mechanisms of substrate channeling, and to accurately predict reaction directions and metabolite concentrations. Interestingly, predicted flux distributions are multimodal, leading to discrete hypotheses on *E. coli*’s metabolic capabilities. Python and MATLAB packages available at <https://gitlab.com/csb.ethz/pta>. Supplementary data are available at Bioinformatics online [6].

References

- [1] Robert A. Alberty. *Biochemical Thermodynamics*. Hoboken, NJ, USA: John Wiley & Sons, Inc., Mar. 2006. ISBN: 0-471-75798-5.
- [2] Daniel A. Beard and Hong Qian. “Relationship between Thermodynamic Driving Force and One-Way Fluxes in Reversible Processes”. In: *PLoS ONE* 2.1 (Jan. 3, 2007), e144.
- [3] Anthony P. Burgard, Priti Pharkya, and Costas D. Maranas. “Optknock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization”. In: *Biotechnology and Bioengineering* 84.6 (Dec. 20, 2003), pp. 647–657. ISSN: 1097-0290.
- [4] Sadi Carnot. *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance*. fr. Bachelier Libraire, 1824.
- [5] Claus Jonathan Fritzscheier, Daniel Hartleb, Balázs Szappanos, Balázs Papp, and Martin J. Lercher. “Erroneous Energy-Generating Cycles in Published Genome Scale Metabolic Networks: Identification and Removal”. In: *PLOS Computational Biology* 13.4 (Apr. 18, 2017), e1005494. ISSN: 1553-7358.
- [6] Mattia G Gollub, Hans-Michael Kaltenbach, and Jörg Stelling. “Probabilistic thermodynamic analysis of metabolic networks”. In: *Bioinformatics* 37.18 (Sept. 2021), pp. 2938–2945.
- [7] Christopher S. Henry, Linda J. Broadbelt, and Vassily Hatzimanikatis. “Thermodynamics-Based Metabolic Flux Analysis”. In: *Biophysical Journal* 92.5 (Mar. 1, 2007), pp. 1792–1805. ISSN: 0006-3495.
- [8] Steffen Klamt and Jörg Stelling. “Two Approaches for Metabolic Pathway Analysis?” In: *Trends in Biotechnology* 21.2 (Feb. 1, 2003), pp. 64–69. ISSN: 0167-7799, 1879-3096. pmid: 12573854.
- [9] Daniel Machado and Markus J. Herrgård. “Co-Evolution of Strain Design Methods Based on Flux Balance and Elementary Mode Analysis”. In: *Metabolic Engineering Communications* 2 (Dec. 1, 2015), pp. 85–92. ISSN: 2214-0301.
- [10] Sayed-Amir Marashi, Laszlo David, and Alexander Bockmayr. “Analysis of Metabolic Subnetworks by Flux Cone Projection”. In: *Algorithms for Molecular Biology* 7 (May 29, 2012), p. 17. ISSN: 1748-7188.
- [11] Michael L Mavrovouniotis. “Identification of Localized and Distributed Bottlenecks in Metabolic Pathways”. In: *Proceedings. International Conference on Intelligent Systems for Molecular Biology* 1 (1993), pp. 275–283. ISSN: 1553-0833. pmid: 7584346.
- [12] Alan Mellors. “The haldane relationship enzymes & equilibrium”. en. In: *Biochem. Educ.* 4.4 (Oct. 1976), p. 71.
- [13] Bastian Niebel, Simeon Leupold, and Matthias Heinemann. “An upper limit on Gibbs energy dissipation governs cellular metabolism”. En. In: *Nature Metabolism* 1.1 (Jan. 2019), p. 125. ISSN: 2522-5812.
- [14] Elad Noor, Arren Bar-Even, Avi Flamholz, Yaniv Lubling, Dan Davidi, and Ron Milo. “An Integrated Open Framework for Thermodynamics of Reactions That Combines Accuracy and Coverage”. In: *Bioinformatics* 28.15 (Aug. 1, 2012), pp. 2037–2044. ISSN: 1367-4803.
- [15] Elad Noor, Arren Bar-Even, Avi Flamholz, Ed Reznik, Wolfram Liebermeister, and Ron Milo. “Pathway Thermodynamics Highlights Kinetic Obstacles in Central Metabolism”. In: *PLOS Computational Biology* 10.2 (Feb. 20, 2014), e1003483. ISSN: 1553-7358.
- [16] Elad Noor, Avi Flamholz, Arren Bar-Even, Dan Davidi, Ron Milo, and Wolfram Liebermeister. “The Protein Cost of Metabolic Fluxes: Prediction from Enzymatic Rate Laws and Cost Minimization”. In: *PLOS Computational Biology* 12.11 (Nov. 3, 2016), e1005167. ISSN: 1553-7358.
- [17] Elad Noor, Avi Flamholz, Wolfram Liebermeister, Arren Bar-Even, and Ron Milo. “A Note on the Kinetics of Enzyme Action: A Decomposition That Highlights Thermodynamic Effects”. In: *FEBS Letters. A century of Michaelis - Menten kinetics* 587.17 (Sept. 2, 2013), pp. 2772–2777. ISSN: 0014-5793.
- [18] Elad Noor, Hulda S. Haraldsdóttir, Ron Milo, and Ronan M. T. Fleming. “Consistent Estimation of Gibbs Energy Using Component Contributions”. In: *PLOS Computational Biology* 9.7 (July 2013), e1003098. ISSN: 1553-7358.

- [19] Elad Noor, E Nathan Lewis, and Ron Milo. “A Proof for Loop-Law Constraints in Stoichiometric Metabolic Networks”. In: *BMC Syst. Biol.* 6 (Jan. 2012).
- [20] Nathan D. Price, Iman Famili, Daniel A. Beard, and Bernhard Ø Palsson. “Extreme Pathways and Kirchhoff’s Second Law”. In: *Biophysical Journal* 83.5 (Nov. 1, 2002), pp. 2879–2882. ISSN: 0006-3495. PMID: 12425318.
- [21] Jan Schellenberger, Nathan E. Lewis, and Bernhard Ø. Palsson. “Elimination of Thermodynamically Infeasible Loops in Steady-State Metabolic Models”. In: *Biophysical Journal* 100.3 (Feb. 2, 2011), pp. 544–553. ISSN: 0006-3495.
- [22] Christophe H. Schilling, David Letscher, and Bernhard Ø. Palsson. “Theory for the Systemic Definition of Metabolic Pathways and Their Use in Interpreting Metabolic Function from a Pathway-Oriented Perspective”. In: *Journal of Theoretical Biology* 203.3 (Apr. 7, 2000), pp. 229–248. ISSN: 0022-5193.
- [23] Naama Tepper and Tomer Shlomi. “Predicting Metabolic Engineering Knockout Strategies for Chemical Production: Accounting for Competing Pathways”. In: *Bioinformatics* 26.4 (Feb. 15, 2010), pp. 536–543. ISSN: 1367-4803.