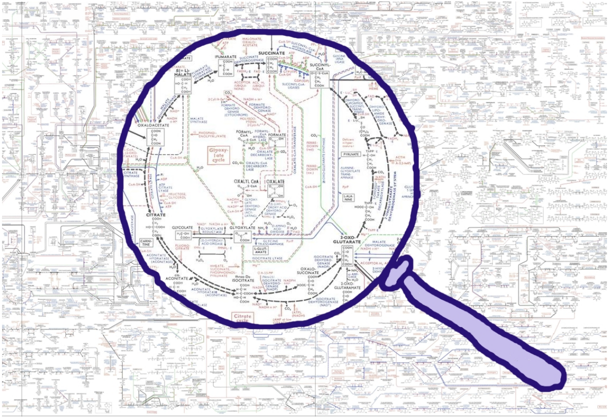
Parameter distributions for kinetic models derived from kinetic, thermodynamic, and metabolic data

Wolfram Liebermeister, MPI-MG Berlin

BTK meeting 2006, Trakai
Session *Modelling the living cell - part II*

From metabolic networks to kinetic models



Our aim, as a starting point for kinetic modelling:
Obtain sensible kinetic parameters from literature knowledge

Image: Roche Applied Science "Biochemical Pathways"

Kinetic data stored in data bases

Parameter		mean $e^{\langle x \rangle}$		e^{σ_x}	# samples, ref
Energy parameter	g	0.79		1.2	142 (1)
Equilibrium constant	k^{eq}	-		212	1309 (2)
Turnover rate	k^{cat}	7.0	1/s	27.1	7559 (3)
M.M. constant	k^{M}	0.17	mM	20.1	44766 (3)
Inhibition constant	k^{I}	0.06	mM	60.3	4338 (3)
Protein mol./cell		2480		4.7	3868 (4)
Protein concentration	E	$3.6 \cdot 10^{-5}$	mM	4.7	3868 (4)
Metab. concentration	c	0.14	mM	7.0	49 (5)

(1) Mavrovouniotis 1990, Biotechnology and Bioengineering, 1070-1082
(2) http://xpdb.nist.gov/enzyme_thermodynamics/
(3) http://www.brenda.uni-koeln.de/
(4) http://yeastgfp.ucsf.edu/
(5) Albe et al 1990, J. Theor. Biol. 143, 163-195

Parameters automatically extracted from PubMed:

http://sysbio.molgen.mpg.de/KMedDB

Mining parameters for metabolic networks

Problems with literature data

- Conflicting data
- Thermodynamic dependencies → additional hidden conflicts
- Missing data
- Uncertainties in data (measurement, biological variation)

Solution:
Obtain balanced parameters by Bayes estimation,
accounting for special properties of metabolic systems

- Build a model with simple kinetics → parameter dependencies
- Find sets of system parameters that agree with literature data
- Compute joint parameter distribution (instead of a single set)
- Sample and analyse model instances

Which kinetic law do we employ?

The "convenience" kinetics

A general reaction formula:

$$\alpha_1\ A_1 + \alpha_2\ A_2 + \dots \quad \longleftrightarrow \quad \beta_1\ B_1 + \beta_2\ B_2 + \dots,$$

Convenience kinetic law

$$v(\mathbf{a},\mathbf{b}) \quad = \quad E\,\frac{k_{+}^{\mathrm{cat}}\ \prod_i (\frac{a_i}{k_{a_i}^{\mathrm{M}}})^{\alpha_i} - k_{-}^{\mathrm{cat}}\ \prod_j (\frac{b_j}{k_{b_j}^{\mathrm{M}}})^{\beta_j}}{\prod_i (1 + (\frac{a_i}{k_{a_i}^{\mathrm{M}}}) + \dots + (\frac{a_i}{k_{a_i}^{\mathrm{M}}})^{\alpha_i}) + \prod_j (1 + (\frac{b_j}{k_{b_j}^{\mathrm{M}}}) + \dots + (\frac{b_j}{k_{b_j}^{\mathrm{M}}})^{\beta_j}) - 1}$$

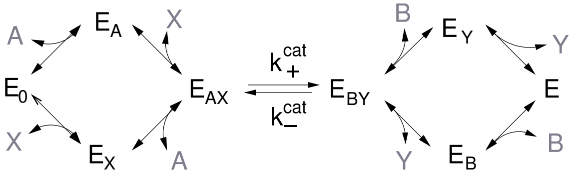
Prefactors for activators/inhibitors

$$h_{\mathrm{A}}(d, k^{\mathrm{A}}) \quad = \quad \frac{d}{k^{\mathrm{A}} + d} \qquad \text{(Activation)}$$

$$h_{\mathrm{I}}(d, k^{\mathrm{I}}) \quad = \quad \frac{k^{\mathrm{I}}}{k^{\mathrm{I}} + d}. \qquad \text{(Inhibition)}$$

Enzyme mechanism behind convenience kinetics

Example reaction: $A + X \rightarrow B + Y$



Assumptions:

- Binding of substrates / products → fast equilibrium
- Slow conversion from substrate to product complex
- No interaction terms in binding energies

The k^{M} values are dissociation constants

The kinetic parameters are constrained by thermodynamics

Haldane relation for equilibrium constant

$$k_l^{\mathrm{eq}} = \frac{k_+^{\mathrm{cat}}}{k_-^{\mathrm{cat}}} \frac{\prod_j (k_{b_j}^{\mathrm{M}})^{\beta_j}}{\prod_i (k_{a_i}^{\mathrm{M}})^{\alpha_i}}$$

k^{eq} fixed $\rightarrow$ parameters are constrained within reaction

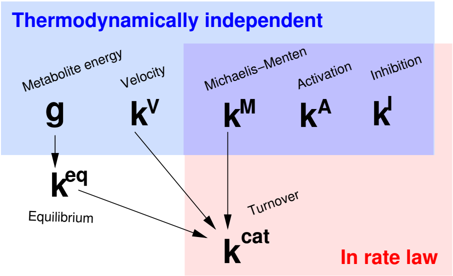
Equilibrium const. depend on Gibbs free energies of formation

$$k_l^{\mathrm{eq}} = \mathrm{e}^{-\Delta G_l^*/(RT)} \qquad \text{where} \quad \Delta G_l^* = \sum_i n_{il} G_i^{(0)}$$

$\rightarrow$ constraints between distant parameters

VERY BAD FOR PARAMETER FITTING, OPTIMISATION ETC.

The solution: thermodynamically independent parameters

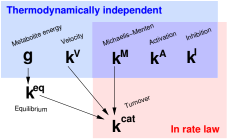


New *basic* parameters

$$g_i \quad = \quad \mathrm{e}^{G_i^{(0)}/(RT)}, \qquad k_l^{\mathrm{V}} = \sqrt{k_{+l}^{\mathrm{cat}}\ k_{-l}^{\mathrm{cat}}}$$

Alternative: use independent subset of k^{eq} values

The parameter dependence graph contains linear relations



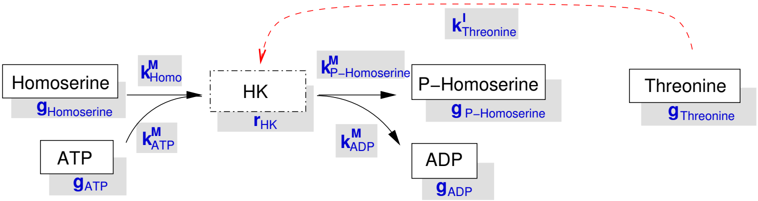
$$x \quad = \quad R^x_\theta \quad \theta$$

logarithmic kinetic data

logarithmic independent parameters

Dependent parameter	Formula
Equilibrium constant	$\log k^{\mathrm{eq}}_l = -\sum_i n_{il} \log g_i$
Turnover rates	$\log k^{\mathrm{cat}}_{\pm l} = \log k^{\mathrm{V}}_l \mp \frac{1}{2}\sum_i n_{il}(\log g_i + \log k^{\mathrm{M}}_{li})$
Maximal velocities	$\log v^{\mathrm{max}}_{\pm l} = \log E_l + \log k^{\mathrm{V}}_l \mp \frac{1}{2}\sum_i n_{il}(\log g_i + \log k^{\mathrm{M}}_{li})$

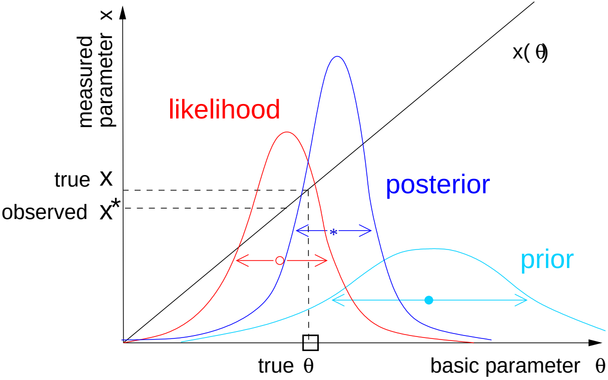
Independent system parameters in the network graph



Parameters in homoserine-kinase reaction

How to obtain parameter values and ranges?

Bayesian parameter estimation:
combining general beliefs with information from noisy data



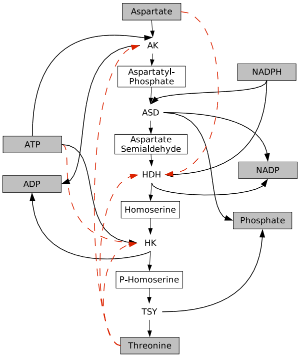
θ: log. model parameters

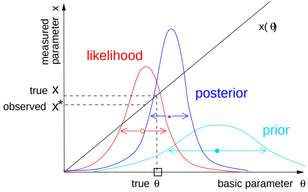
x^*: log. kinetic data

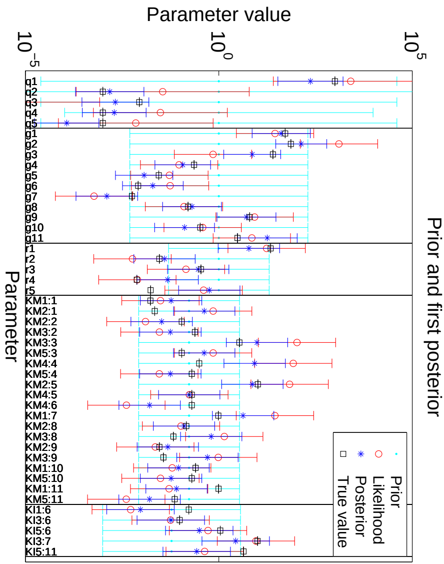
$$p(\theta|x^*) \sim p(x^*|\theta) \; p(\theta)$$

posterior likelihood prior

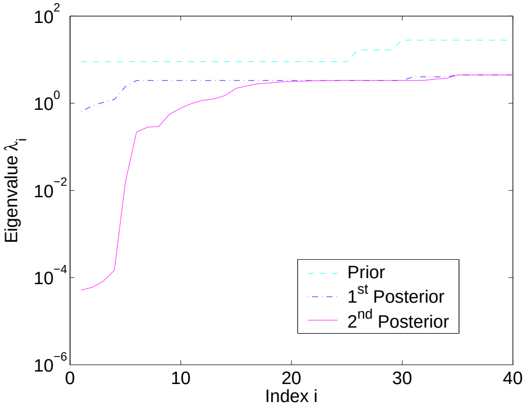
Parameter estimation from artificial data threonine model





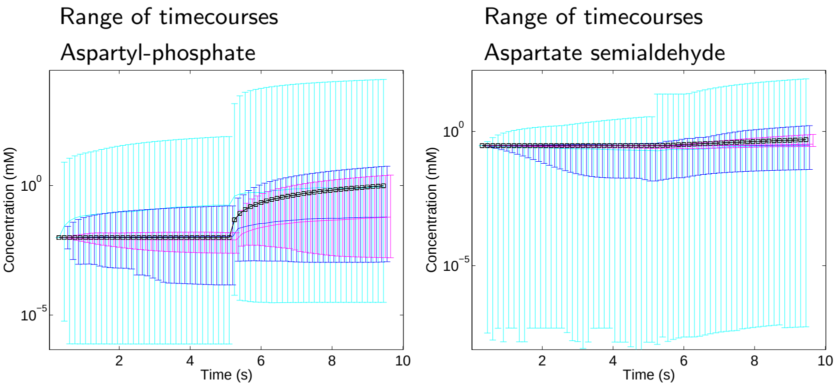


Data integration narrows the joint parameter distribution



Eigenvalues of covariance matrix for
prior, kinetics-based posterior, metabolics-based posterior

Data integration improves model predictions



Original, prior, kinetics-based posterior, metabolics-based posterior

Summary: Main ingredients of the data integration method

System description

- Standard kinetics
- Logarithmic parameters
- Thermodynamically independent system parameters

Parameter balancing

- Joint parameter distribution
- Priors from empirical parameter distributions
- Gaussian distributions (for log values)

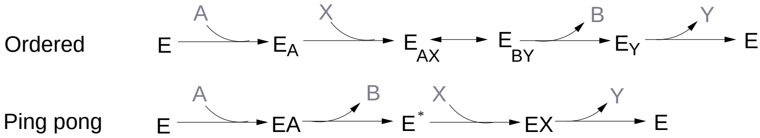
Kinetic Modelling Group, MPI-MG

Edda Klipp

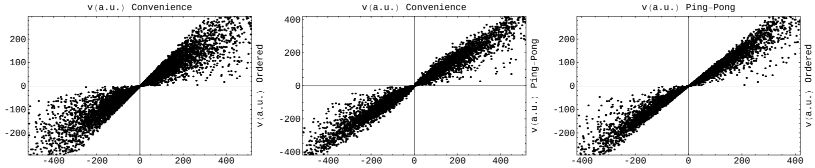
Jörg Schaber
Simon Borger
René Hoffmann
Marija Cvijovic
Zhike Zi
Christian Waltermann
Anselm Helbig

www.molgen.mpg.de/∼ag_klipp/

Comparison to ordered and ping-pong mechanism



Monte Carlo results
(Log-uniform random parameters / concentrations)



Convenience kinetic laws

Reaction formula	Rate law	Therm. indep. form
$A \leftrightarrow B$	$\frac{k_+^{cat}\tilde{a} - k_-^{cat}\tilde{b}}{1+\tilde{a}+\tilde{b}}$	$k^V \frac{\hat{a}-\hat{b}}{1+\tilde{a}+\tilde{b}}$
$A + X \leftrightarrow B$	$\frac{k_+^{cat}\tilde{a}\tilde{x} - k_-^{cat}\tilde{b}}{1+\tilde{a}+\tilde{x}+\tilde{a}\tilde{x}+\tilde{b}}$	$k^V \frac{\hat{a}\hat{x}-\hat{b}}{1+\tilde{a}+\tilde{x}+\tilde{a}\tilde{x}+\tilde{b}}$
$A + X \leftrightarrow B + Y$	$\frac{k_+^{cat}\tilde{a}\tilde{x} - k_-^{cat}\tilde{b}\tilde{y}}{1+\tilde{a}+\tilde{x}+\tilde{a}\tilde{x}+\tilde{b}+\tilde{y}+\tilde{b}\tilde{y}}$	$k^V \frac{\hat{a}\hat{x}-\hat{b}\hat{y}}{1+\tilde{a}+\tilde{x}+\tilde{a}\tilde{x}+\tilde{b}+\tilde{y}+\tilde{b}\tilde{y}}$
$2\,A \leftrightarrow B$	$\frac{k_+^{cat}\tilde{a}^2 - k_-^{cat}\tilde{b}}{1+\tilde{a}+\tilde{a}^2+\tilde{b}}$	$k^V \frac{\hat{a}^2-\hat{b}}{1+\tilde{a}+\tilde{a}^2+\tilde{b}}$
$2\,A \leftrightarrow B + Y$	$\frac{k_+^{cat}\tilde{a}^2 - k_-^{cat}\tilde{b}\tilde{y}}{1+\tilde{a}+\tilde{a}^2+\tilde{b}+\tilde{y}+\tilde{b}\tilde{y}}$	$k^V \frac{\hat{a}^2-\hat{b}\hat{y}}{1+\tilde{a}+\tilde{a}^2+\tilde{b}+\tilde{y}+\tilde{b}\tilde{y}}$
$2\,A + X \leftrightarrow B$	$\frac{k_+^{cat}\tilde{a}^2\tilde{x} - k_-^{cat}\tilde{b}}{(1+\tilde{a}+\tilde{a}^2)(1+\tilde{x})+\tilde{b}}$	$k^V \frac{\hat{a}^2\hat{x}-\hat{b}}{(1+\tilde{a}+\tilde{a}^2)(1+\tilde{x})+\tilde{b}}$

Abbreviations: $\tilde{a} = a/k_A^M$ $\hat{a} = a(g_A/k_A^M)^{1/2}$