



Systems Biology for the development of microbial cell factories

Eugénio Campos Ferreira



IBB – Institute for Biotechnology and Bioengineering
Centre of Biological Engineering, BioPSE group
Universidade do Minho



Systems Biology approaches for modelling, optimization, and control of microbial cell factories

- Cellular Models for Metabolic Engineering: gene networks
- Inference of Biological Networks
 - From Genome-scale metabolic models
 - From experimental data
 - From literature data mining
- *In Silico* Metabolic Engineering Platforms: Optimization of Microbial strains – OptFlux tool





- Metabolic Engineering can gain major benefits from the systems biology approach
- Systems Biology does not investigate individual cellular components at a time, but the behaviour and relationships of all of the elements in a particular biological system while it is functioning



Systems biology involves the use of computer simulations of cellular subsystems (such as the networks of metabolites and enzymes which comprise metabolism, signal transduction pathways and gene regulatory networks) to both analyze and visualize the complex connections of these cellular processes.

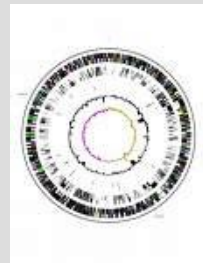
INTRODUCTION METABOLIC ENGINEERING

INTRODUCTION
CELLULAR MODELS
INFERENCE OF BIOLOGICAL
NETWORKS
OPTIMIZATION TOOLS

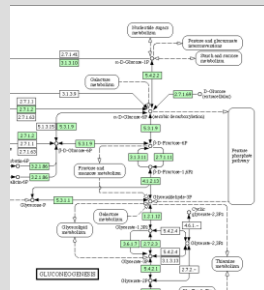
- In order to economically produce desired compounds like antibiotics, therapeutic proteins, food and feed ingredients, fuels, vitamins and other chemicals from microbial cell factories it is generally necessary to retrofit the metabolism
- Metabolic engineering envisages the introduction of directed genetic modifications leading to desirable metabolic phenotypes, as opposed to traditionally used random mutagenesis and screening



Cell factory



Genome



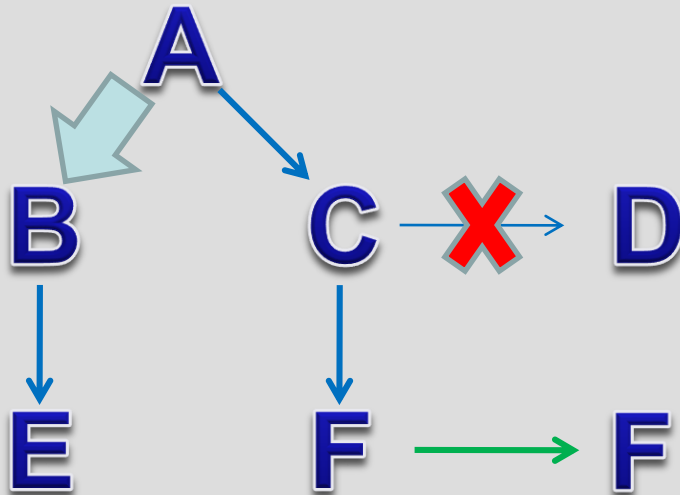
Metabolism / Phenotype



INTRODUCTION

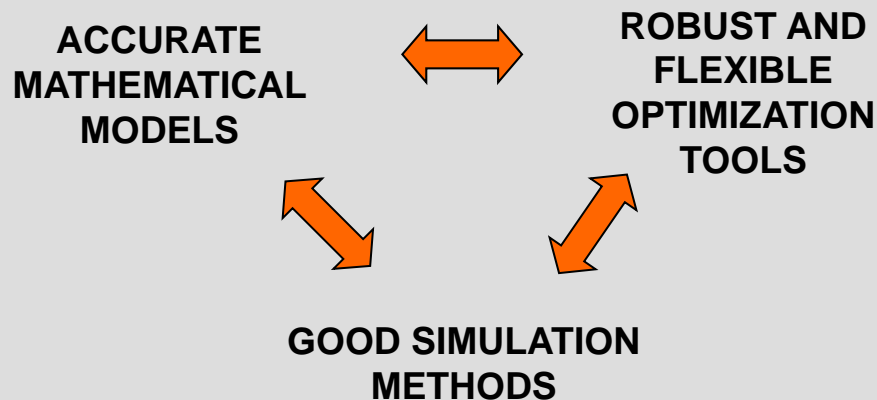
METABOLIC ENGINEERING STRATEGIES

- Gene Deletion
- Gene Addition
- Gene Under/Overexpression
- Manipulation of environmental conditions





- In metabolic engineering problems, it is often difficult to identify a priori which genetic manipulations will originate a given desired phenotype
- In order to rationally design production strains with enhanced capabilities, it is essential to have:



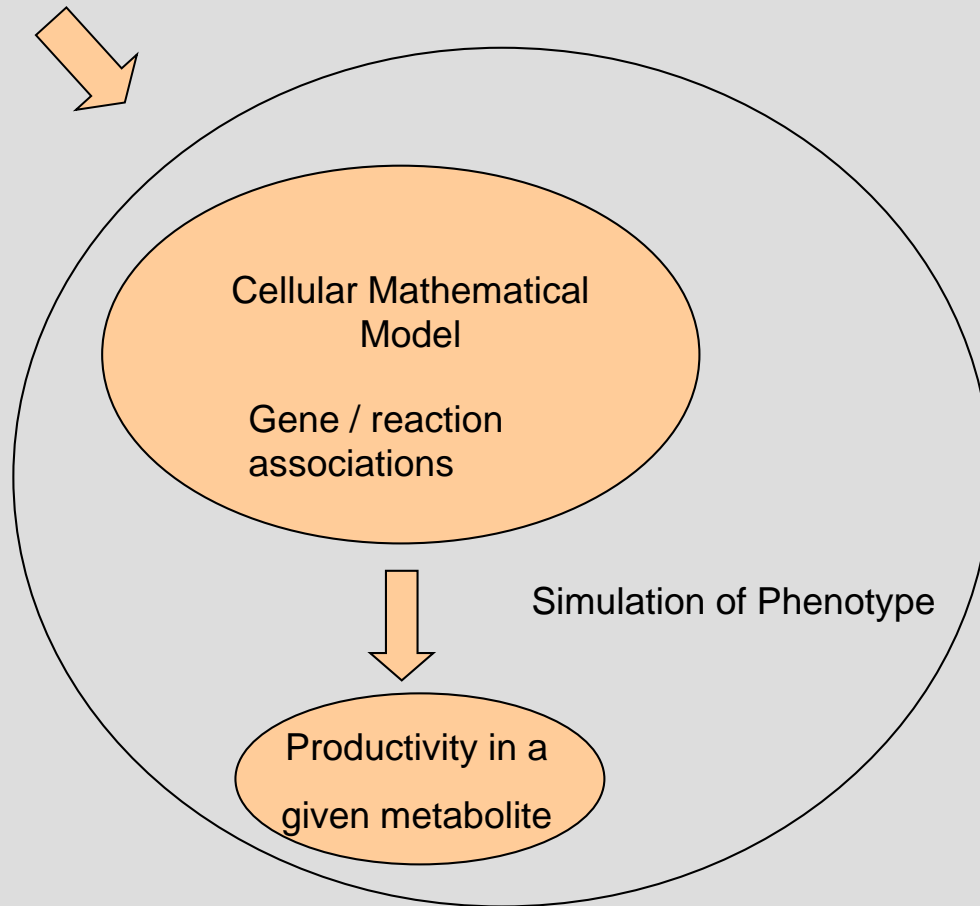
INTRODUCTION METABOLIC ENGINEERING



INTRODUCTION
CELLULAR MODELS
INFERENCE OF BIOLOGICAL
NETWORKS
OPTIMIZATION TOOLS

Optimization
Algorithm

Genome modifications and manipulation of
environmental conditions
Objective function – productivity in a given metabolite



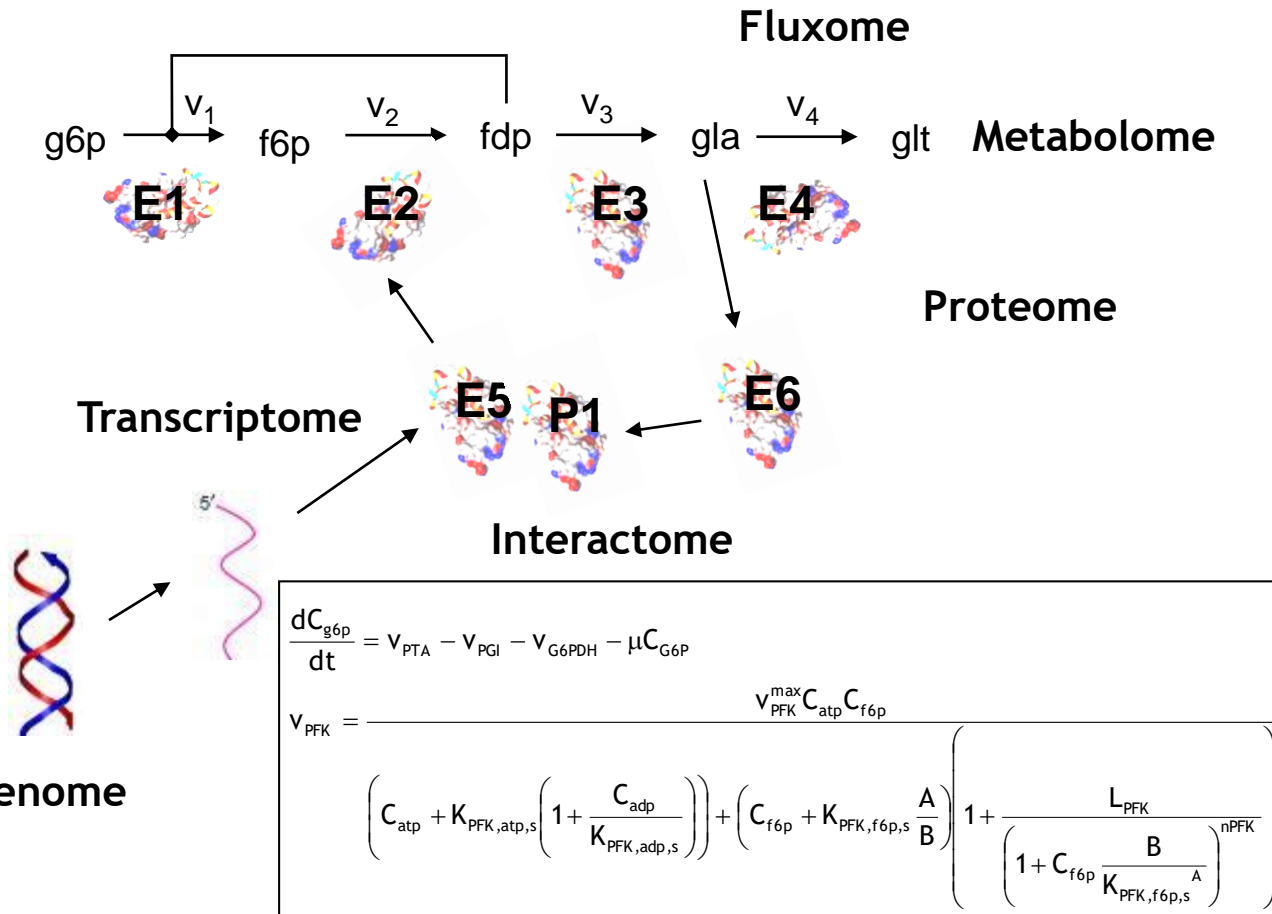
CELLULAR MODELS LEVELS OF INFORMATION

μ $S_{,max}$ $S + K_d$

INTRODUCTION
CELLULAR MODELS
INFERENCE OF BIOLOGICAL
NETWORKS
OPTIMIZATION TOOLS

Models should comprise different levels of information:

- reactions stoichiometry
- reactions kinetics
- regulatory information

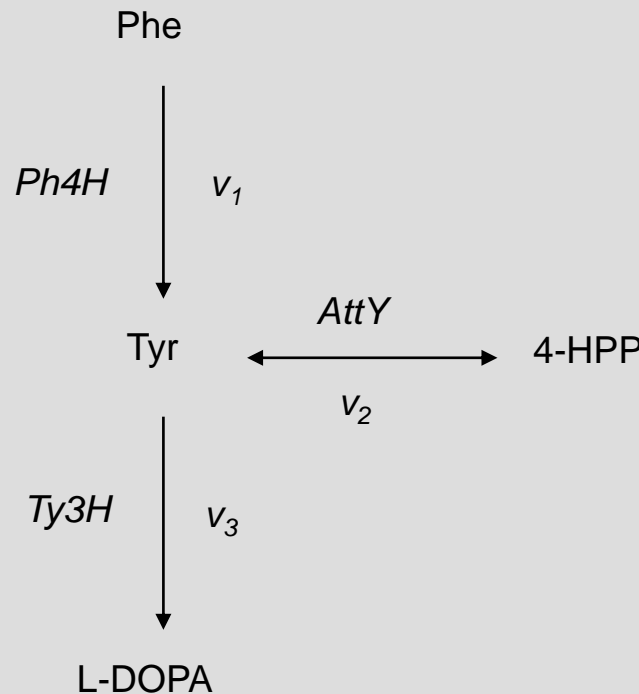
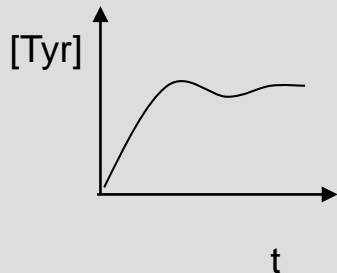


Genome



- There are several ways to represent the **chemical conversions** associated to metabolic reactions
- Kinetic or mechanistic models use **deterministic** differential equations relating the amount of reactants with the quantity of products, according to a given reaction rate and other parameters
- Given an initial state, the trajectory of metabolite can be obtained by numerical simulation

$$\frac{d[\text{Tyr}]}{dt} = \dots$$

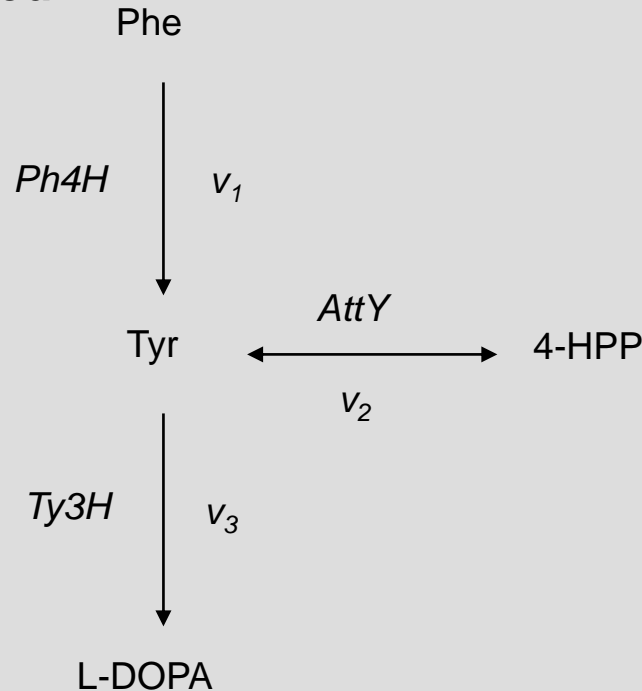




- A (pseudo)steady state condition is usually assumed inside the cell

$$\frac{d[\text{Tyr}]}{dt} = v_1 - v_2 - v_3$$

$$v_1 - v_2 - v_3 \approx 0$$



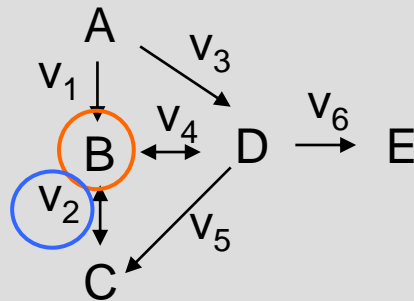
This procedure is repeated for *all* considered metabolites and will originate the so-called stoichiometric model

The result is a Linear Equations system described by stoichiometric matrix *S*.

$$S v = 0$$

$$\beta_j \leq v_j \leq \alpha_j$$

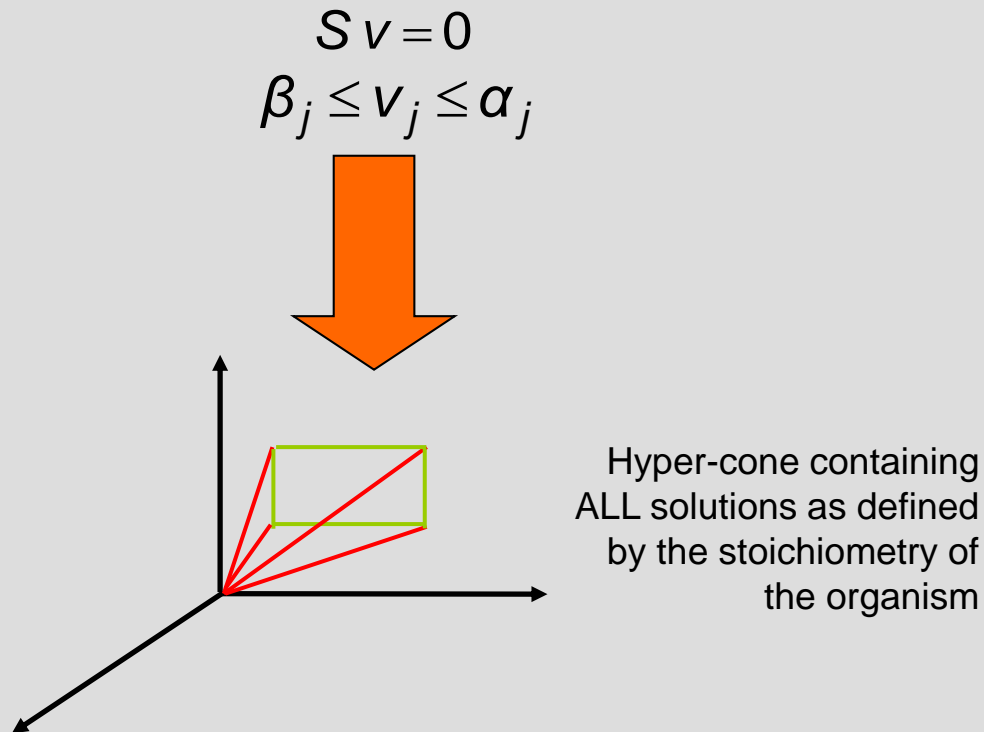
For an identified reaction set:



$$S = \begin{matrix} & v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & \dots & \text{flux } n \\ \begin{bmatrix} -1 & 0 & -1 & 0 & 0 & 0 & \dots & 0 \\ 1 & -1 & 0 & -1 & 0 & 0 & \dots & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & 1 & -1 & -1 & \dots & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & \begin{matrix} A \\ B \\ C \\ D \\ E \\ \dots \\ \text{metabolite } m \end{matrix} \end{matrix}$$

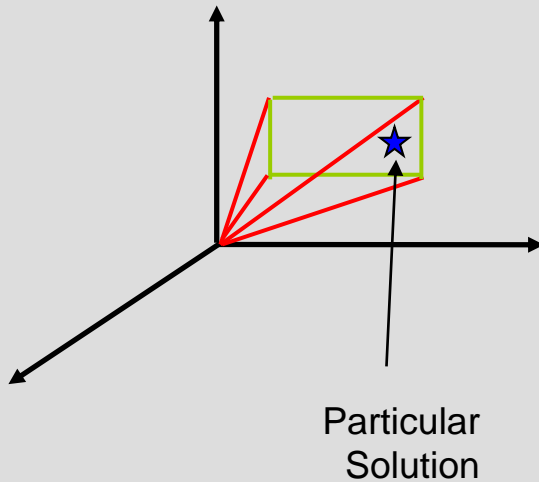
$$\begin{aligned} 0 &\leq v_1 \leq +\infty \\ -\infty &\leq v_2 \leq +\infty \\ 0 &\leq v_3 \leq +\infty \\ -\infty &\leq v_4 \leq +\infty \\ 0 &\leq v_5 \leq +\infty \\ 0 &\leq v_6 \leq +\infty \\ &\dots \\ \beta &\leq v_n \leq \alpha \end{aligned}$$

- Stoichiometric models typically have more fluxes than balanced metabolites.
- The equation system, $S \cdot v = 0$, then has more variables than equations. This is a so-called under-determined equation system with infinitely many solutions:

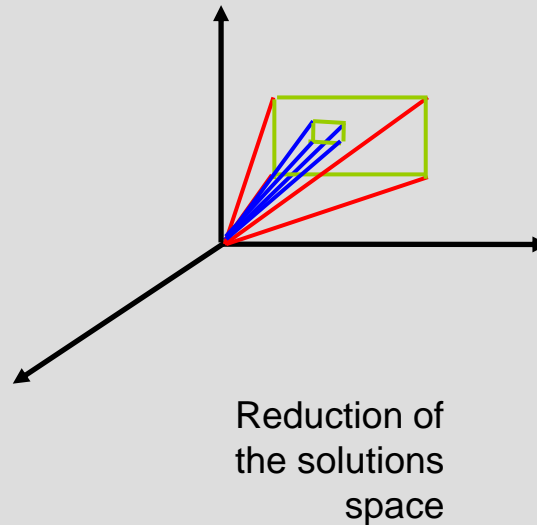




- How can we reduce the cone of solutions?
 - By optimizing a given criterion – FBA, MoMA, ROOM...
 - By the introduction of regulatory information (ex: Gene Networks)



Particular
Solution



Reduction of
the solutions
space

FBA: Flux Balance Analysis
ROOM: Regulatory On/Off
Minimization
MoMA: Minimization of Metabolic
Adjustment



- The idea is to find *one* solution to the under-determined system

$S \cdot v = 0$ by **optimization of a given criterion.**

Maximize:

$$z = c^T v = v_{prod}$$

c = row vector containing weights specifying what combination of fluxes to optimize

Subject to:

$$S v = 0$$

Constraints from stoichiometry

$$\beta_j \leq v_j \leq \alpha_j$$

- A vector containing the values of each individual metabolic flux is obtained

LINEAR
PROGRAMMING
PROBLEM!

BUT WHAT SHOULD WE OPTIMIZE?

- Studies in several organisms demonstrated that their metabolic network has evolved for optimization of the specific growth rate under several carbon source limiting conditions
- Thus, for simulating cellular behaviour, the most common objective function is the maximization of biomass production (BPCY: Biomass-Product Coupled Yield)



- For mutants and organisms grown on unusual carbon sources the hypothesis of optimal growth is not always real
- Such strains may undergo minimal redistribution of fluxes with respect to the wild-type strains (MoMA)
- The problem is the search of a flux set (x) that has a minimal distance from the wild-type flux vector (w) obtained with FBA.
- The distance between w and x is given by the Euclidean distance:

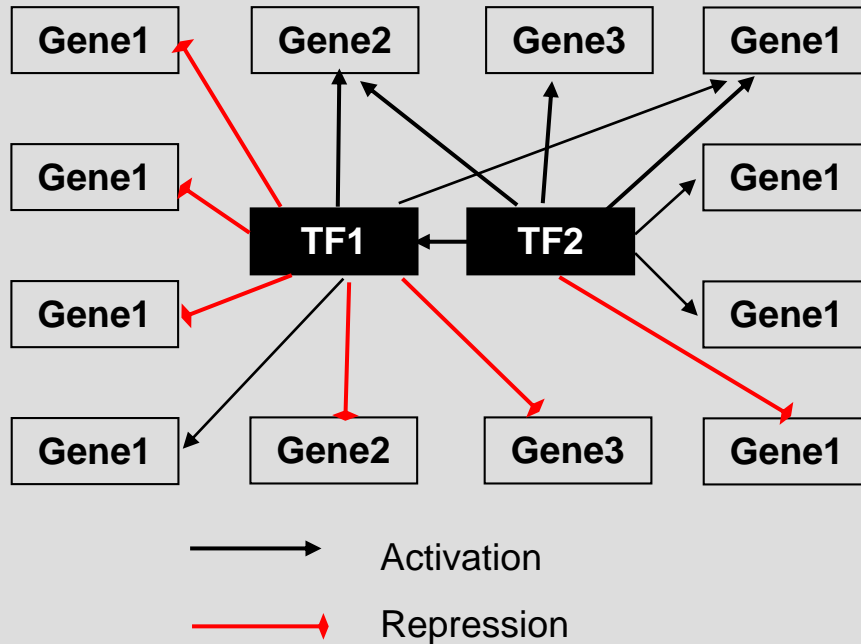
$$D(w, x) = \sqrt{\sum_{i=1}^N (w_i - x_i)^2}$$

- The minimization of that distance can be formulated as a QP problem

Segre *et al.* (2002), PNAS

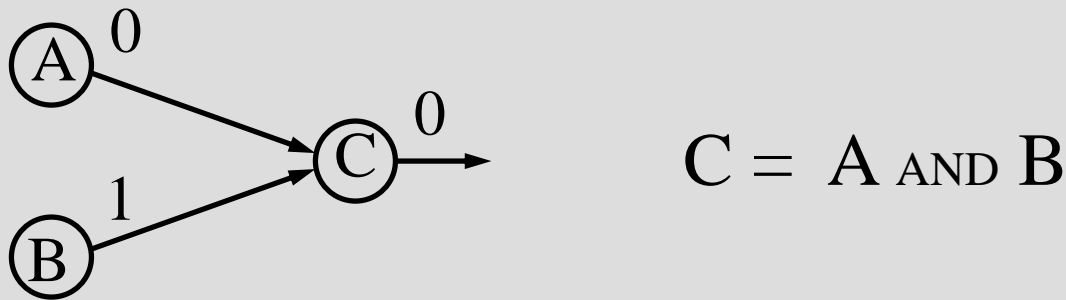


- Gene Regulatory Networks represent regulatory elements and their interactions
- A regulatory network will direct the activation or repression of a set of genes in response to a specific environmental stimulus, like O₂ or pH





- The simulation of a genetic network can be performed in several ways
- The simplest one is to consider Boolean Networks, where ON/OFF gene states are assumed.



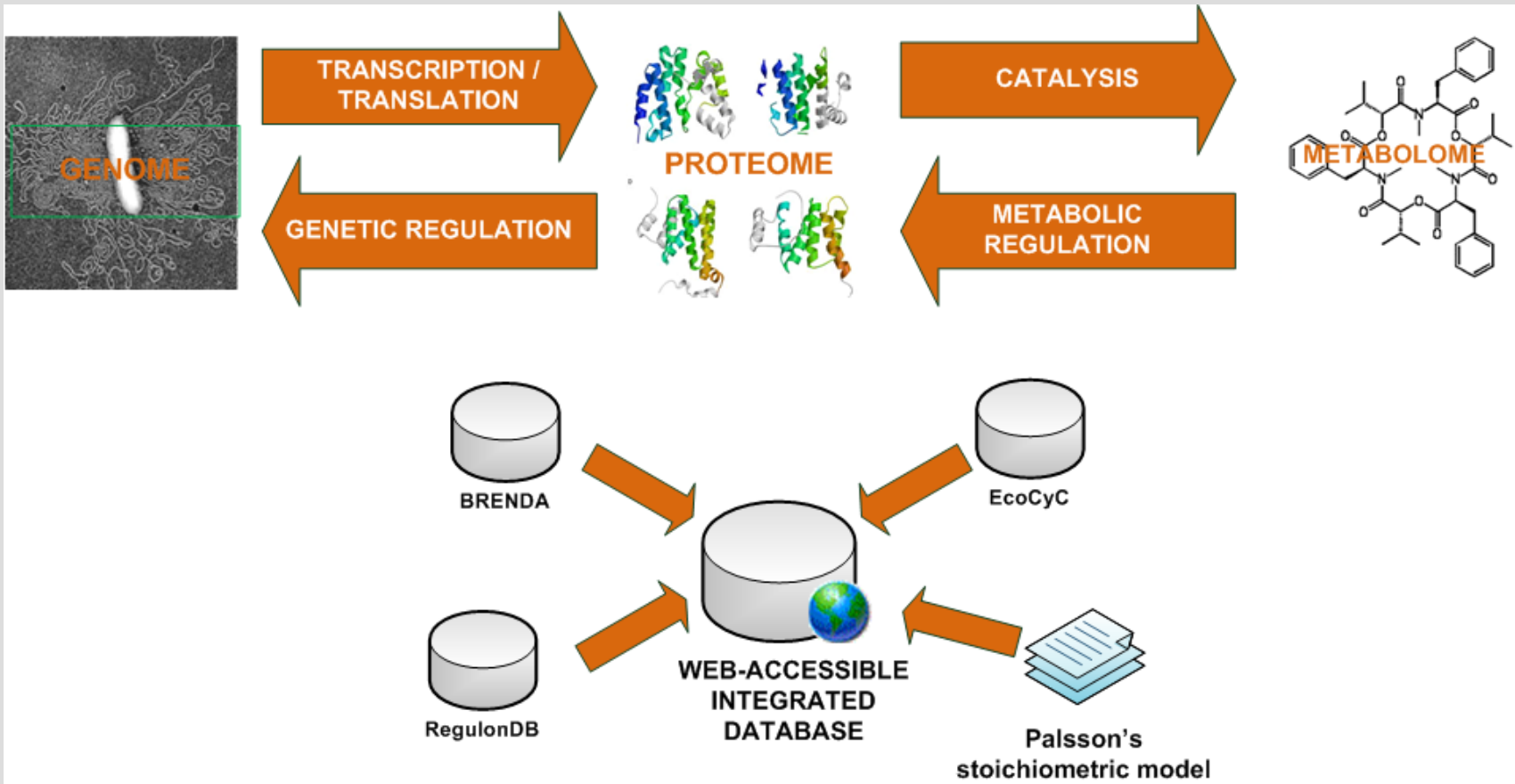
- This approach:
 - Allows analysis at the network-level
 - Provides useful insights in network dynamics

CELLULAR MODELS

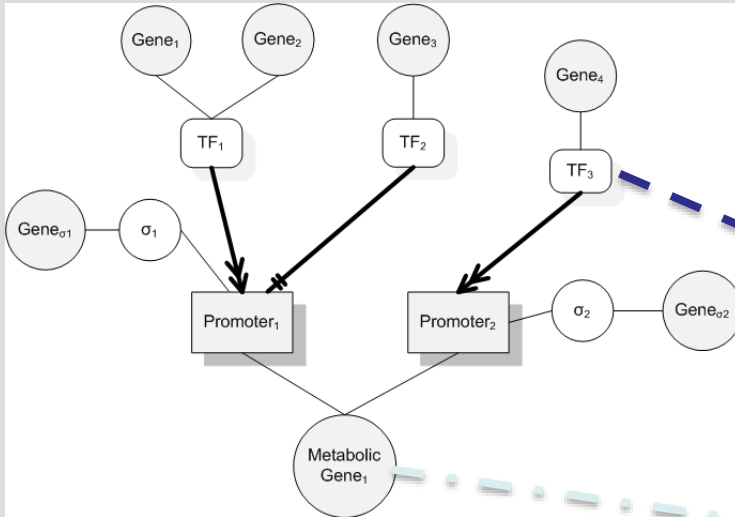
GENE NETWORKS



WORK IN PROGRESS:
Reconstruction of *E. coli*
regulatory network and integration
with stoichiometric model



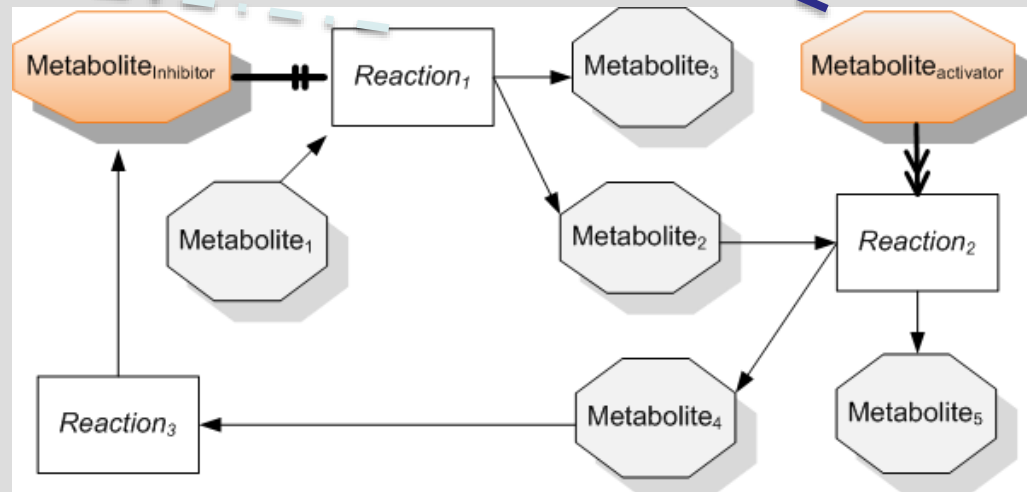
WORK IN PROGRESS: Reconstruction of *E. coli* regulatory network and integration with stoichiometric model



Metabolic
Transcriptional
Regulation

Inhibition /
Activation

Genetic
Regulation



CELLULAR MODELS *GENE NETWORKS*

WORK IN PROGRESS: Information System of
Biochemical and Regulatory Data on *Escherichia coli*



Research Activities on Bioinformatics



- Home
- Projects
- Collaborations
- Publications
- Resources
- Members
- Contacts

Welcome analia! — [Logout](#)

Regulatory Models

This is our **database of regulatory models** for *Escherichia coli*. Its primary goal is to assist researchers in their daily activities, providing the means to manage the catalogue as well as advanced analysis features. Database management covers the insertion, update and removal of information at each entity level, i.e., genes, transcription factors, pathways, reactions, regulatory events and so on. While providing some means of consultation and search these options are not meant for analysis purposes. Research needs in terms of data crossover and analysis led to the construction of more sophisticated and specific reports.

Database Management

- External Factors
- Genes
- Metabolites
- Pathways
- Reactions
- Regulatory Events
- Ri functions
- Sigma factors
- Transcription Factors

Reports

- Encoding Genes
- Metabolites
- Reactions
- Transcription Factors

© All rights reserved.

CELLULAR MODELS GENE NETWORKS

WORK IN PROGRESS: Information System of
Biochemical and Regulatory Data on *Escherichia coli*



Research Activities on Bioinformatics

Home
Projects
Collaborations
Publications
Resources
Members
Contacts
Welcome analia! — Logout

Regulatory Models

Database Management

- External Factors
- Genes
- Metabolites
- Pathways
- Reactions
- Regulatory Events
- Ri functions
- Sigma factors
- Transcription Factors

Reports

- Encoding Genes
- Metabolites
- Reactions
- Transcription Factors

Reactions

[New Record](#) | [Regulated Gene](#) | [Equation](#) | [Export TXT](#) | [Print](#) | [Refresh](#)

2008-5-15 (14:08)

Records: 0 - 20
Total of Records: 931

	Pathway	Reaction	ECnumber	Equation
	Alanine and aspartate metabolism	ALAR	EC-5.1.1.1	[c]ala-L <==> ala-D
	Alanine and aspartate metabolism	ALARi	EC-5.1.1.1	[c]ala-L --> ala-D
	Alanine and aspartate metabolism	ALATA_L	EC-2.6.1.2	[c]akg + ala-L <==> glu-L + pyr
	Alanine and aspartate metabolism	ASNN	EC-3.5.1.1	[c]asn-L + h2o --> asp-L + nh4
	Alanine and aspartate metabolism	ASNS1	EC-6.3.5.4	[c]asp-L + atp + gln-L + h2o --> amp + asn-L + glu-L + h + ppi
	Alanine and aspartate metabolism	ASNS2	EC-6.3.1.1	[c]asp-L + atp + nh4 --> amp + asn-L + h + ppi
	Alanine and aspartate metabolism	ASPT	EC-4.3.1.1	[c]asp-L --> fum + nh4
	Alanine and aspartate metabolism	ASPTA	EC-2.6.1.1	[c]akg + asp-L <==> glu-L + oaa
	Alanine and aspartate metabolism	DAAD	EC-1.4.99.1	[c]ala-D + fad + h2o --> fadh2 + nh4 + pyr

http://sysbio.di.uminho.pt - Record Visualisation - Mozilla Firefox

Visualising record at table reactions

[Update](#) | [Remove](#) | [Print](#) | [Close](#)

Pathway: Alanine and aspartate metabolism
Reaction: ALAR
ECnumber: EC-5.1.1.1
Equation: [c]ala-L <==> ala-D

Reagents: L-Alanine
Products: D-Alanine

Reversible: 1
Boolean rule: CRP AND NOT Lrp
Comment: Activity of alanine racemase in E. coli is due to two distinct gene products. One alanine racemase (Alr) is constitutive; it is encoded by alr. The other DadX is induced by D- or L-alanine and repressed by glucose; it is, and is encoded by the dadX. Alr is less abundant than DadX. Boolean rule de acordo com as posições de ligação do CRP e Lrp (ecocyc)

Encoding Genes:

- alr - b4053
- dadX - b1190

Regulatory Genes:

- crp - b3357
- lrp - b0889

2008-5-15 (14:09)

Done



- HOW CAN WE BUILD THE MODELS IN AN AUTOMATED WAY?
- Ideally, it should be possible to extract all the knowledge necessary to construct biological models from the information obtained during genome sequencing
- However, the knowledge extracted is still very limited...



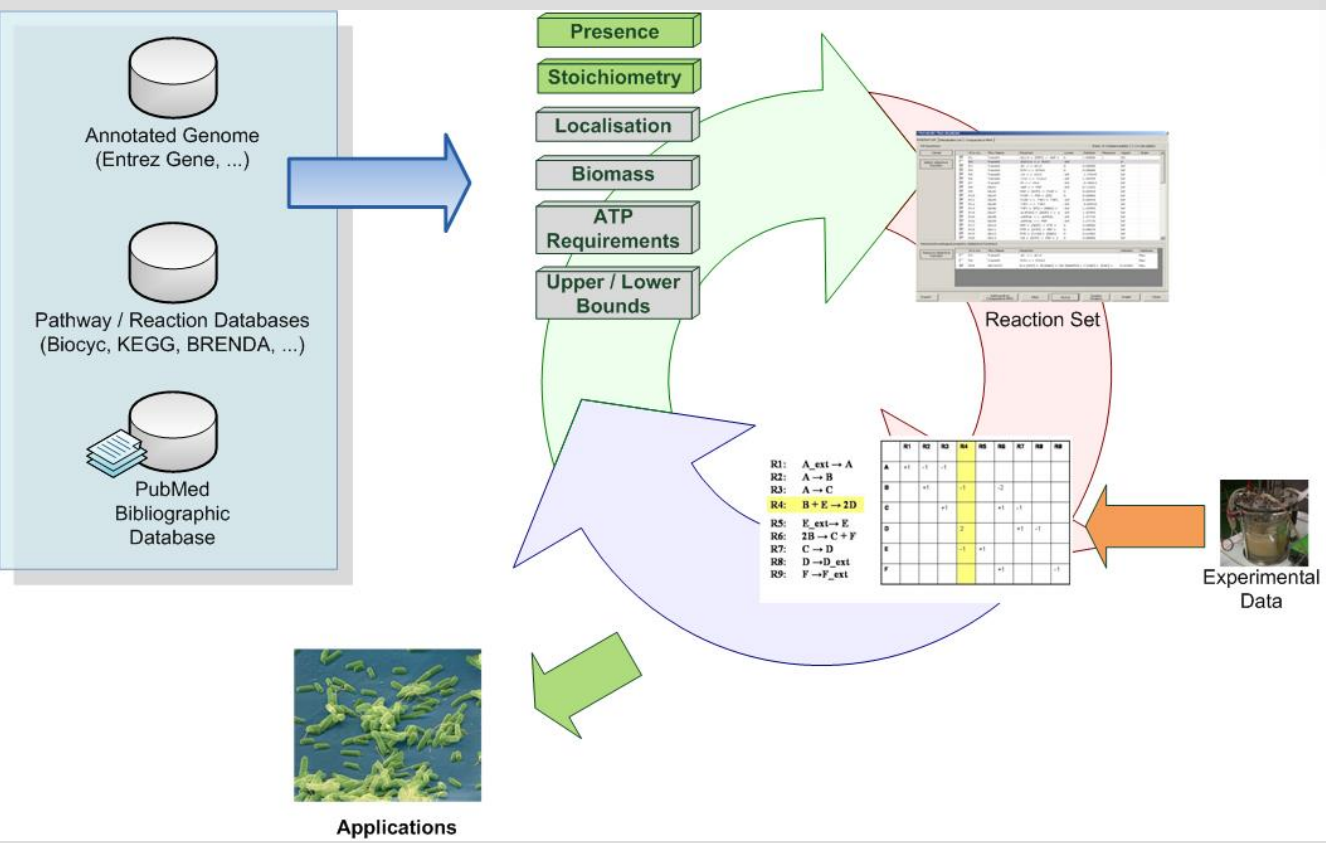
- Presently, the methodology of obtaining stoichiometric models from genome annotation is quite developed

INFERENCE OF BIOLOGICAL NETWORKS

GENOME-SCALE METABOLIC MODELS



WORK IN PROGRESS:
 Reconstruction of Metabolic Networks of:
 -*H. pylori*
 -*K. lactis*
 -*Streptococcus faecalis*



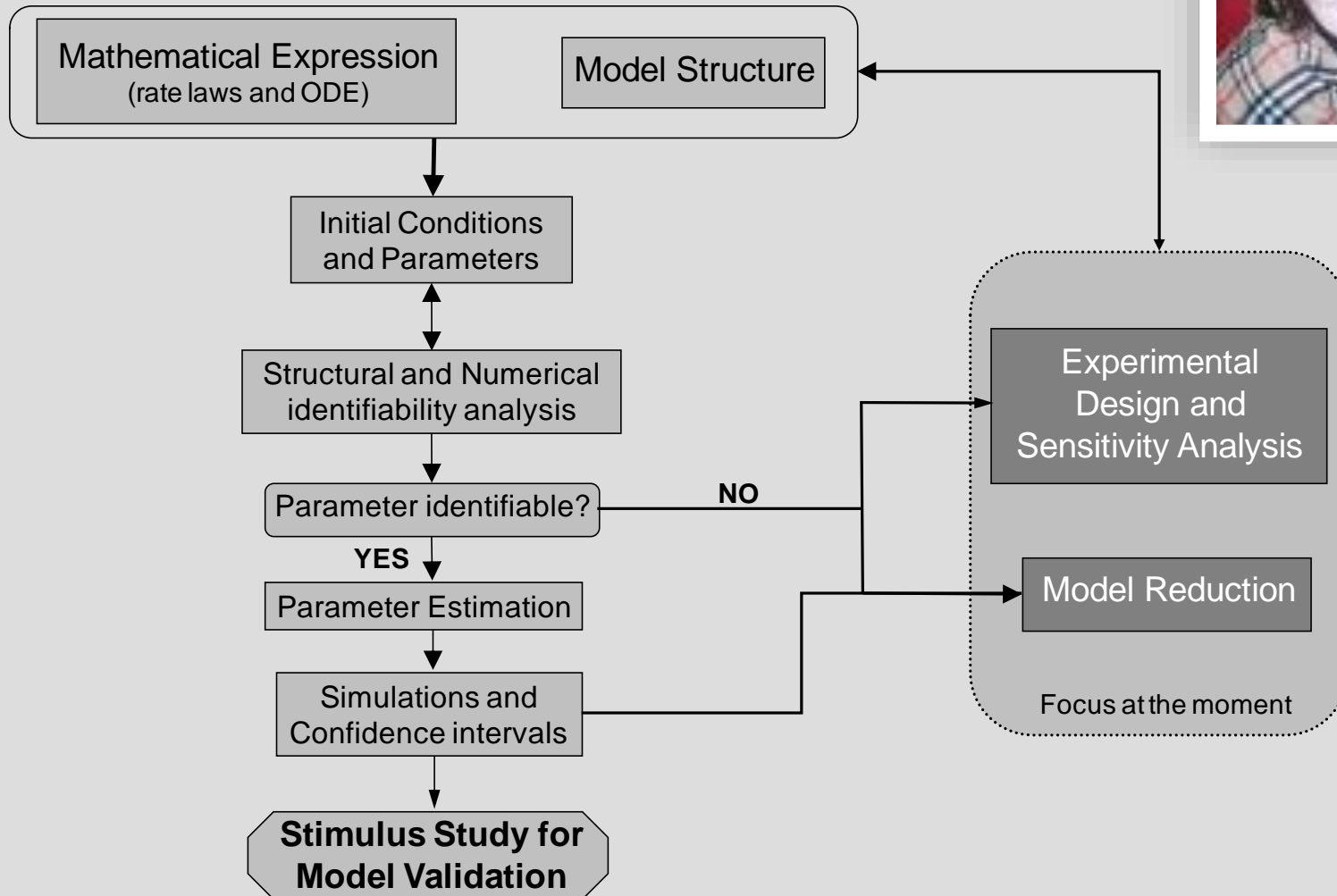
INFERENCE OF BIOLOGICAL NETWORKS

INFERENCE FROM EXPERIMENTAL DATA

- Inference of Biological Networks can also be performed, from **experimental data**.
 - Flux and metabolomic data allow, in principle, to estimate model parameters for kinetic deterministic metabolic models
 - However, the number of experiments and measurements to be performed is very high!
 - Also, the structure of the kinetic equations has to be imposed a priori
 - In this field, optimal experiment design play an important role
- An alternative is to use **Text Mining** tools to automatically search in the literature for biological relations

INFERENCE OF BIOLOGICAL NETWORKS

WORK IN PROGRESS: Experimental Design for inferring model structure and parameters from flux data



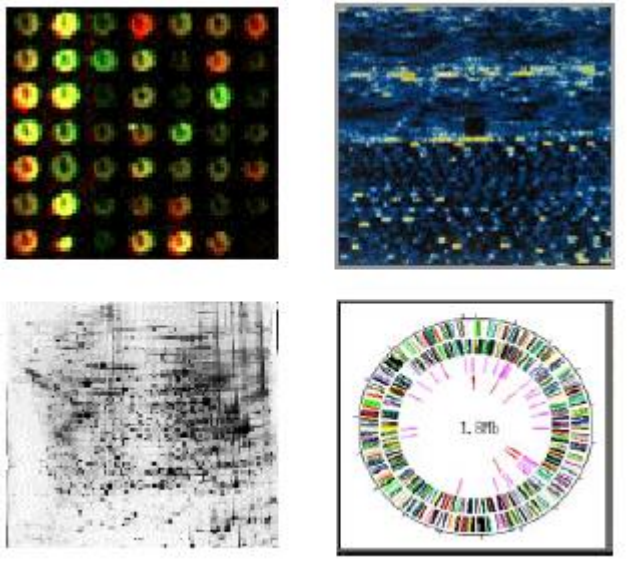
Model Reduction based on dynamic sensitivity analysis



$$\frac{dC_i}{dt} = \sum_j v_{ij}r_j - \mu C_i$$

INTRODUCTION
CELLULAR MODELS
INFERENCE OF BIOLOGICAL
NETWORKS
OPTIMIZATION TOOLS

From *in vivo* to *in silico* and back



High-Throughput Data

Stoichiometric
and/or
Dynamic
Modelling



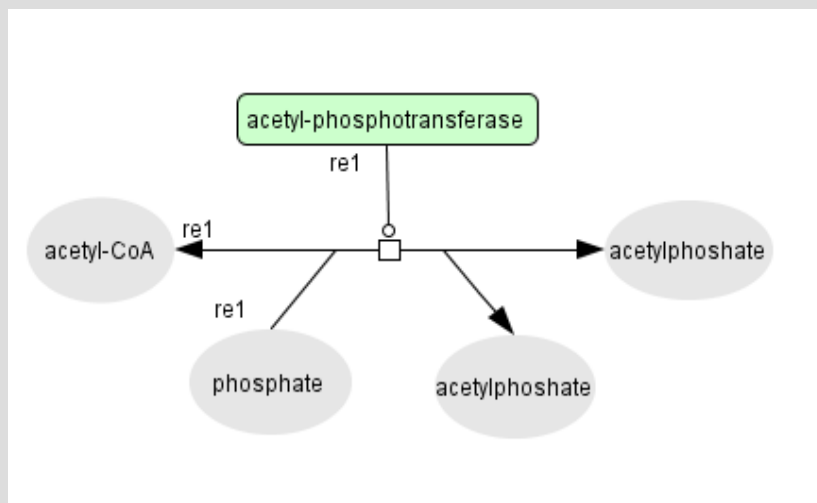
Metabolism
Complexity

A complete kinetic description

Complexity of dynamic modelling



$$\frac{dC_i}{dt} = \sum_j v_{ij} r_j - \mu C_i$$



- $re1_{PTA} = f(\text{metabolites, enzyme, parameters, regulators, ...})$
- Model fluxes and concentrations over time

Drawbacks

- Lots of parameters
- Measured *in vitro* (valid *in vivo*?)
- Nearly impossible to get all parameters at genome scale model

Obstacle for their effective use in optimization and control processes

$$re1_{PTA} = \frac{r_{PTA}^{\max} \left(\frac{1}{K_{i,acetyl-CoA} K_{pta,p}} \right) \left(C_{acetyl-CoA} C_p - \frac{C_{acetyl-P} C_{CoA}}{K_{pta,eq}} \right)}{1 + \frac{C_{acetyl-CoA}}{K_{i,acetyl-CoA}} + \frac{C_p}{K_{i,p}} + \frac{C_{acp}}{K_{i,acp}} + \frac{C_{CoA}}{K_{i,CoA}} + \left(\frac{C_{acetyl-CoA} C_p}{K_{i,acetyl-CoA} K_{pta,p}} \right) + \left(\frac{C_{acp} C_{CoA}}{K_{pta,acetyl-P} K_{i,CoA}} \right)}$$

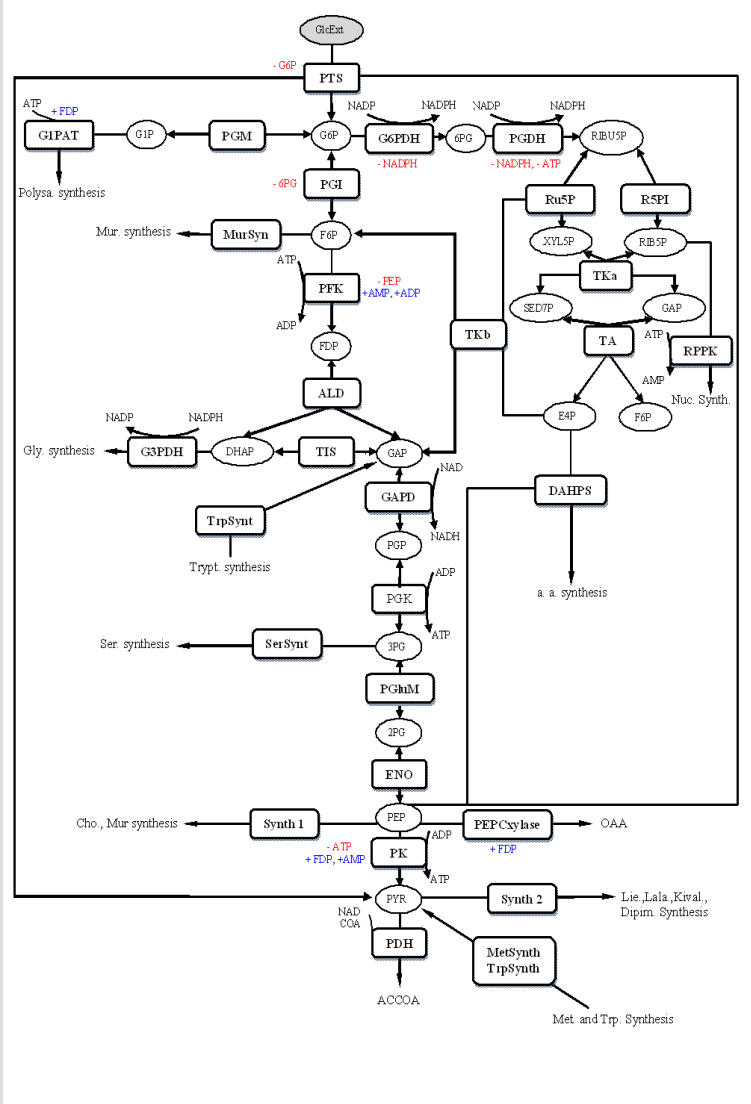
Motivation for model reduction



$$\frac{dC_i}{dt} = \sum_j v_{ij}f_j - \mu C_i$$

Complex *E. coli* dynamic model describing the carbon central metabolism with 116 parameters was used to:

- Identify key parameters that have more impacts on the global systems – **Sensitivity analysis**
- Study a **model reduction** strategy based on univariate analysis of the Euclidean-norm to consider the effect to all metabolites.





$$\frac{dC_i}{dt} = \sum_j v_{ij} r_j - \mu C_i$$

- Nonlinear ODE model

$$\frac{dX_i}{dt} = \sum_j v_{ij} r_j - \mu X_i \quad X_i(t) = f(t, p_j, X_0)$$

- Computing sensitivity analysis

$$S_{i,j}(t) = \frac{X_i(p_j + \Delta p_j) - X_i(p_j - \Delta p_j)}{2\Delta p_j} \times \frac{p_j}{X_i(p)} \approx \frac{\partial \ln X_i(t, p)}{\partial \ln p_j}$$

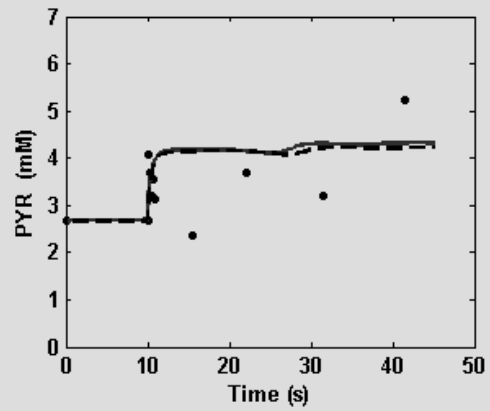
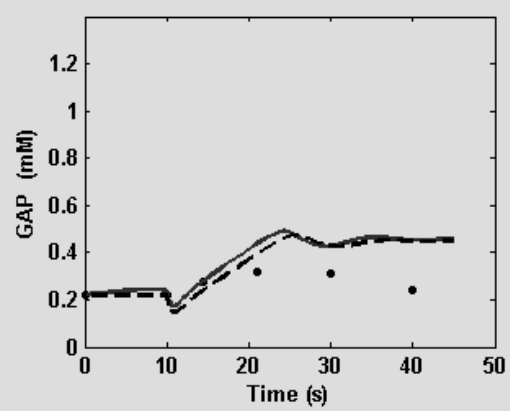
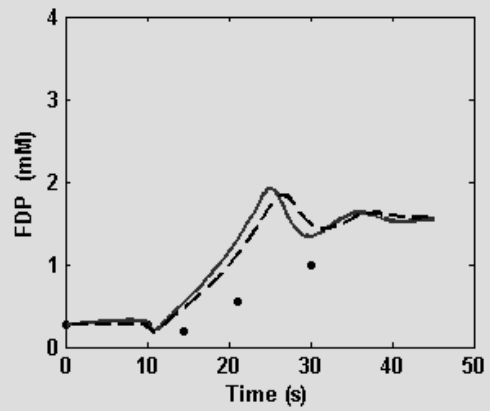
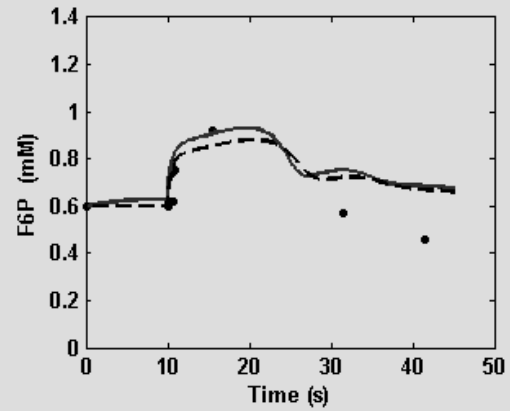
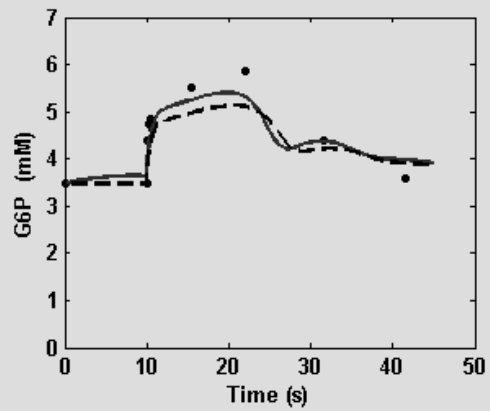
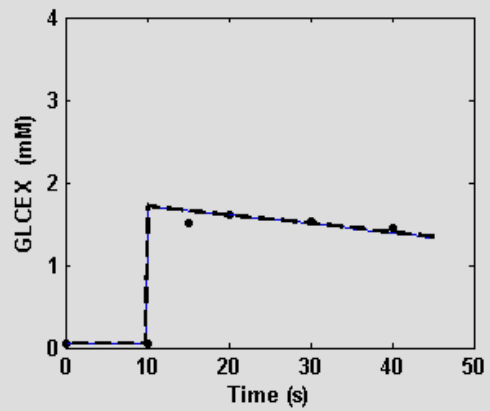
- Dynamic sensitivity analysis based on Euclidean-norm

$$OS_j = \frac{1}{n} \sqrt{\sum_{k=1}^n \sum_{i=1}^p |S_{i,j}(t)|^2}$$

Comparison Original and reduced Model - Metabolite



$$\frac{dC_i}{dt} = \sum_j v_{ij}r_j - \mu C_i$$



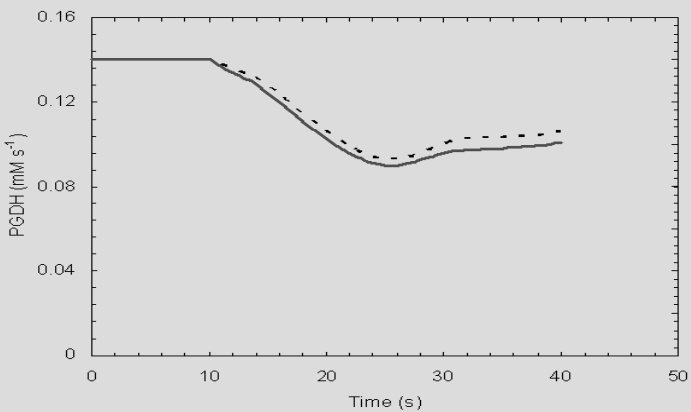
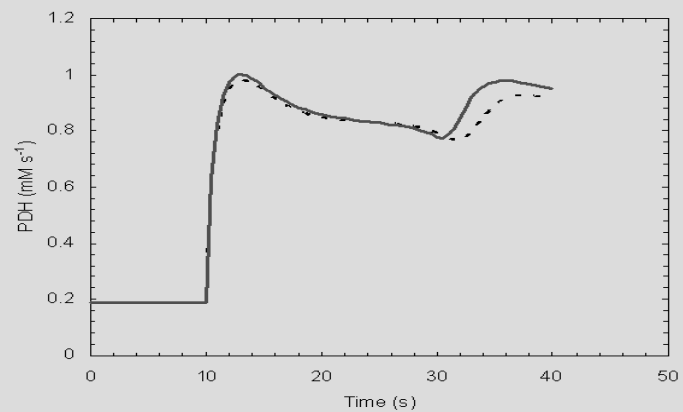
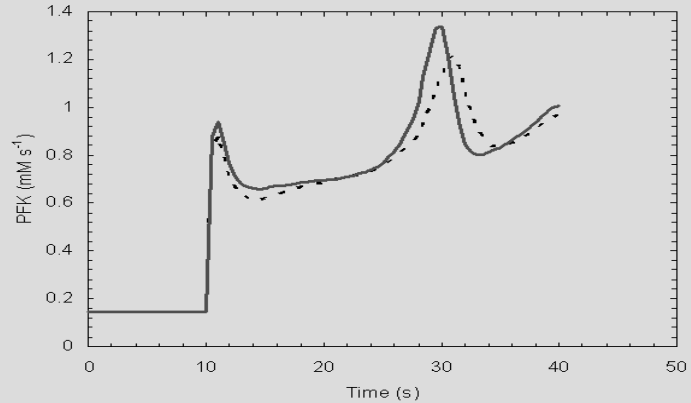
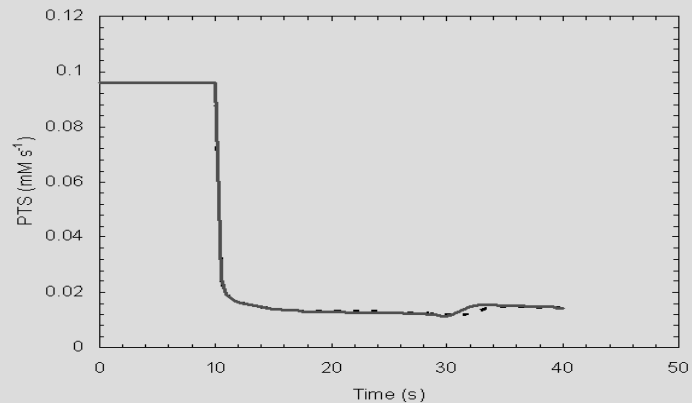
dotted line = original model
 Solid line = reduced model

41 (35.3%) parameters were rejected

Comparison Original and reduced Model - Fluxes



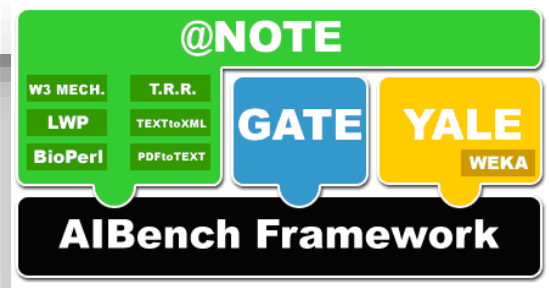
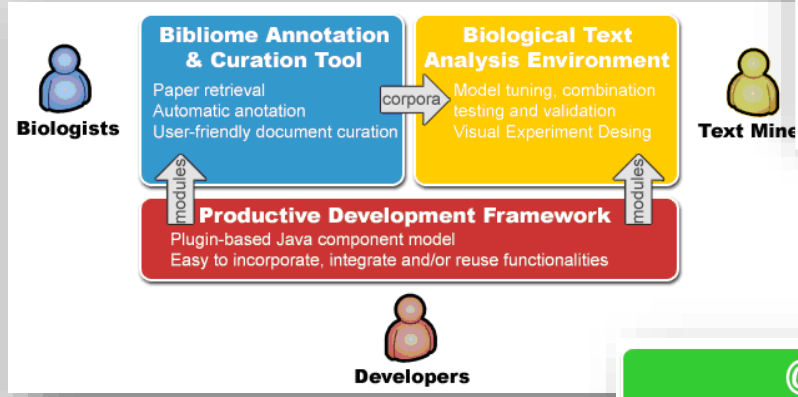
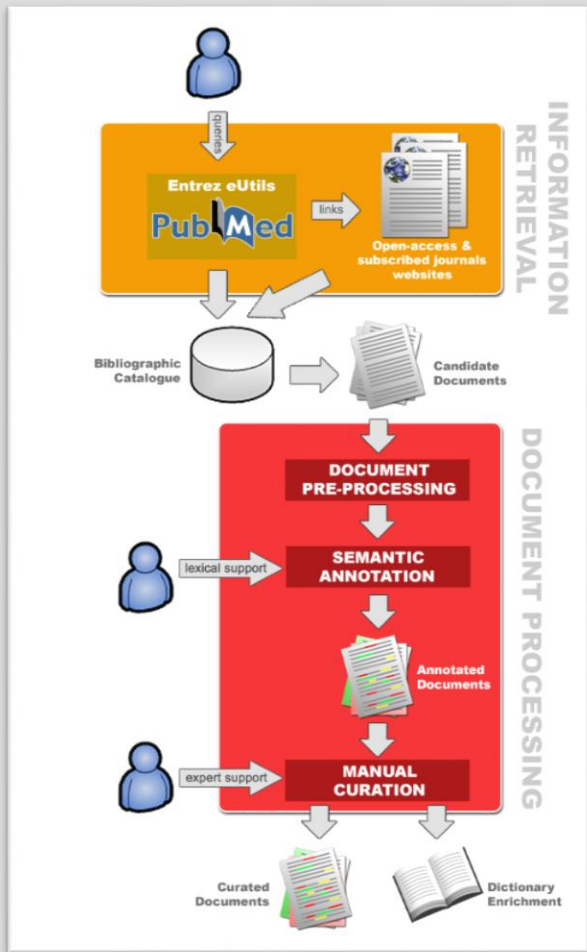
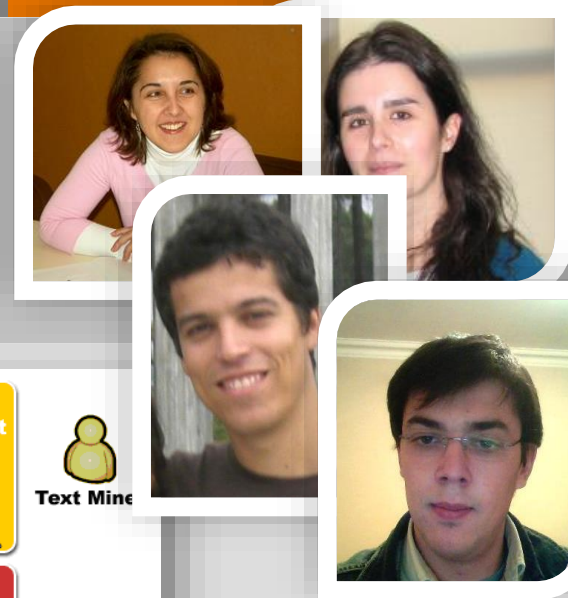
$$\frac{dC_i}{dt} = \sum_j v_{ij}r_j - \mu C_i$$



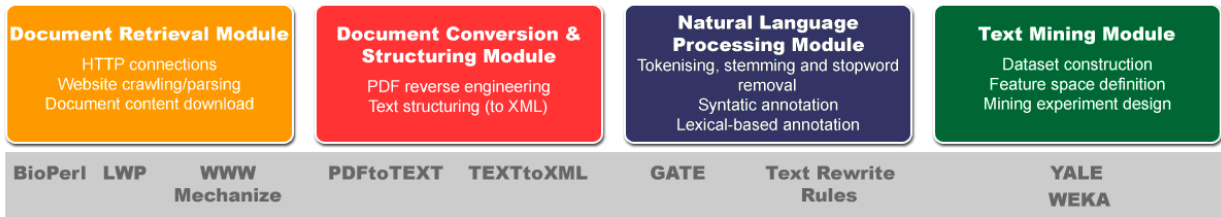
dotted line = original model
Solid line = reduced model

INFERENCE OF BIOLOGICAL NETWORKS

WORK IN PROGRESS: Development of tools for automatically inferring regulatory networks from literature data



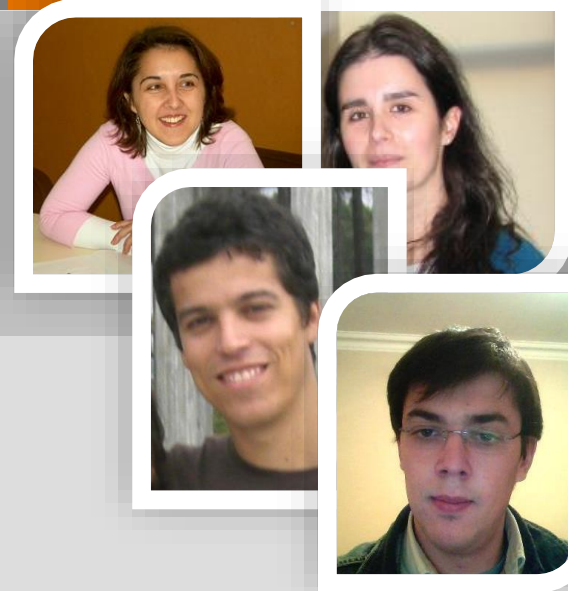
FUNCTIONAL MODULES



TECHNOLOGIES

INFERENCE OF BIOLOGICAL NETWORKS

WORK IN PROGRESS: Development of tools for automatically inferring regulatory networks from literature



@Note (AlBench Framework v2.0b2)

File Database Documents Settings YALE Documentor

Clipboard

Publication: 16961923

Journal: BMC Genomics

Title: Global gene expression during stringent response in *Corynebacterium glutamicum* in presence and absence of the *rel* gene encoding *ppGpp* synthesis

Date: 2006

Volume: 7

Status: MEDLINE

Authors: Brockmann-Gretza O., Kalinow J.

Keywords: Escherichia coli stringent response

Abstract: BACKGROUND: The stringent response is the initial reaction of microorganisms to nutritional stress. During stringent response the small nucleotides, ppGpp act as global regulators and reprogram bacterial transcription. In this work, the genetic network controlled by the stringent response was characterized in the amino acid producing *Corynebacterium glutamicum*. RESULTS: The transcription of 4 C. glutamicum *rel* gene deletion mutant, unable to synthesize ppGpp and to induce the stringent response, was compared with that of the *rel*-proficient parent strain by microarray analysis. A total of 337 genes were found to be transcribed differentially in the *rel*-deficient mutant strain. In a second experiment the stringent response was induced by the addition of D,L-lysine hydrochloride (Lys) in early exponential growth phase. The time point of maximal effect on transcription was determined by real-time RT-PCR using the midline and saline

Selected: 294 Total: 294

Get Publications

Download Non-avail

Results Table Viewer

AlBench

Kasai K, Nishizawa T, Takah...	2006	<input checked="" type="checkbox"/>
Milon P, Tischenko E, Tomsic...	2006	<input checked="" type="checkbox"/>
Brockmann-Gretza O, Kalinow...	2006	<input checked="" type="checkbox"/>
Berney M, Wellenmann HU, E...	2006	<input checked="" type="checkbox"/>
Dabrowska G, PrusA_ska J, ...	2006	<input checked="" type="checkbox"/>
Dabrowska G, PrusA_ska J, ...	2006	<input checked="" type="checkbox"/>
Mouery K, Rader BA, Gaynor ...	2006	<input checked="" type="checkbox"/>
Nakanishi N, ...	2006	<input checked="" type="checkbox"/>

Guanosine tetraphosphate (*ppGpp*) is a key mediator of stringent control, an adaptive response of bacteria to amino acid starvation, and has thus been termed a bacterial alarmone. Previous X-ray crystallographic analysis has provided a structural basis for the transcriptional regulation of RNA polymerase activity by *ppGpp* in the *thermophilic bacterium Thermus thermophilus*. Here we investigated the physiological basis of the stringent response by comparing the changes in intracellular *ppGpp* levels and the rate of RNA synthesis in stringent (*rel*⁻, wild type) and relaxed (*relA*⁻ and *relC*⁻; mutant) strains of *T. thermophilus*. We found that in wild-type *T. thermophilus*, as in other bacteria, *serine hydroxycarbonate*, an amino acid analog of *ppGpp*, and RNAse overproduction elicited a stringent response characterized in part by intracellular accumulation of *ppGpp* and that this response was completely blocked in a *relA* null mutant and partially blocked in a *relC* mutant harbor a *relA* complement. Subsequently, *ppGpp* was isolated from wild-type mutant strains containing the *ppGpp* synthase *relC* gene deletion or tetraacycline inhibitors (*p*₇₀)*ppGpp* synthesis in an *in vitro* system. *ppGpp* acts as a transcriptional inhibitor of RNA polymerase catalyzed 23S/5S rRNA gene transcription but at a concentration much higher than that observed in the cell. On the other hand, changes in *ppGpp* gene promoter activity tightly correlated with changes in the GTP but not ATP concentration. Also, (*p*₇₀)*ppGpp* exerted a potent inhibitory effect on IMP dehydrogenase activity. The present data thus complement the earlier structural analysis by providing physiological evidence that *T. thermophilus* does produce *ppGpp* in response to amino acid starvation in a ribosome-dependent (i.e., *RelA*-dependent) manner. However, it appears that in *T. thermophilus*, rRNA promoter activity is controlled directly by the GTP pool size, which is modulated by *ppGpp* via inhibition of IMP dehydrogenase activity. Thus, unlike the case of *Escherichia coli*, *ppGpp* may not inhibit *T. thermophilus* RNA polymerase activity directly *in vivo*, as recently proposed for *Bacillus subtilis* rRNA transcription (L.

@Note (AlBench Framework v2.0b2)

File Database Documents Settings YALE Documentor

Clipboard

ANoteProject

Catalogue

ResultSet

WorkingSets

PublicationSet (instance 0)

Ner Box List (instance 0)

NerBox (instance 0)

12511489.xml

12949113.xml

Full Text Box (instance 0)

12969985.txt

ANoteProjectConnection

Connection

AlBench

ANoteNerBox (instanc... 12511489.xml)

JOURNAL OF BACTERIOLOGY

Jan. 2003, p. 444-452 Vol. 185, No. 20021-919303:008 00 DOI: 10.1128/JB.185.2.444-452.2003 Copyright © 2003, American Society for Microbiology. All Rights Reserved.

Expression of *spoT* in *Borrelia burgdorferi* during Serum Starvation

Mar. B. Conception and David R. Nelson *

Department of Cell and Molecular Biology, University of Rhode Island, Kingston, Rhode Island 02881 Received 3 June 2002 / Accepted 17 October 2002

Abstract:

Borrelia burgdorferi, the causative agent of Lyme disease, is transmitted by the tick *Ixodes scapularis*. A 29-kb fragment containing a putative *spoT* gene was isolated from *B. burgdorferi* genomic DNA by PCR amplification and cloned into a pBAD24 vector. The cloned gene complemented *Escherichia coli* mutant strain CF1693, which contains deletions of both the *relA* and *spoT* genes. The *spoT* gene product is a sigma factor capable of synthesizing and degrading (*pp*)*ppGpp*, which mediates the stringent response. This-layer chromatography was used to show that *spoT* expression in *E. coli* CF1693 complemented the stringent response. Northern blot analysis revealed that *spoT* expression was induced in *B. burgdorferi* cells starved for serum in RPMI. Expression of *spoT* that starvation conditions was maintained. Further, expression of *spoT* decreased transcriptase PCR (RT-PCR) was used to detect *spoT* mRNA from 106 cells starved for serum for 15 min. Northern blot analysis suggests that *spoT* transcript was 900 nucleotides in length.

RT-PCR amplification of the transcript using several sets of primers confirmed this finding. The 2,001-bp *spoT* open reading frame was able to complement *E. coli* CF1693. The data to serum starvation and during incubation in tick saline.

bio_classes

- technique (21)
- western blot (1)
- thin layer chromatography (1)
- cloning (7)
- centrifugation (5)
- agarose gel (5)
- electrophoresis (4)
- gel electrophoresis (3)
- southern blot (4)
- southern analysis (2)
- hybridization (4)
- electroporation (3)
- gene
- compound
- pathway
- regulatory_gene
- state
- metabolic_gene
- organism
- enzyme
- verb
- transcription_factor
- rna
- rRNA (3)
- irRNA (3)
- mRNA (8)
- protein
- Sall (7)
- TnaA (1)
- PstI (7)
- SacI (1)
- rRNA polymerase, sigma

Structure

- SECTIONS
- ACKNOWLEDGMENT
- REFERENCES
- DISCUSSION
- RESULTS
- MATERIALS AND METHODS
- AUTHORS
- ABSTRACT
- JOURNAL
- TITLE

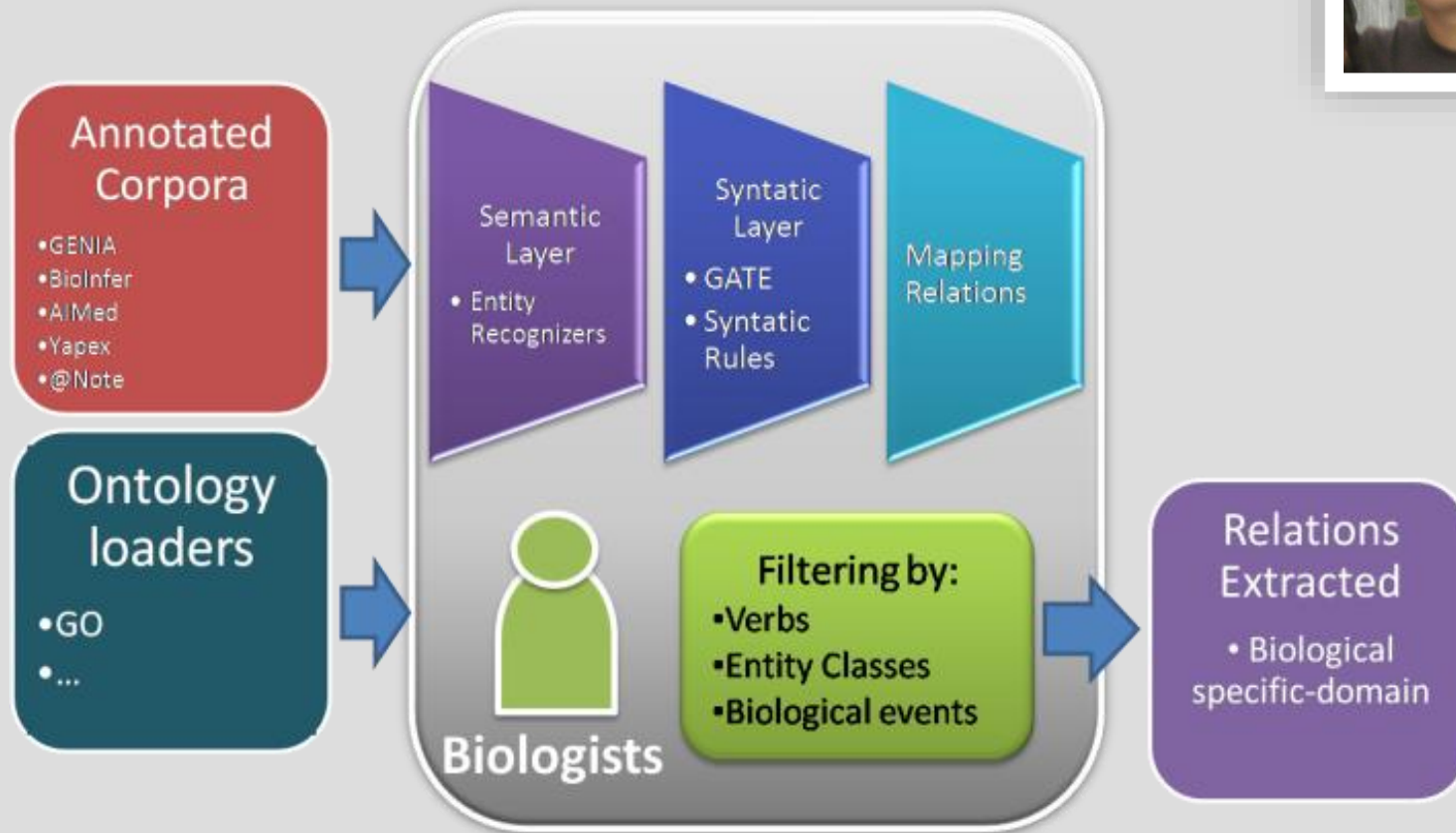
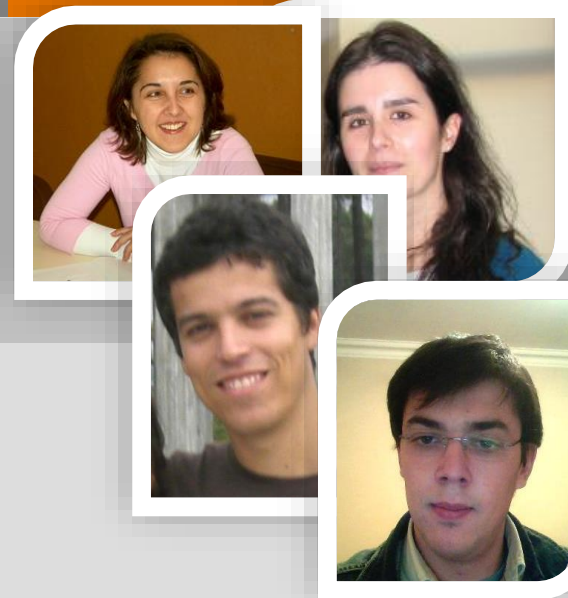
Document View

Legend:

- technique
- compound
- state
- reaction
- pathway
- metabolic_gene
- gene
- regulatory_gene
- organism
- transcription_factor
- protein

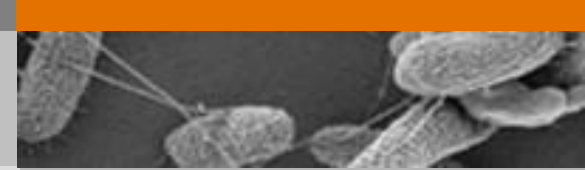
INFERENCE OF BIOLOGICAL NETWORKS

WORK IN PROGRESS: Development of tools for automatically inferring regulatory networks from literature data

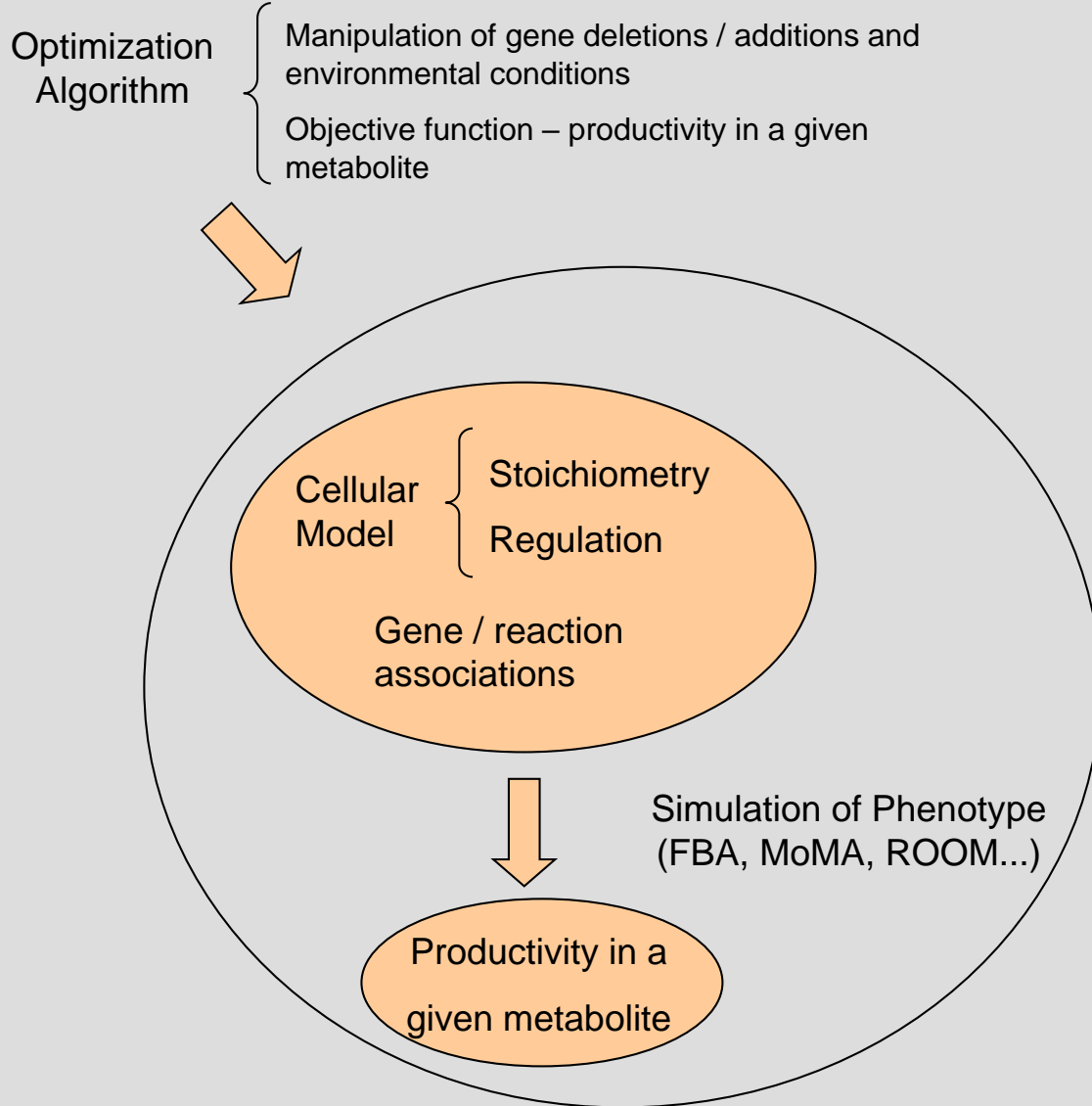


OPTIMIZATION TOOLS

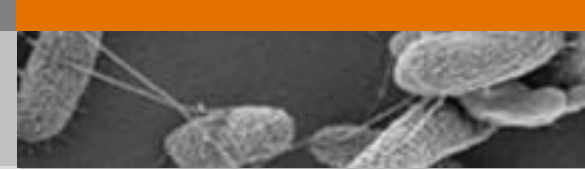
OPTIMIZATION PROBLEM



INTRODUCTION
CELLULAR MODELS
INFERENCE OF BIOLOGICAL
NETWORKS
OPTIMIZATION TOOLS

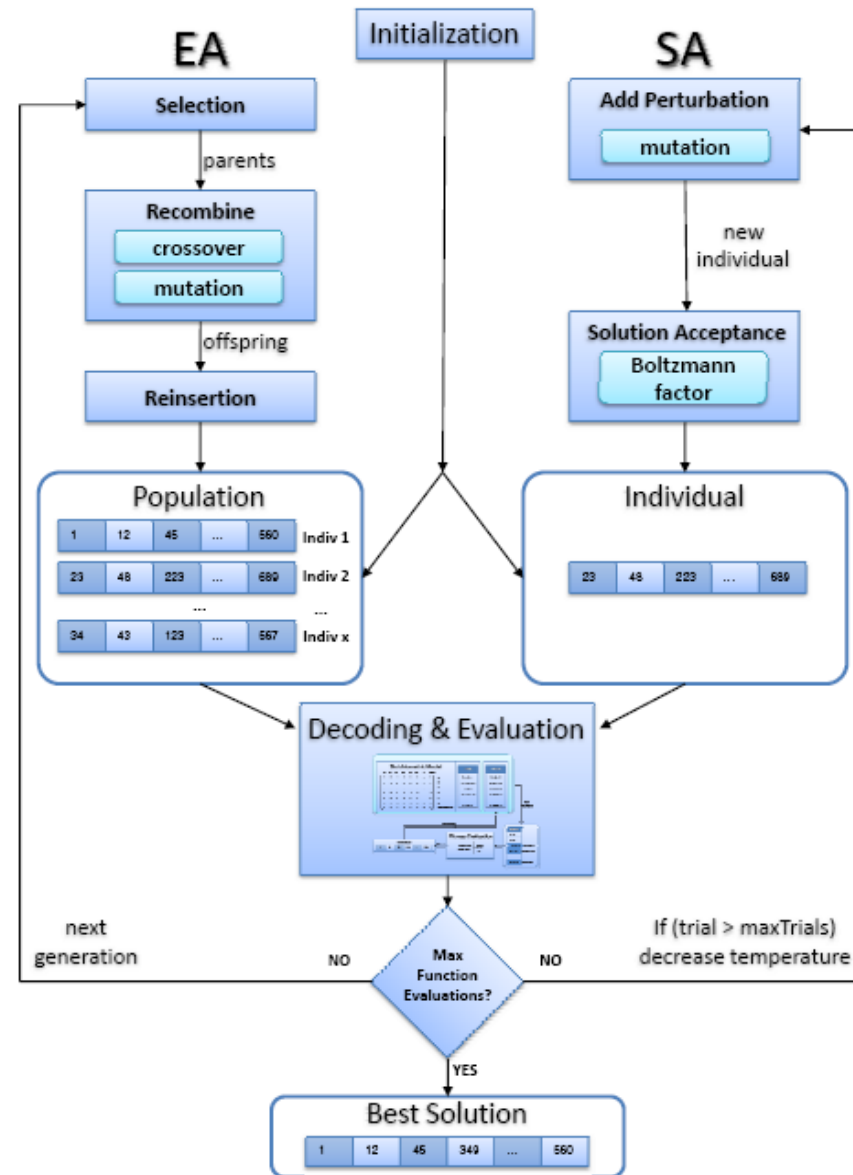


FBA: Flux Balance Analysis
ROOM: Regulatory On/Off
Minimization
MoMA: Minimization of Metabolic
Adjustment

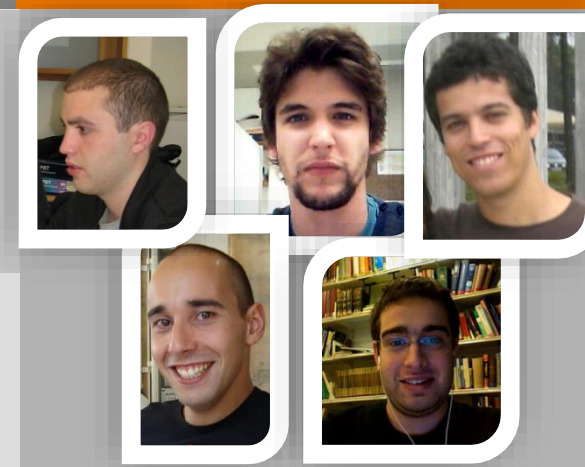


The only reported algorithms are:

- OptKnock (Burgard *et al.*, Biotech Bioeng 2003)
 - Based on MILP
 - Only applicable to relatively small stoichiometric models
- OptGene (Patil *et al.*, BMC Bioinf 2005) & **OptFlux** (Rocha *et al.*, BMC Bioinf 2008)
 - Evolutionary Algorithms
 - Applicable to different types of (large-scale) models
- Additional algorithms being applied:
 - Local Search
 - Simulated Annealing



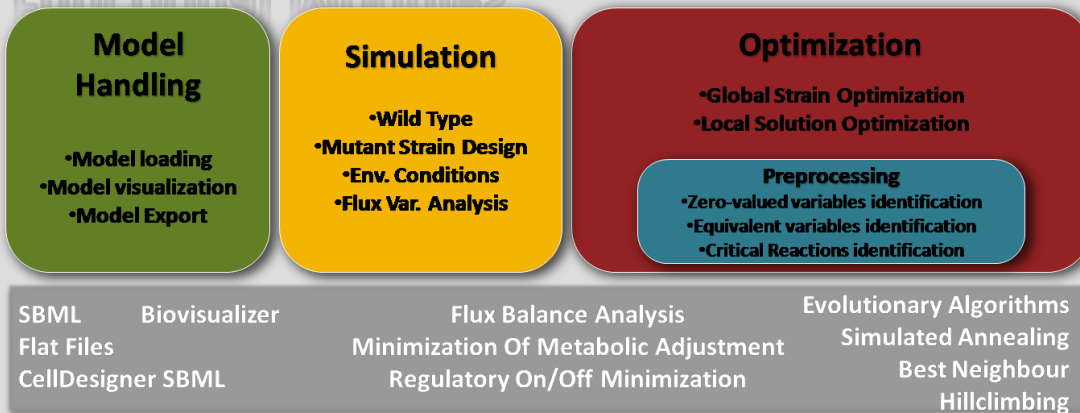
WORK IN PROGRESS: OptFlux – a software for the Optimization of microbial strains



- **OptFlux** is an open-source, user-friendly and modular software aimed at being the reference computational tool for metabolic engineering applications. It allows the use of stoichiometric metabolic models for simulation and optimization purposes.

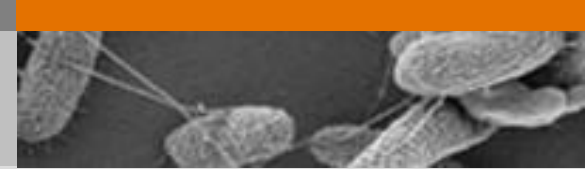
www.optflux.org

Functional Modules



Methods & Technologies

OPTIMIZATION TOOLS
OptFlux - Conceptual overview



Strain
Optimization

- ✓ Evolutionary algorithms
- ✓ Simulated Annealing
- ✓ Local optimization

Objective
function:
Maximize the
production of a
given metabolite

Phenotype
Simulation

- ✓ FBA, MoMA, ROOM
- ✓ Boolean net simulation
- ✓ ...

Determine the
set of fluxes (in
steady state)

Environmental
Conditions

Genetic conditions
(e.g. knockouts)

Override Model

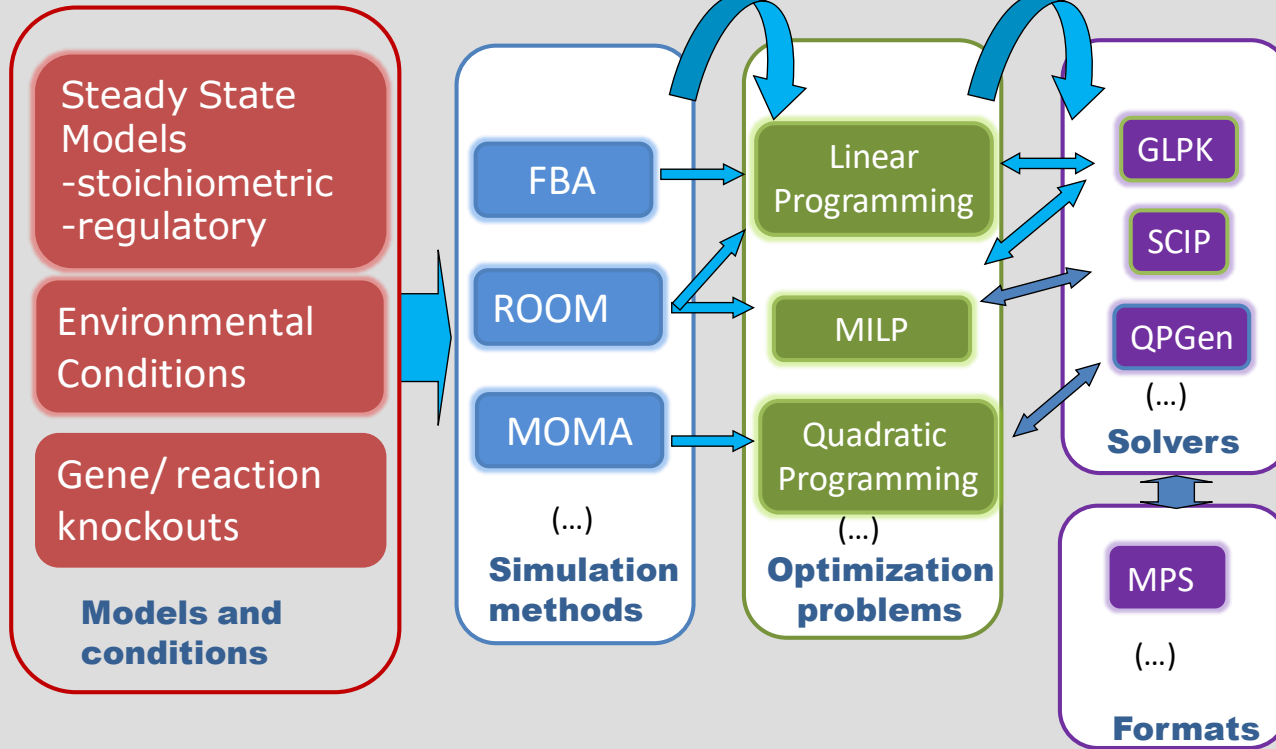
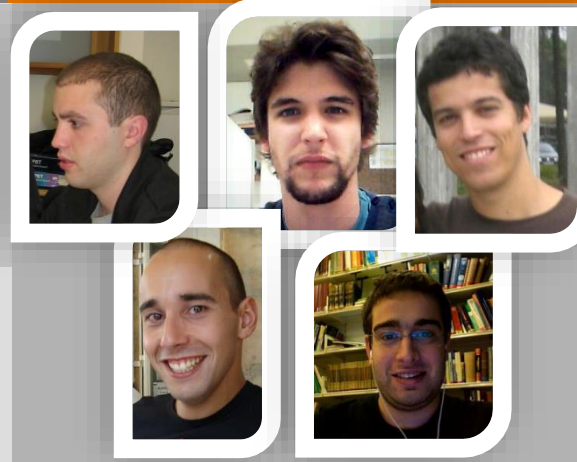
Model

- ✓ Stoichiometric
- ✓ Regulatory
- ✓ ...

Original Model

OPTIMIZATION TOOLS

WORK IN PROGRESS: OptFlux – a software for the Optimization of microbial strains



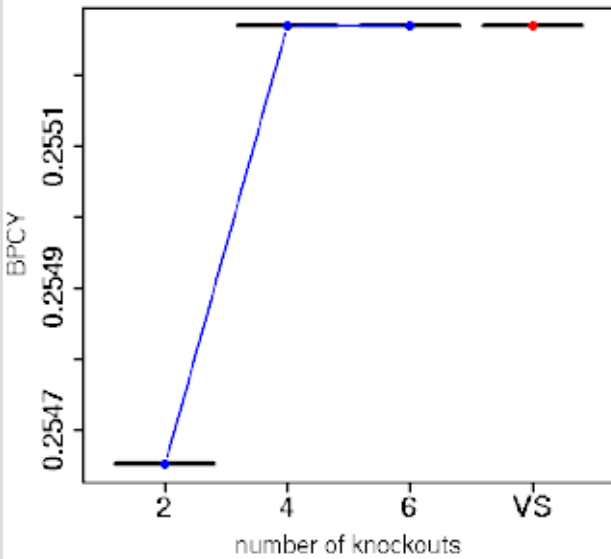
www.optflux.org

OPTIMIZATION TOOLS

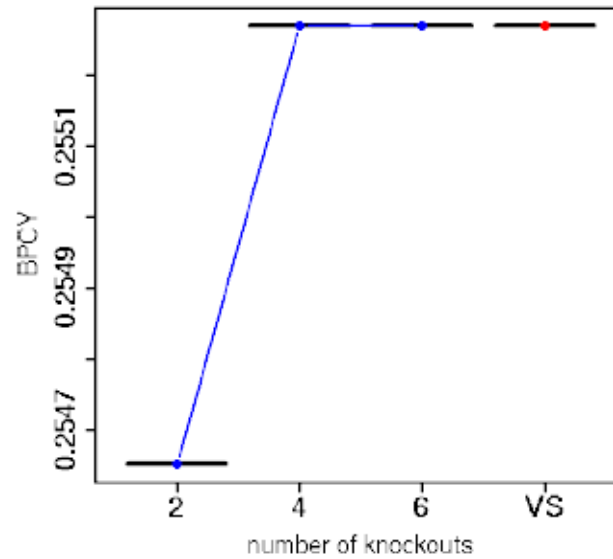
OPTIMIZATION PROBLEM



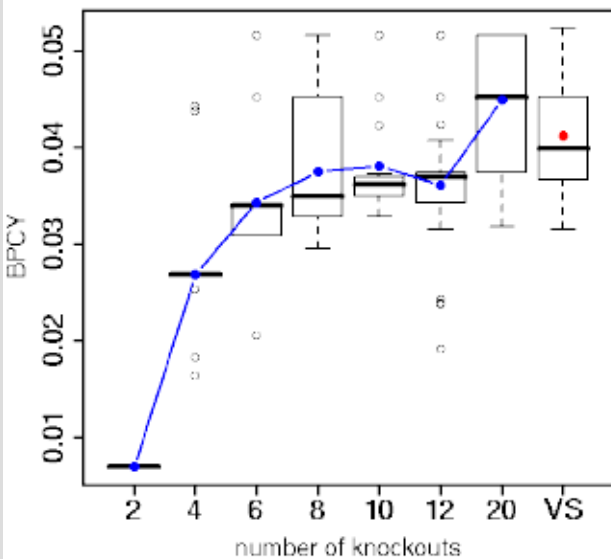
Anaerobic *E.coli*, lactate, EA



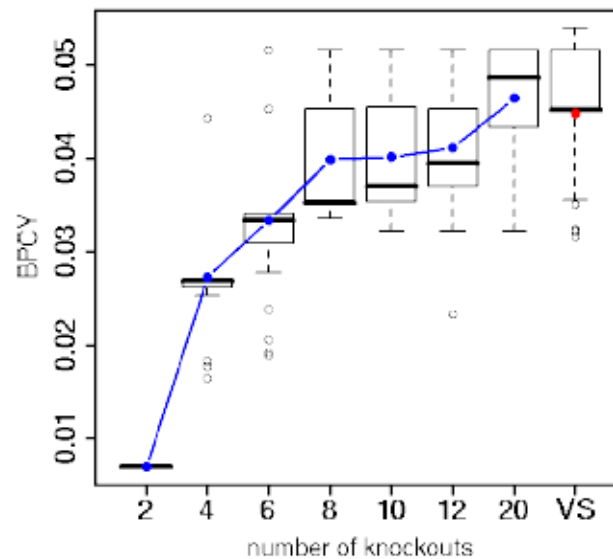
Anaerobic *E.coli*, lactate, SA



S.cerevisiae, succinate, EA



S.cerevisiae, succinate, SA



BPCY: Biomass-Product Coupled Yield
VS: Variable size

Acknowledgments - BioPSE group



IBB – Institute for Biotechnology and Bioengineering
Centre of Biological Engineering
Universidade do Minho

Faculty:
Eugénio Ferreira
Isabel Rocha
Miguel Rocha
Rui Mendes
Ana Veloso
Anália Lourenço

Post-Docs:
Zita Soons

PhD students:
Sónia Carneiro
Rafael Costa
José Pedro Pinto
Daniel Machado
Oscar Dias
Carla Portela
Daniela Correia

Other researchers:
Paulo Maia
Pedro Evangelista
Rafael Carreira
Simão Soares
José P. Faria
Paulo Vilaça
Rui Pereira

<http://biopseg.deb.uminho.pt>



Universidade do Minho

INTRODUCTION COLLABORATIONS

