

Biologia de Sistemas para o desenvolvimento de fábricas celulares microbianas

Eugénio Campos Ferreira

Centre of Biological Engineering
University of Minho
Braga
Portugal





- Introduction to Systems Biology
- Metabolic Models
- Simulation
- Systems Metabolic Engineering (in silico)
- Bioprocess Engineering
- Conclusions

- **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
- **Metabolic Engineering** can gain major benefits from the systems biology approach



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.



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

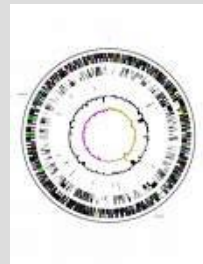
- Cellular Models for Metabolic Engineering: gene networks
- Inference of Biological Networks
 - From Genome-scale metabolic models
 - From experimental data
 - From literature data mining
- *In Silico* Metabolic Engineering Platforms: Optimization of Microbial strains – **OptFlux** tool



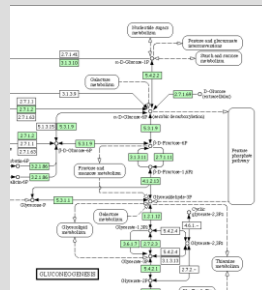
- Metabolic engineering - introduction of directed genetic modifications leading to desirable metabolic phenotypes



Cell factory



Genome

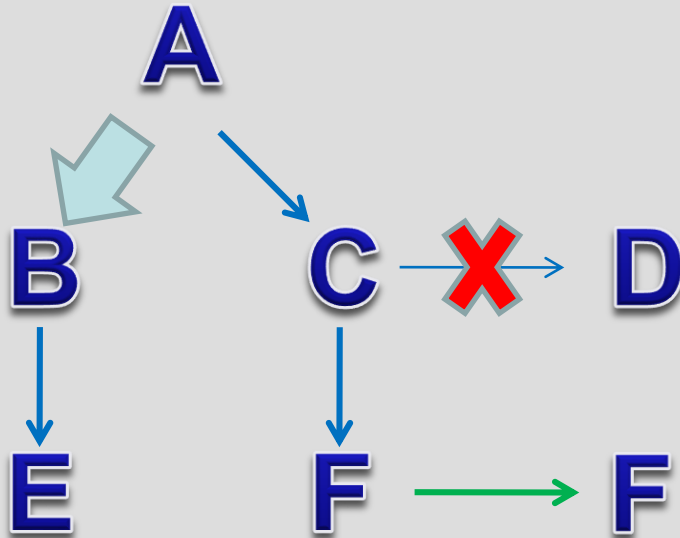


Metabolism / Phenotype



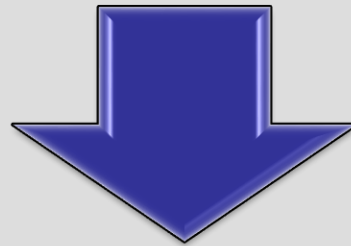
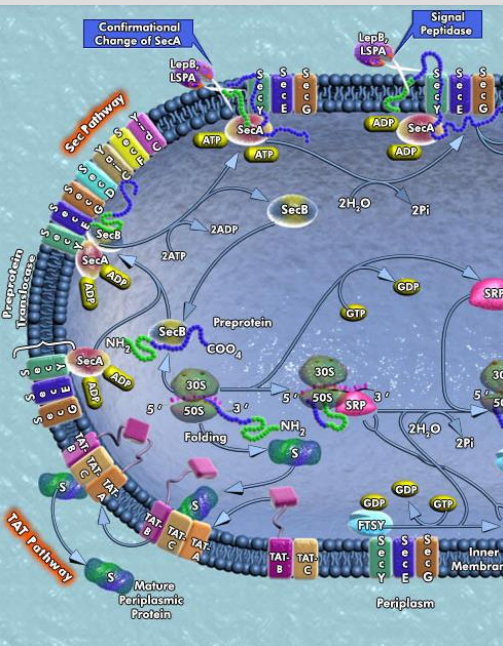
Metabolic Engineering Strategies:

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



INTRODUCTION IN SILICO METABOLIC ENGINEERING

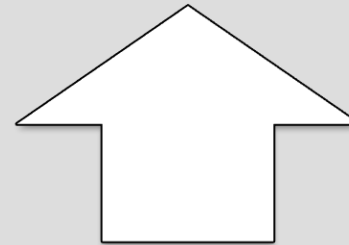
INTRODUCTION
METABOLIC MODELS
SIMULATION
IN SILICO METABOLIC ENGINEERING
PROCESS ENGINEERING
CONCLUSIONS



Bioprocess
Engineering



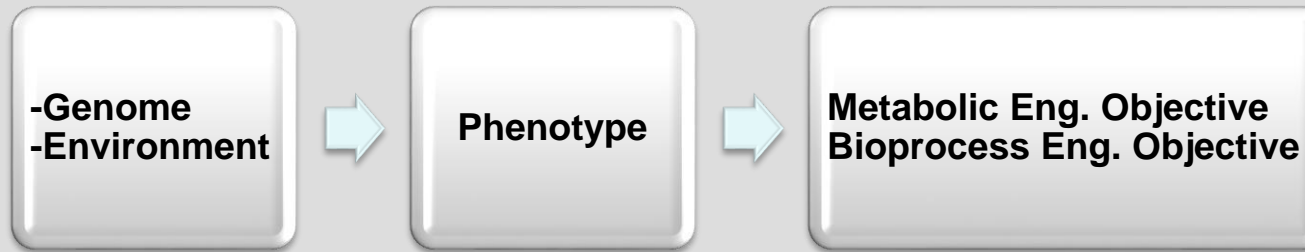
Metabolic
Engineering





- A view of the Metabolic Engineering / Bioprocess Engineering Problem

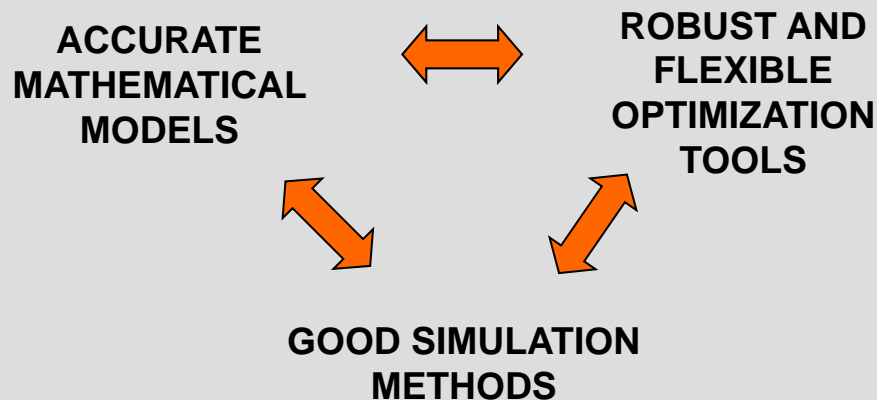
Simulation



**Optimization,
Control strategies**



- 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:



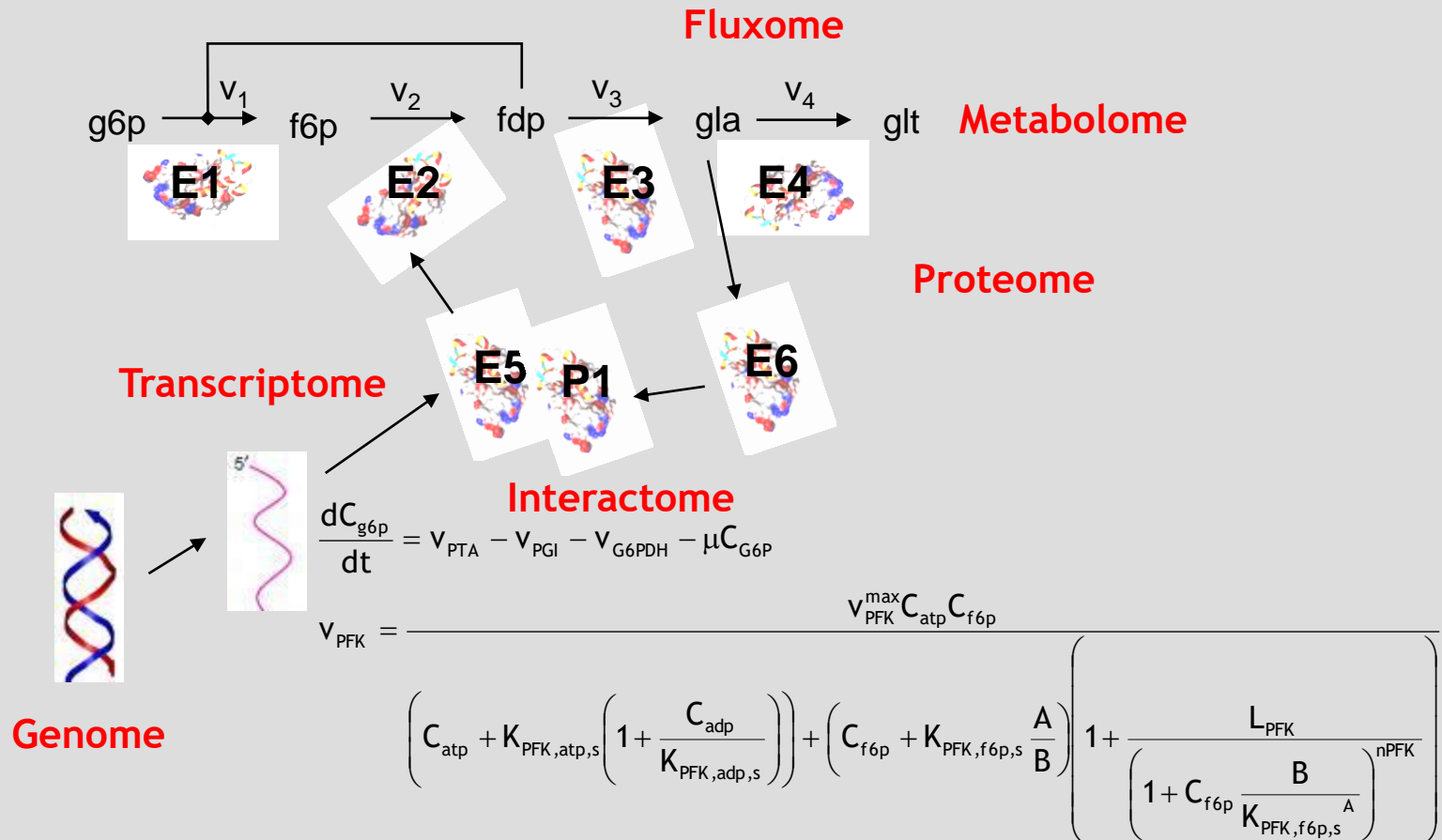
METABOLIC MODELS LEVELS OF INFORMATION



INTRODUCTION
METABOLIC MODELS
SIMULATION
IN SILICO METABOLIC ENGINEERING
PROCESS ENGINEERING
CONCLUSIONS

Models should comprise different levels of information:

- Reactions stoichiometry
- Reactions kinetics
- Regulatory information



METABOLIC MODELS

EXISTING MODEL TYPES



INTRODUCTION
METABOLIC MODELS
SIMULATION
IN SILICO METABOLIC ENGINEERING
PROCESS ENGINEERING
CONCLUSIONS

- Different model types are at different development stages...

	Stoichiometric	Regulatory	Kinetic
Genome-scale	**	***	*
Simul. Accuracy	*	**	***
Gene Deletions	***	***	***
Gene Over/ under express.	-	-	***



Framework for calculation of intracellular metabolic (net) fluxes is based on:

- **Mass balance** over intracellular metabolites

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

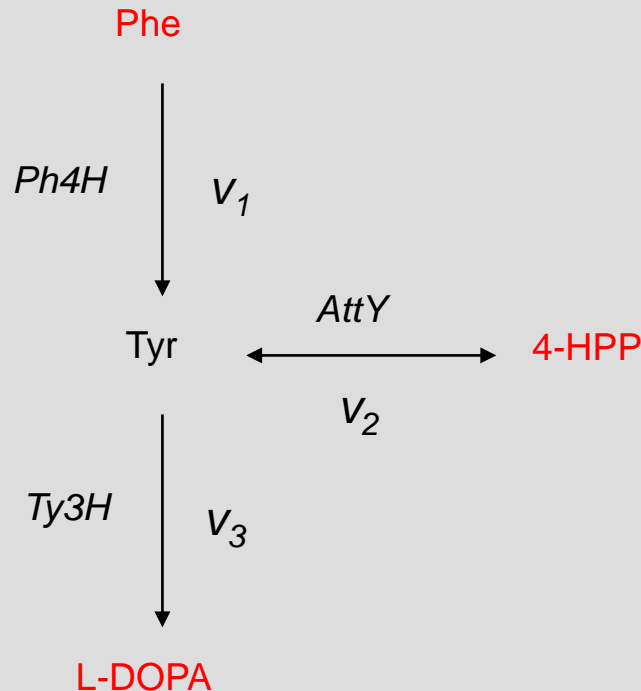
- Assumption of (pseudo) **steady state**

$$v_1 - v_2 - v_3 = 0$$

- For all metabolites:

$$S \cdot v = 0$$

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



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

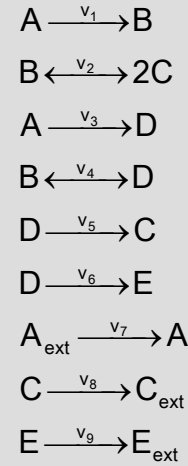
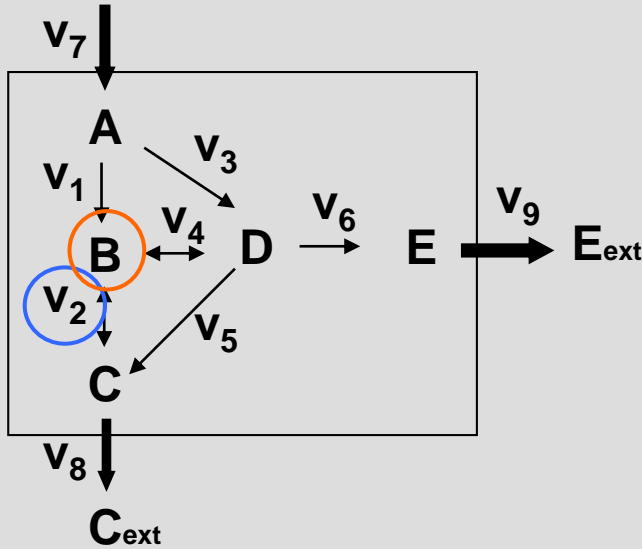
The result is a Linear Equations system described by stoichiometric matrix *S*.

METABOLIC MODELS

STOICHIOMETRIC MODELS



For an identified reaction set:



$$\begin{aligned}
 \text{A: } &-v_1 - v_3 + v_7 = 0 \\
 \text{B: } &v_1 - v_2 - v_4 = 0 \\
 \text{C: } &2v_2 + v_5 - v_8 = 0 \\
 \text{D: } &v_3 + v_4 - v_5 - v_6 = 0 \\
 \text{E: } &v_6 - v_9 = 0
 \end{aligned}$$

$$\begin{aligned}
 &0 \leq v_1 \leq +\infty \\
 &-\infty \leq v_2 \leq +\infty \\
 &0 \leq v_3 \leq +\infty \\
 &-\infty \leq v_4 \leq +\infty \\
 &0 \leq v_5 \leq +\infty \\
 &0 \leq v_6 \leq +\infty \\
 &0 \leq v_7 \leq a \\
 &0 \leq v_8 \leq +\infty \\
 &0 \leq v_9 \leq +\infty
 \end{aligned}$$

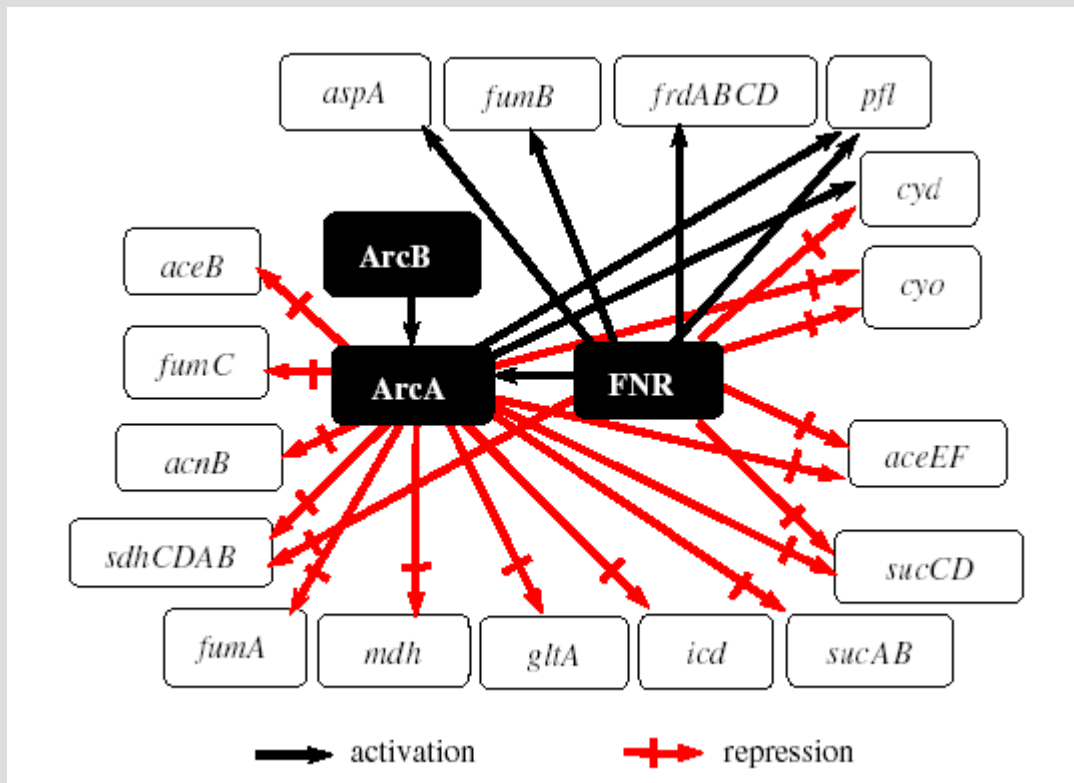
$$\begin{bmatrix}
 -1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\
 1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 2 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\
 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1
 \end{bmatrix}
 \begin{bmatrix}
 v_1 \\
 v_2 \\
 v_3 \\
 v_4 \\
 v_5 \\
 v_6 \\
 v_7 \\
 v_8 \\
 v_9
 \end{bmatrix}
 =
 \begin{bmatrix}
 0 \\
 0 \\
 0 \\
 0 \\
 0
 \end{bmatrix}$$

$$S \cdot v = 0$$

METABOLIC MODELS

GENE REGULATORY NETWORKS

- 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
- In the figure, ArcA and FNR are transcription factors



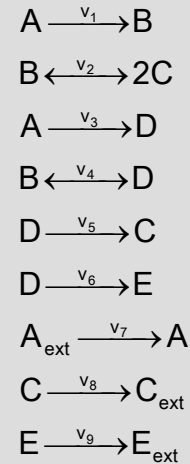
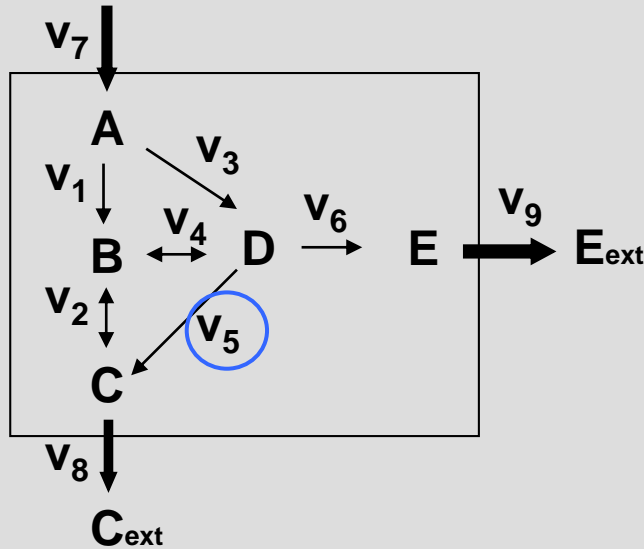
METABOLIC MODELS

STOICHIOMETRIC-REGULATORY MODELS



INTRODUCTION
 METABOLIC MODELS
 SIMULATION
 IN SILICO METABOLIC ENGINEERING
 PROCESS ENGINEERING
 CONCLUSIONS

For an identified reaction set:



If for condition Z, reaction 5 does not occur

$$\begin{aligned}
 \text{A: } &-v_1 - v_3 + v_7 = 0 \\
 \text{B: } &v_1 - v_2 - v_4 = 0 \\
 \text{C: } &2v_2 + v_5 - v_8 = 0 \\
 \text{D: } &v_3 + v_4 - v_5 - v_6 = 0 \\
 \text{E: } &v_6 - v_9 = 0
 \end{aligned}$$

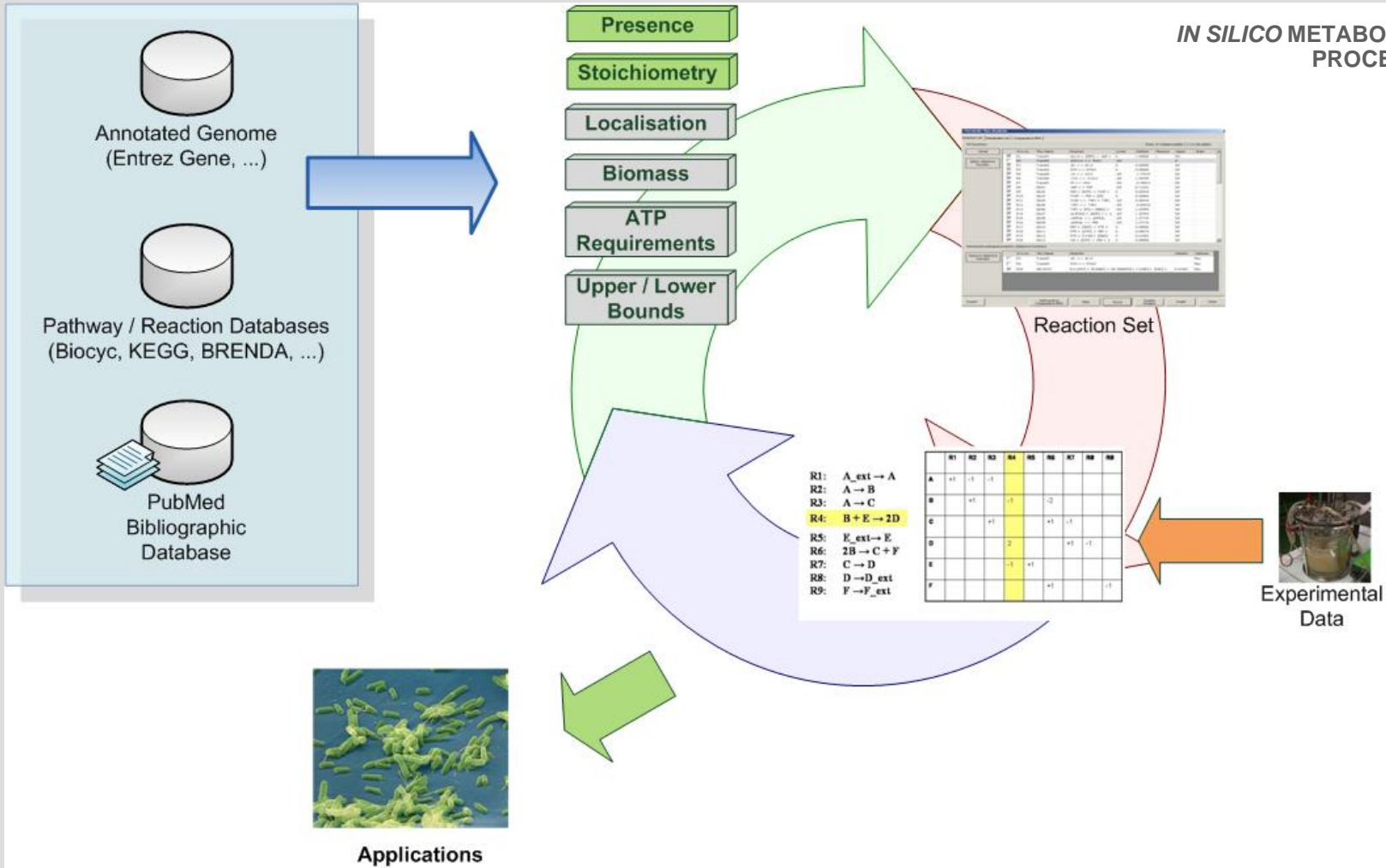
$$\begin{aligned}
 0 &\leq v_1 \leq +\infty \\
 -\infty &\leq v_2 \leq +\infty \\
 0 &\leq v_3 \leq +\infty \\
 -\infty &\leq v_4 \leq +\infty \\
 \mathbf{0 \leq v_5 \leq 0} \\
 0 &\leq v_6 \leq +\infty \\
 0 &\leq v_7 \leq a \\
 0 &\leq v_8 \leq +\infty \\
 0 &\leq v_9 \leq +\infty
 \end{aligned}$$

$$\begin{bmatrix} -1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \\ v_9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

METABOLIC MODELS RECONSTRUCTION



INTRODUCTION
METABOLIC MODELS
SIMULATION
IN SILICO METABOLIC ENGINEERING
PROCESS ENGINEERING
CONCLUSIONS



METABOLIC MODELS GENOME-SCALE MODELS



INTRODUCTION
METABOLIC MODELS
SIMULATION
IN SILICO METABOLIC ENGINEERING
PROCESS ENGINEERING
CONCLUSIONS

Microorganism	On-line availability
<i>Haemophilus influenzae</i>	http://gcrq.ucsd.edu/organisms/hinfluenzae.html
<i>Escherichia coli</i>	http://gcrq.ucsd.edu/organisms/ecoli_reactions.html
<i>Helicobacter pylori</i>	http://gcrq.ucsd.edu/organisms/hpylori.html
<i>Saccharomyces cerevisiae</i>	http://www.cpb.dtu.dk/models/yeastmodel.html http://systemsbiology.ucsd.edu/organisms/yeast.html
<i>Aspergillus niger</i>	http://blackwellpublishing.com/products/journals/suppmat/EJB/EJB3798/EJB3798sm.htm
<i>Plasmodium falciparum</i>	http://plasmocyc.stanford.edu
...	

Applications of Genome-scale metabolic models:

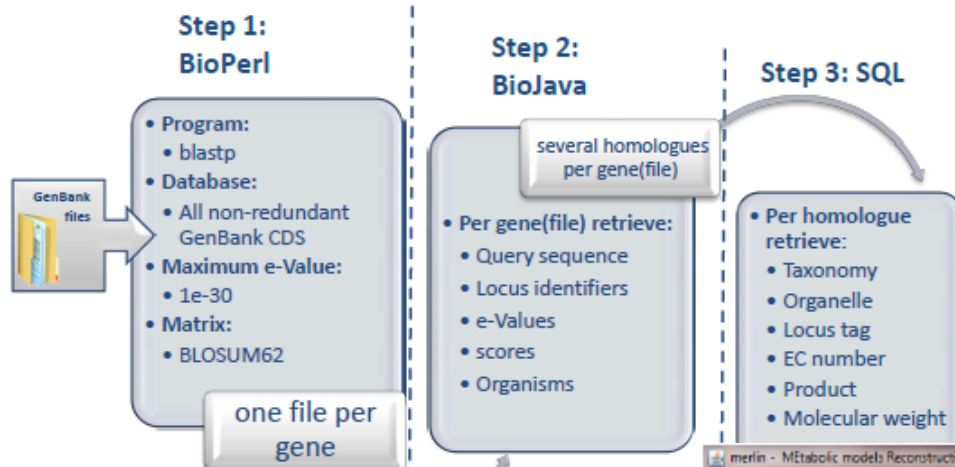
- Design of industrial strains for industrial biotechnology
- Growth medium design
- Discovery of new gene functions
- Better understanding of microbial physiology
- Identification of potential drug targets in pathogens
- ...

METABOLIC MODELS

GENOME ANNOTATION IN RECONSTRUCTION



- Annotation available usually after sequencing
- However, it can be old or incomplete!



merlin:
Metabolic Models Reconstruction
using Genome-Scale Information

merlin - Metabolic Models Reconstruction using genome-scale Information (powered by AIBench Framework v X.x)

Clipboard

Genes	Name	Chromosome	Product	Score (%)	EC Numbers
YGL028C	SCW11	VII	beta-glucosidase 5	4.95 %	3.2.1.-
YGL029W	CGR1	VII	Cgr1p	25 %	
YGL030W	RPL30	VII	60S ribosomal protein L30	25 %	
YGL031C	RPL24A	VII	60S ribosomal protein L24	21.78 %	
YGL032C	AGA2	VII	alpha-2-macroglobulin subunit	50 %	
YGL033W	HOP2	VII	Hop2p	27.27 %	
YGL035C	MG1	VII	Mg1p	30 %	
YGL036W		VII	Putative protein of unknown fun...	16.67 %	
YGL037C	PNC1	VII	Nicotinamidase that converts ni...	10 %	3.5.1.19
YGL038C	OCH1	VII	alpha-1,6-mannosyltransferase	7.46 %	2.4.1.232
YGL039W		VII	potential oxidoreductase	2.97 %	1.1.1.-
YGL040C	HEM2	VII	delta-aminolevulinic acid dehyd...	26.73 %	4.2.1.24
YGL041C-B		VII	delta-aminolevulinic acid dehydrat...	21.57 %	

Select Genes: All genes Metabolic

Export:

Alpha Value: 0.2

Commit to Database:

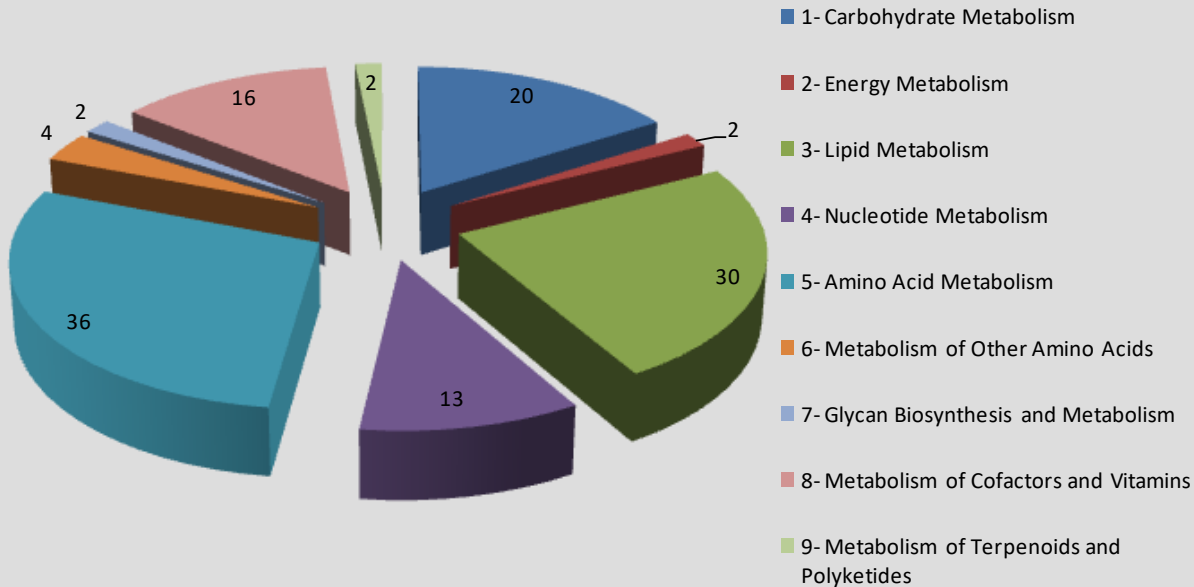
Dias *et al.* (2010) Computers Application in Biotechnology

METABOLIC MODELS

K. LACTIS RECONSTRUCTION



Metabolism - new enzymes

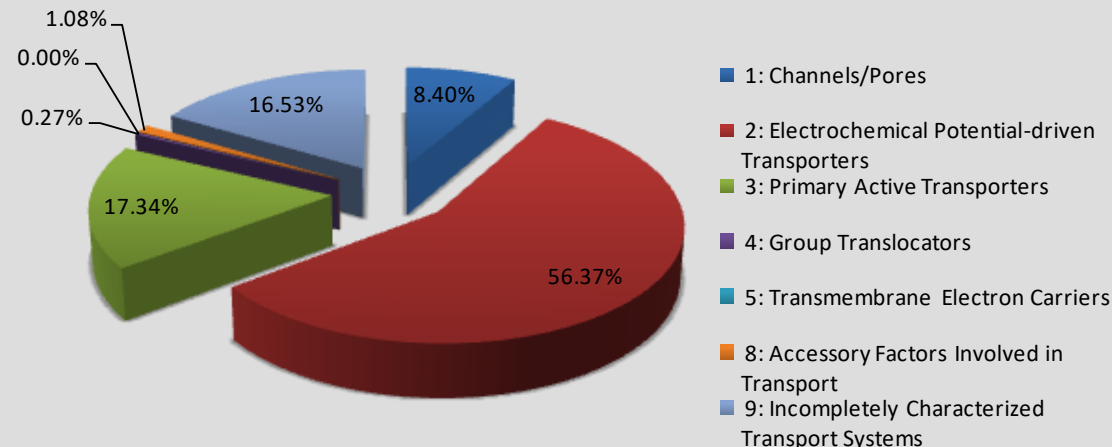


MOTIVATION

- Growth on **lactose** as a sole carbon source
- Various industrial applications, especially in the **dairy industry** but also host for **recombinant proteins**
- Molecular tools that make it amenable to **genetic manipulation**
- Evolutionary distance to *S. cerevisiae* allows to perform **comparative studies** between these two species

# of <i>K. lactis</i> genes with:	distinct	total
Yeast metabolic homologues	1627	1725
<i>K. lactis</i> transporter classification (TC) annotation	6	6
Other metabolic homologues	62	70

TC numbers



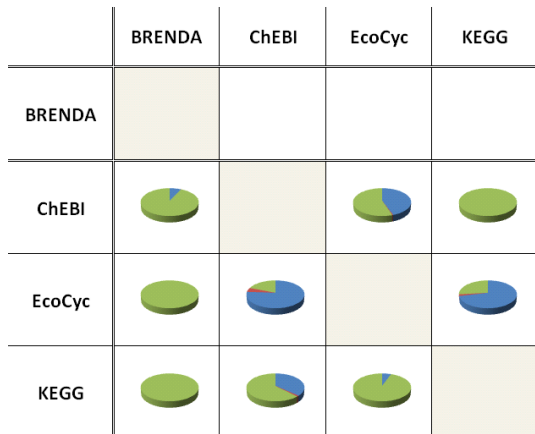
METABOLIC MODELS

DATA INTEGRATION DURING RECONSTRUCTION

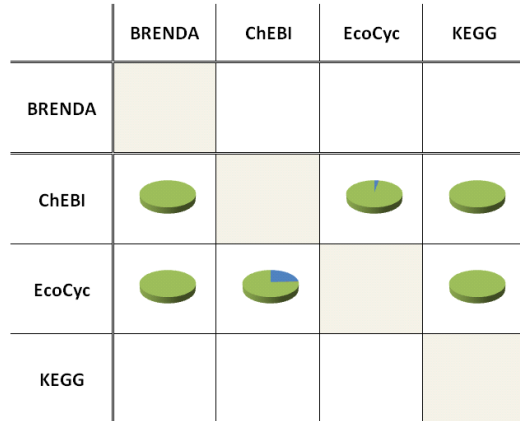


Matching Problems for compounds in *E. coli* in different databases

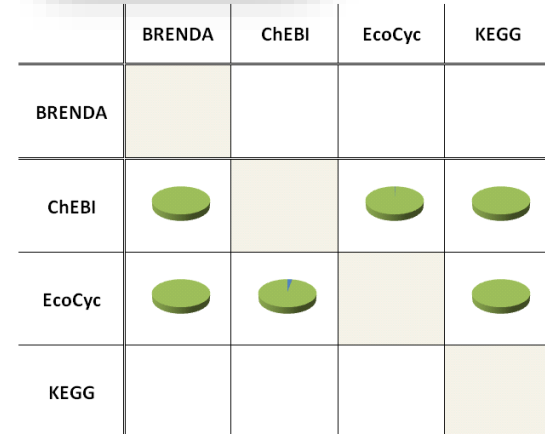
a) CAS



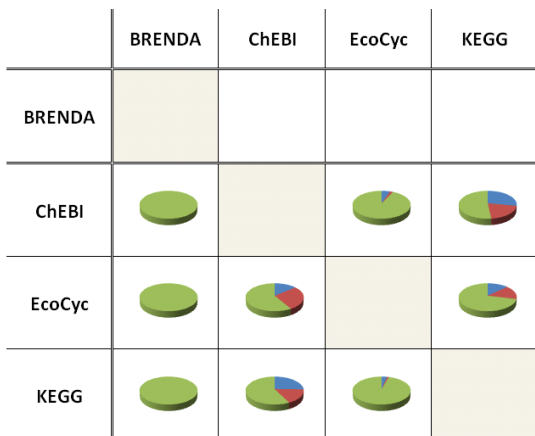
b) InCHI



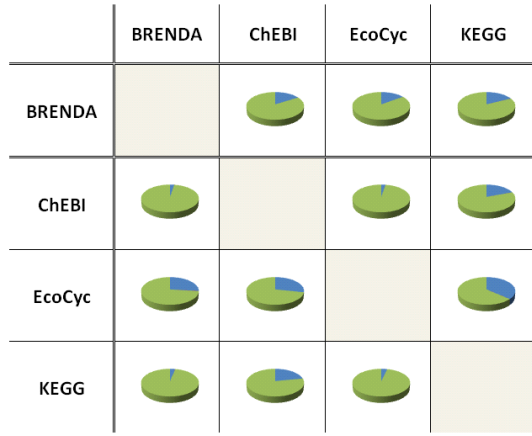
c) SMILES



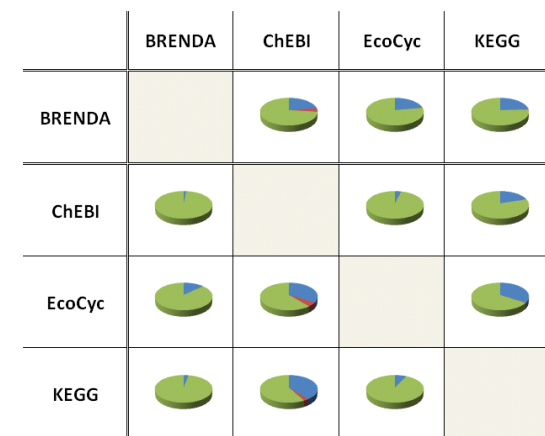
d) Formulas



e) Common names



f) All names



METABOLIC MODELS TEXT MINING FOR AIDING RECONSTRUCTION

Tools for automatically inferring metabolic and regulatory networks from literature data - @Note



Lourenço et al. (2009),
J Biomedical Informatics
42, 710-720

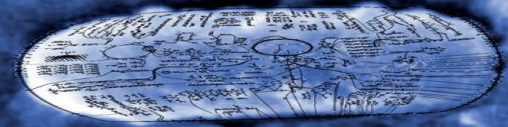
The screenshot shows the @Note software interface. On the left, there's a 'Clipboard' and 'ANoteProject' tree. The main window displays search results for the keyword 'Escherichia coli stringent response'. A specific article is highlighted: 'Global gene expression during stringent response in Corynebacterium glutamicum'. The abstract is visible, and a green arrow points to a 'D06' tag in the text. Below the results, there's a 'Results Table Viewer' showing 294 total results.

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 rRNA 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 rates of rRNA 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 hydroxamate, an amino acid analogue that inhibits rRNA synthesis, triggered a stringent response characterized in part by intracellular accumulation of ppGpp and this response was completely blocked in a relA null mutant, and partially blocked in a relC mutant harboring a mutation that allows the ribosome to synthesize ppGpp. Substrate synthesis of ppGpp was isolated from wild-type and relA relC mutant strains using a relA⁻ relC⁻ SubC⁻ synthetized ribosome. The formation of relA or relC null blocks that trigger this mechanism was further characterized by demonstrating that the relA relC tetrahydroline inhibits (ppGpp) synthesis in an in vitro system. The ppGpp synthetase catalyzed 235 / 55 rRNA gene transcription but at a concentration much lower than that observed in the relA relC null mutant. On the other hand, changes in ppGpp gene promoter activity tightly correlated with changes in the GTP but not ATP concentration. Also, (p) 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* rRNA polymerase activity directly in vivo, as recently proposed for *Bacillus subtilis* rRNA transcription (L.

This screenshot shows a detailed view of a document within the @Note software. The document is titled 'Expression of spoT in Borrelia burgdorferi during Serum Starvation' from the Journal of Bacteriology. The abstract is visible, and a context menu is open over the text, showing various actions like 'Add Tag', 'Correct Tag', etc. The right sidebar shows a hierarchical view of 'bio_classes' including 'technique', 'gene', 'compound', 'pathway', 'regulatory_gene', 'state', 'metabolic_gene', 'organism', 'enzyme', 'verb', 'transcription_factor', and 'rna'. A legend at the bottom identifies the colors used for these classes: technique (pink), compound (purple), state (green), reaction (yellow), pathway (light blue), metabolic_gene (dark blue), gene (orange), regulatory_gene (cyan), and organism (magenta).

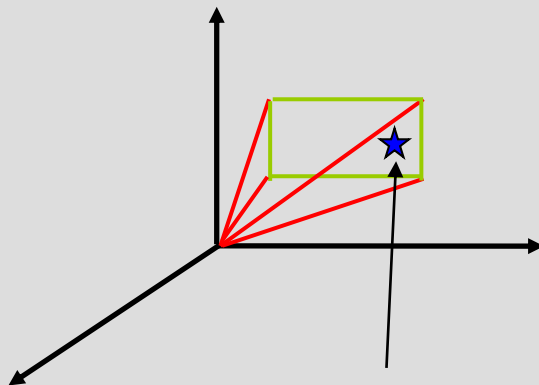
SIMULATION

STOICHIOMETRIC-REGULATORY MODELS

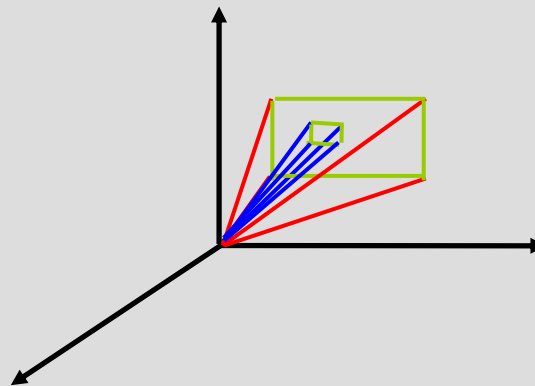


Ways to reduce the cone of solutions given by the stoichiometric model

- By optimizing a given criterion – FBA, MOMA, ROOM...
- By the introduction of regulatory information (ex: Gene Networks)



Particular Solution



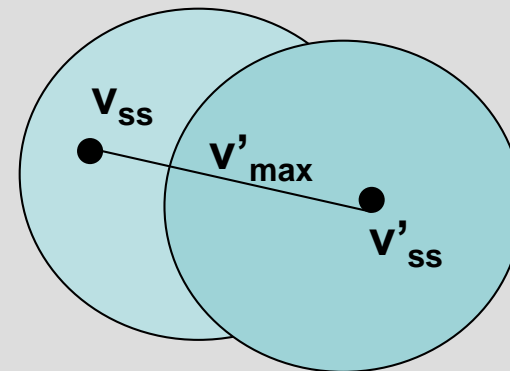
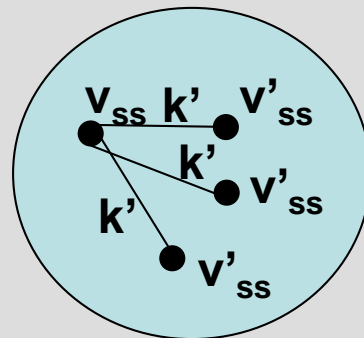
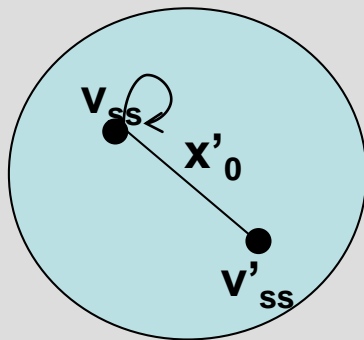
Reduction of the solutions space

INTRODUCTION
METABOLIC MODELS
SIMULATION
IN SILICO METABOLIC ENGINEERING
PROCESS ENGINEERING
CONCLUSIONS

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



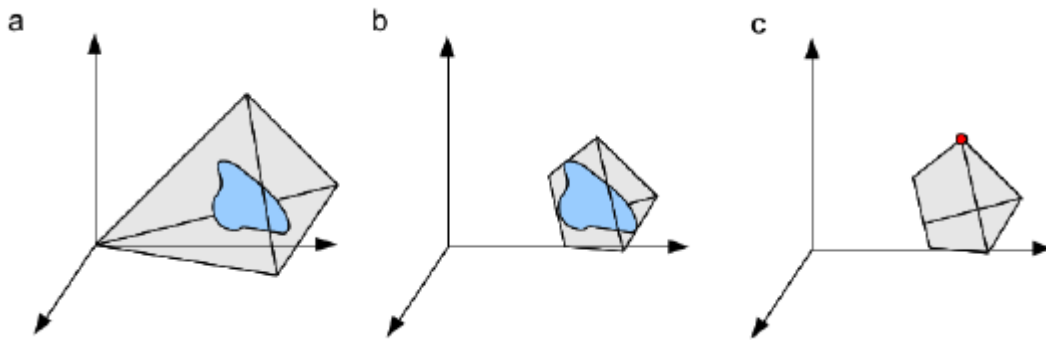
- Often, we have kinetic information for parts of the network
- What can we learn from the kinetics of those parts of the network that we know?
 - If reaction rates in the stoichiometric models are constrained by v_{\max} , then the flux space given by the stoichiometry should be reachable by changing the kinetic parameters...
 - Or not?



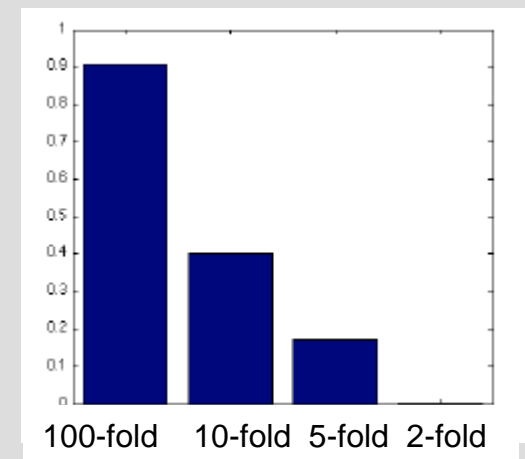
GAP BETWEEN DYNAMIC AND STOICHIOMETRIC MODELS



- Construction of a kinetically feasible flux cone:
 - a) Limiting the range of the kinetic constants results in a smaller feasible space.
 - b) The flux cone can be adjusted to fit the feasible space.
 - c) Simulation methods such as FBA can use the reduced flux cone to search for optimal solutions.



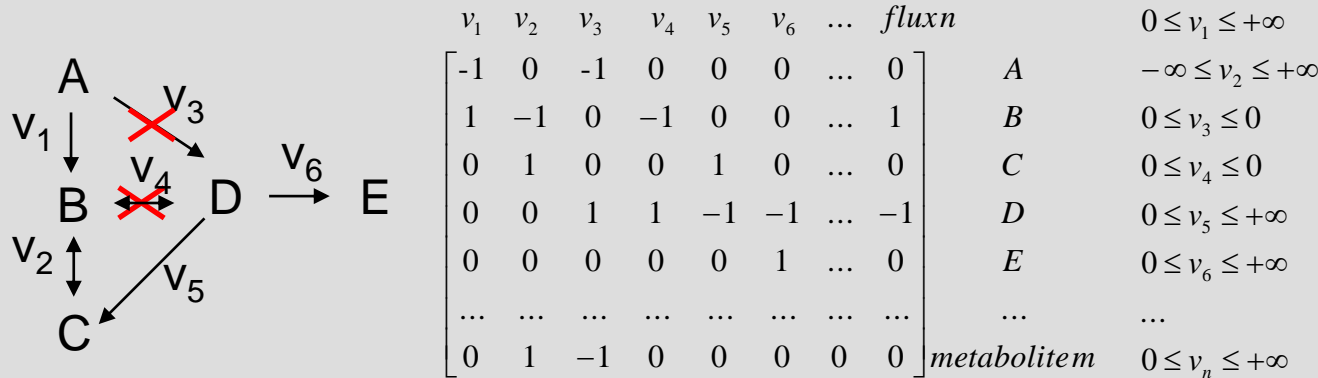
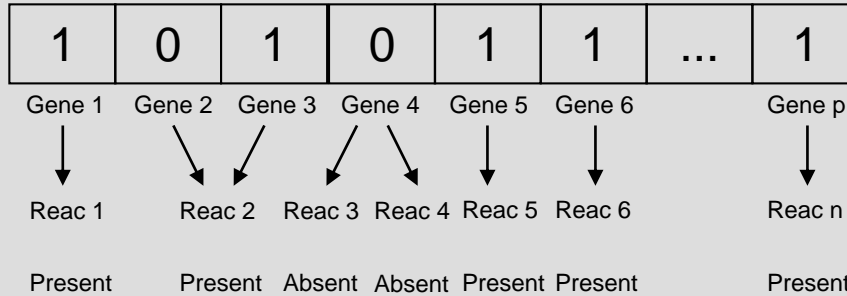
- Effect of constraining the range of variation of the kinetic constants of the dynamic model of the central carbon metabolism in the volume of the solution spaces



IN SILICO METABOLIC ENGINEERING OPTGENE



Modified "Genome"



Representation of the
metabolic genotype

Modified reaction network

Modified metabolic model

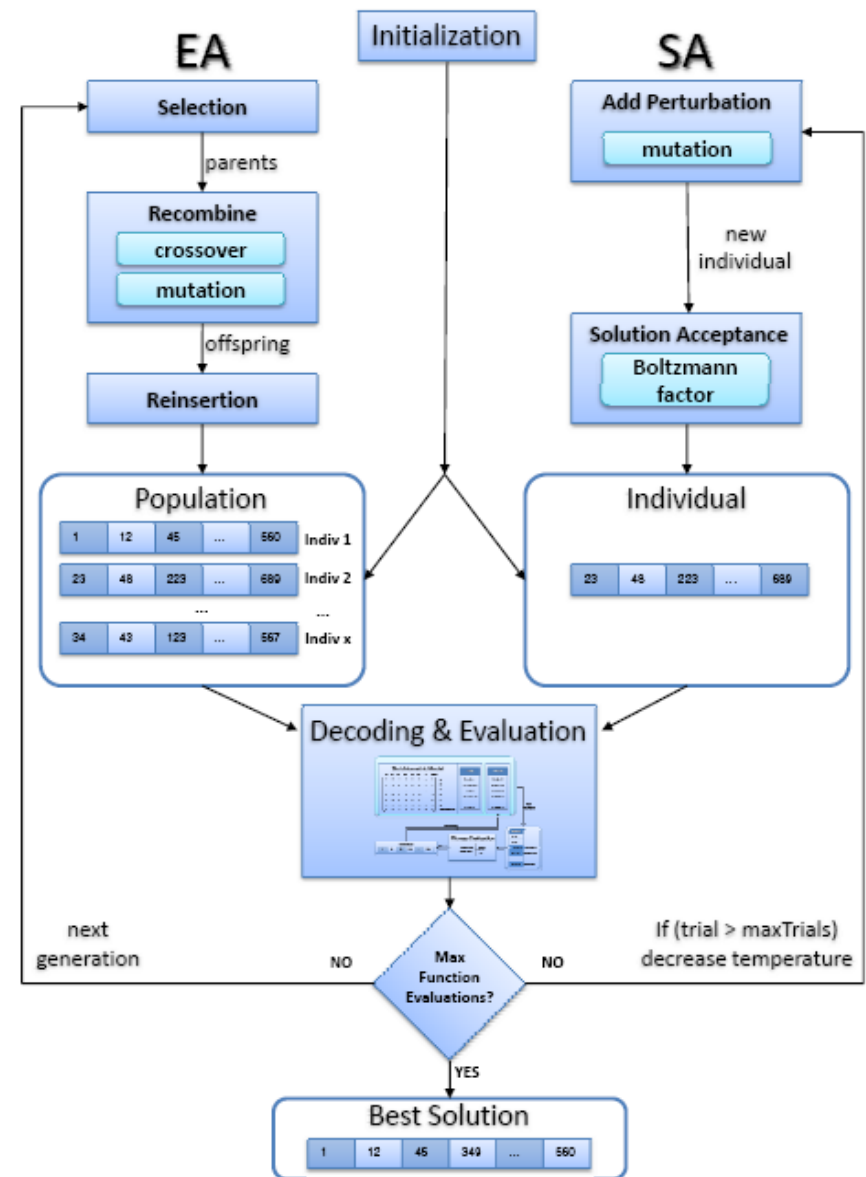


Tools and Algorithms for Optimization (OptGene Algorithms):

- EAs – Evolutionary Algorithms
- SA – Simulated Annealing algorithms
- Local Search

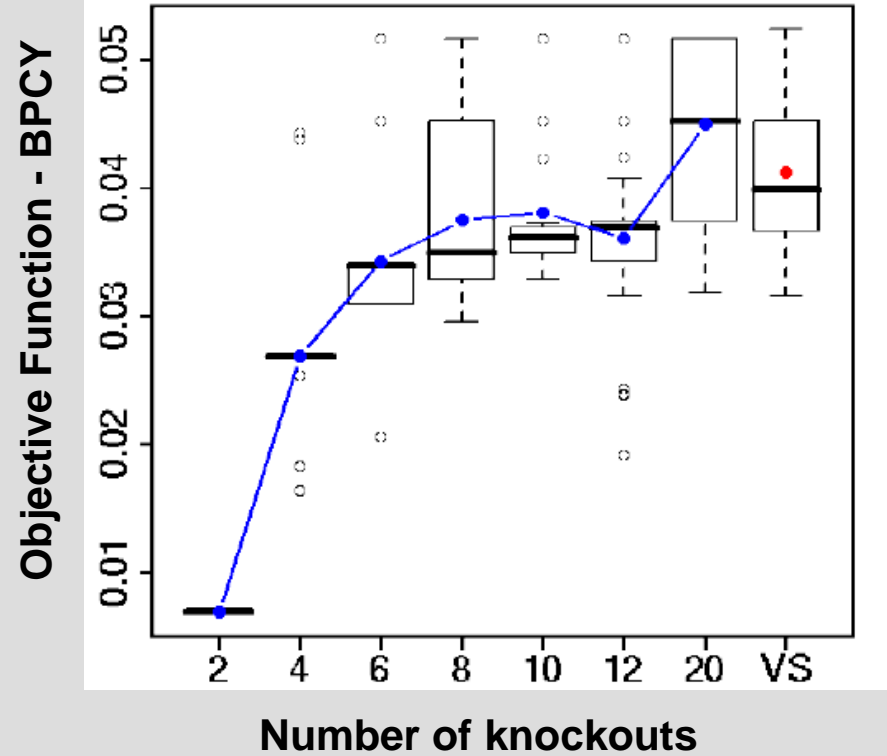
Patil et al (2005) *BMC Bioinf* 6

Rocha et al (2008) *BMC Bioinf* 9



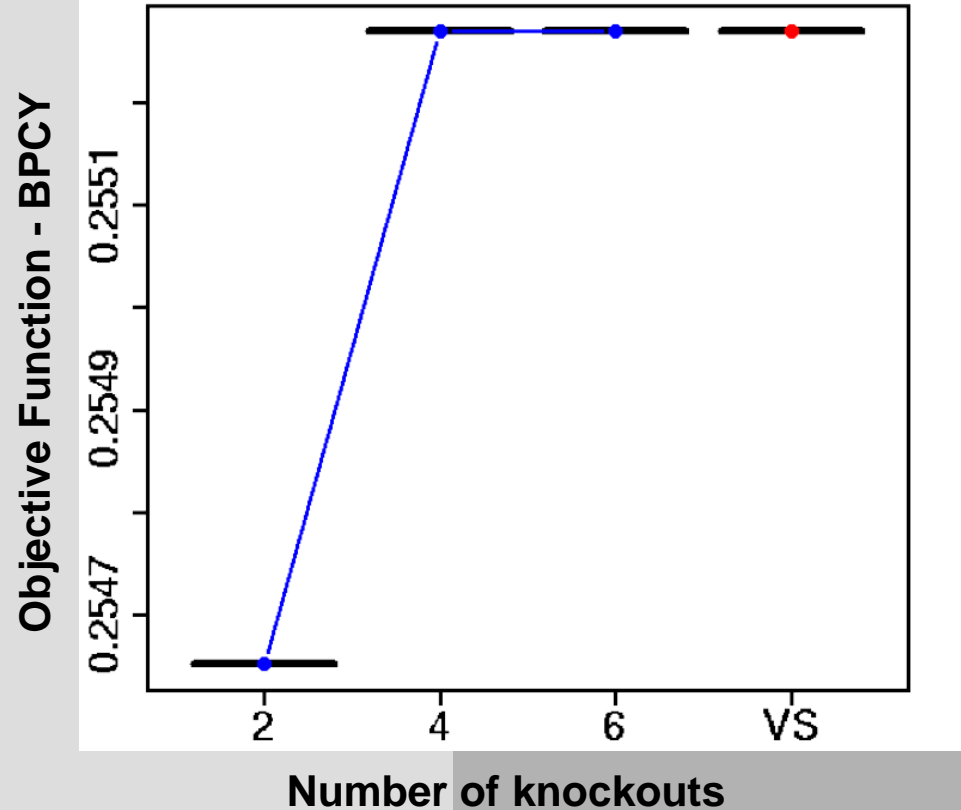


***S. cerevisiae*, succinate, EA**

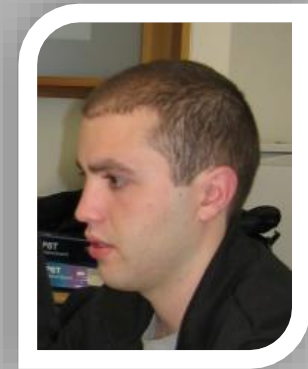


Rocha et al (2008) *BMC Bioinf* 9

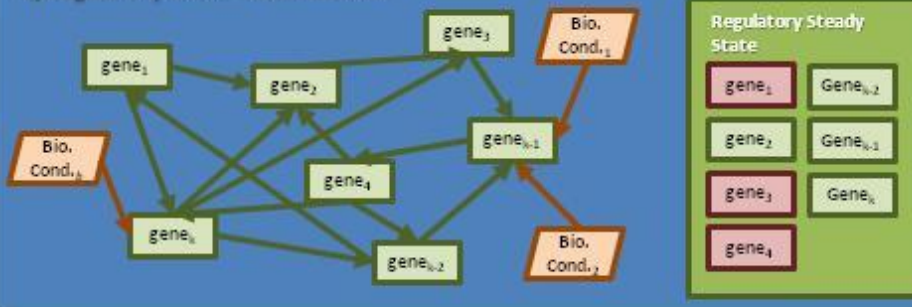
***E. coli*, anaerobic lactate, SA**



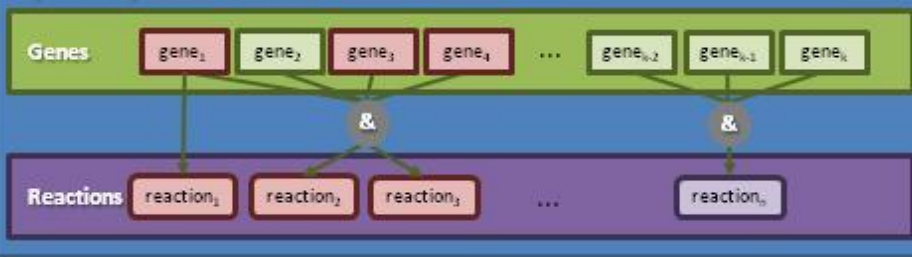
IN SILICO METABOLIC ENGINEERING INCORPORATION OF REGULATORY MODELS



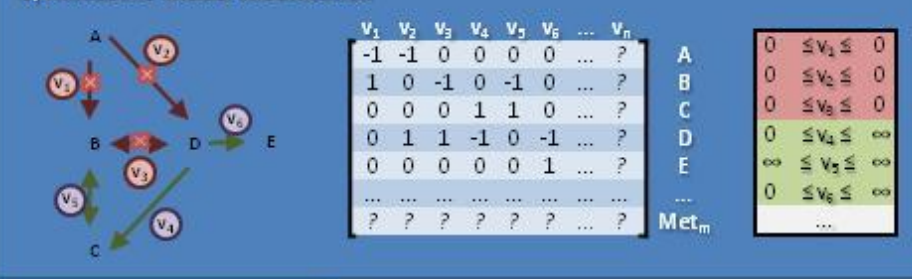
1) Regulatory Network Simulation



2) Transcriptional Information



3) Metabolic Model Modification



4) Mutant Simulation

FBA
MOMA
ROOM

Results showing the differences obtained by deleting genes or reactions for lactate production with *E. coli*

Optimization Type	Algorithm	Yield	Nr knock.
Reactions	SA	0,348	17
Genes	SA	0,293	12

Vilaça et al., *Biosystems* (in press)

IN SILICO METABOLIC ENGINEERING SUCCESSFUL APPLICATIONS OF THE ALGORITHMS



INTRODUCTION
METABOLIC MODELS
SIMULATION
IN SILICO METABOLIC ENGINEERING
PROCESS ENGINEERING
CONCLUSIONS

- Production of Succinate with *S. cerevisiae*
- Production of Succinate with *E. coli*
- Production of Sesquiterpenes with *S. cerevisiae*
- Production of aminoacids with *E. coli* (ongoing)



Open-source

- allows all users to use the tool freely
- invites the contribution of other researchers

User-friendly

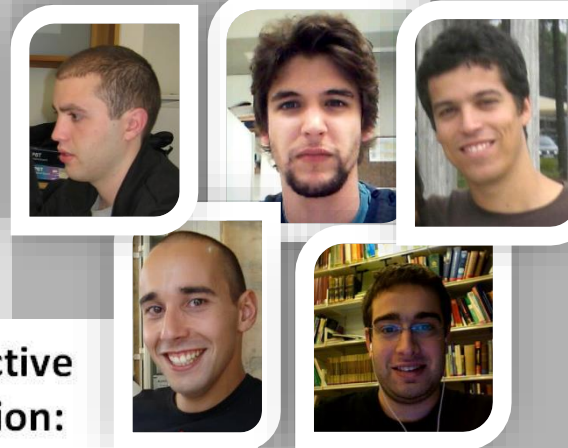
- facilitates its use by users with no/little background in modeling/informatics
- integrates a tool that allows the visualization of the metabolic models and results

Modular

- facilitates the addition of specific features by computer scientists / bioinformaticians
- based on the general-purpose AIBench platform

Compatible with standards

- SBML- Systems Biology Markup Language
- Cell Designer layouts



Objective function:
Maximize the production of a given metabolite

Determine the set of fluxes (in steady state)

Override Model

Original Model

Strain Optimization

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

Phenotype Simulation

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

Environmental Conditions

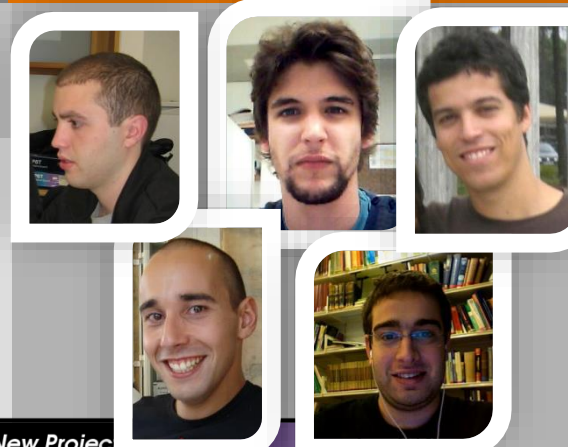
Genetic conditions (e.g. knockouts)

Model

- ✓ Stoichiometric
- ✓ Regulatory
- ✓ ...



IN SILICO METABOLIC ENGINEERING OPTFLUX



File Simulation Optimization Options Help

New Project Wizard Simulation Wizard OPT Optimization Wizard

Clipboard

OptFlux

Project

- Small S. cerevisiae **A**
- Original Model
 - Flux Bounds
 - Metabolites
 - Stoichiometric Model
- Simulation Results
- Optimization Results
- Environmental Conditions
 - glucose minimal medium
 - Variability Analysis Results
- Critical Reactions
- Model Graph

Model Graph **B**

Metabolic pathway diagram showing the conversion of D-Lactate to Ethanol and Acetyl phosphate. Key enzymes and metabolites are highlighted with colored boxes (blue for genes, green for metabolites). A table on the right lists metabolites with their IDs and types.

Na...	Type	Sp...	id
Pyr...	SIM...	s58	sa60
D-L...	SIM...	s60	sa62
NAD	SIM...	s61	sa63
NA...	SIM...	s62	sa64
H	SIM...	s63	sa65
Ace...	SIM...	s65	sa67
Co...	SIM...	s56	sa69
PI	SIM...	s66	sa68
H2	SIM...	s69	sa73
CO2	SIM...	s68	sa71
H	SIM...	s63	sa72
NAD	SIM...	s61	sa77
H	SIM...	s63	sa79
Eth...	SIM...	s70	sa75
NADH	SIM...	s71	sa79
ADP	SIM...	s77	sa87
Ace...	SIM...	s72	sa82
ATP	SIM...	s73	sa83
AMP	SIM...	s74	sa84
PPI	SIM...	s75	sa85
Ace...	SIM...	s57	sa59
For...	SIM...	s55	sa57
Co...	SIM...	s56	sa58
1.1...	PR...	s64	sa66
2.3...	PR...	s67	sa70
no ...	PR...	s98	sa74
1.2...	PR...	s102	sa...
2.3...	PR...	s99	sa...
6.2...	PR...	s76	sa86
2.7...	PR...	s78	sa88
tdcF	GENE	s103	sa...

Model Graph

Show Genes
 Show Proteins
 Show Simple Molecules
 Show Complex Molecules
 Show Secondary Reactions

New Project

Please select the type of SBML and the container file:

Pure SBML
 CellDesigner SBML

SBML File:

Reactions

Name	Reactants	Direct...	Products
R_SDHcompl...	FAD + SUC	----->	FADH2 + FUM
R_ZWF	G6P + NADPcyt	----->	NADPHcyt + G15L
R_FBA	F16P	<-----	DHAP + GA3P
R_LSC1LSC2	SUCCOA + ADP	----->	SUC + ATP
R_SUC	SUC	----->	
R_PDC	PYR	----->	CO2 + ACA
R_NADHX	24.0 x ADP + 20.0 x NADHmit	----->	20.0 x NADmit + 24.0 x ATP
R_ACETR	ACE	----->	CI
R_CIT	ACCOAmit + OAA	----->	CO2 + ACCOAmit + NADHmit
R_PDH	NADmit + PYR	----->	MAL
R_FUM1	FUM	<-----	
R_PFK	ATP + F6P	----->	F16P + ADP
R_TALL	S7P + GA3P	<-----	
R_ATPX	ATP	----->	ADP

Knockouts

search: fum

Name	Lower Bound	Upper Bound
R_FUM1	0.0	10000.0
R_PFK	-10000.0	10000.0
R_TALL	0.0	10000.0
R_PFX	0.0	10000.0
R_FRD52	0.0	10000.0
R_KG1KGD2	0.0	10000.0
R_FAGW	0.0	10000.0
R_BIOMASSX	0.0	10000.0
R_ACO	-10000.0	10000.0
R_DAR	0.0	10000.0
R_ACS	0.0	10000.0
R_DAR	0.0	10000.0
R_MAE1	0.0	10000.0
R_PIC	0.0	10000.0
R_DND	-10000.0	10000.0
R_ENO	-10000.0	10000.0
R_PFK	0.0	10000.0
R_PFI	-10000.0	10000.0
R_PFK	1.0	1.0
R_GLD	-10000.0	10000.0
R_PFK	-10000.0	10000.0
R_CAT2	0.0	10000.0

add selected flux to knockouts list

Data: Original Model Remove

Algorithm: FBA MOMA ROOM

Flux to maximize: R_BIOMASSX minimize?

use env. conditions: glucose minimal medium

Cancel Run

E

Project

Small S. cerevisiae

EA Parameters

representation: Set-Based Representation
 population size: 10
 generations: 10
 knockouts: 8
 variable size:
 use essential genes:

Simulation Parameters

flux to maximize in FBA/MOMA/ROOM: R_BIOMASSX
 desired flux: R_SUC
 substrate: R_HXX
 simulation method: ROOM
 objective function: YIELD
 minimum biomass %: 80%
 use env. conditions: glucose minimal medium

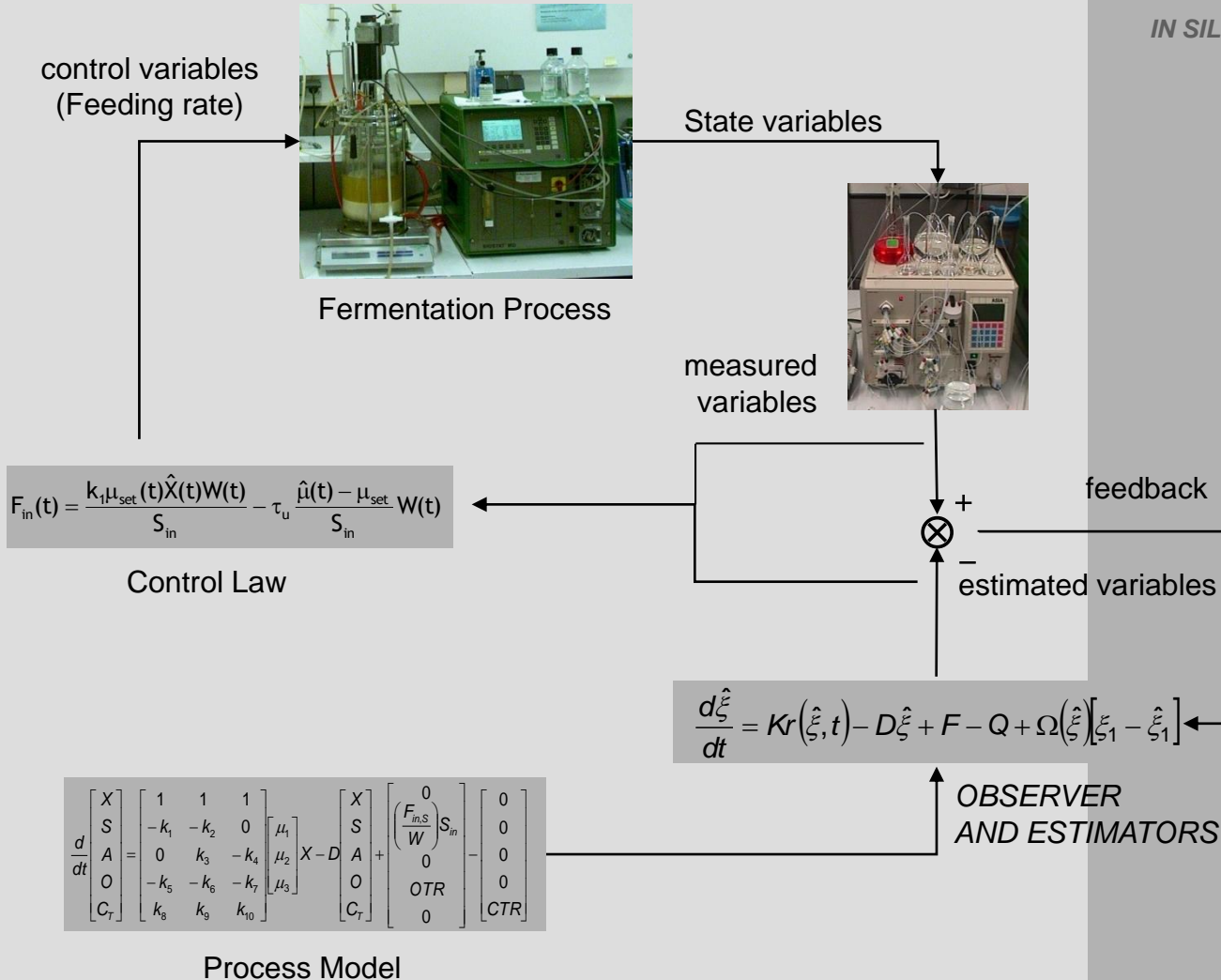
cancel run

PROCESS ENGINEERING

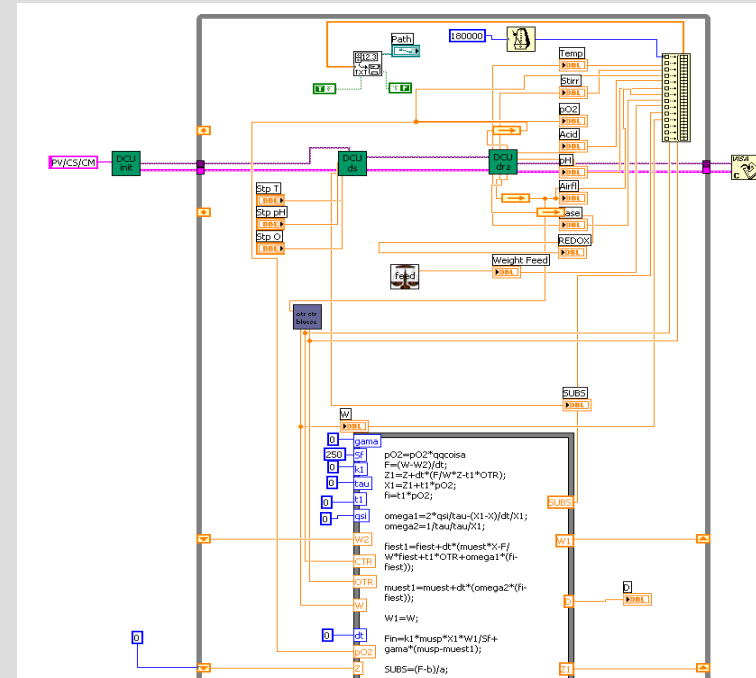
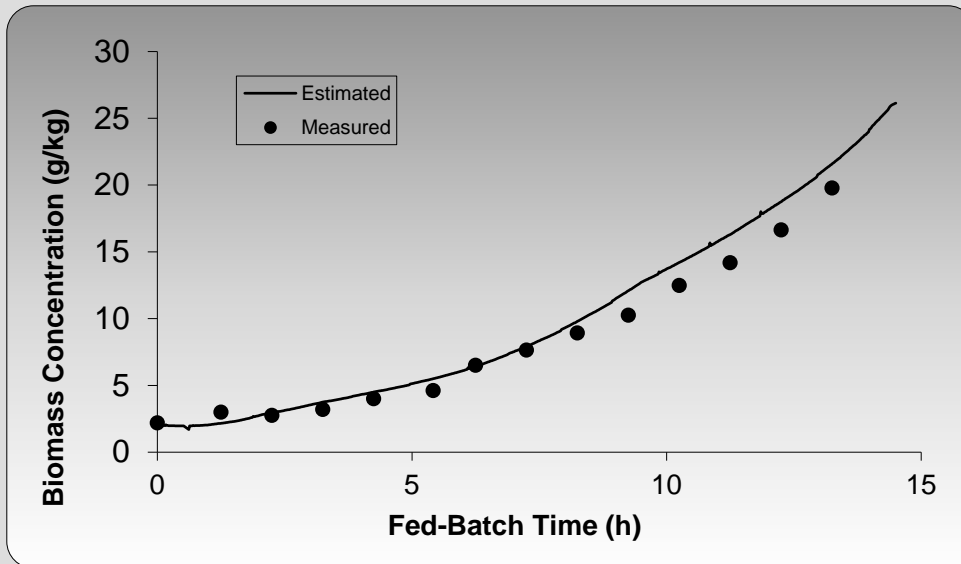
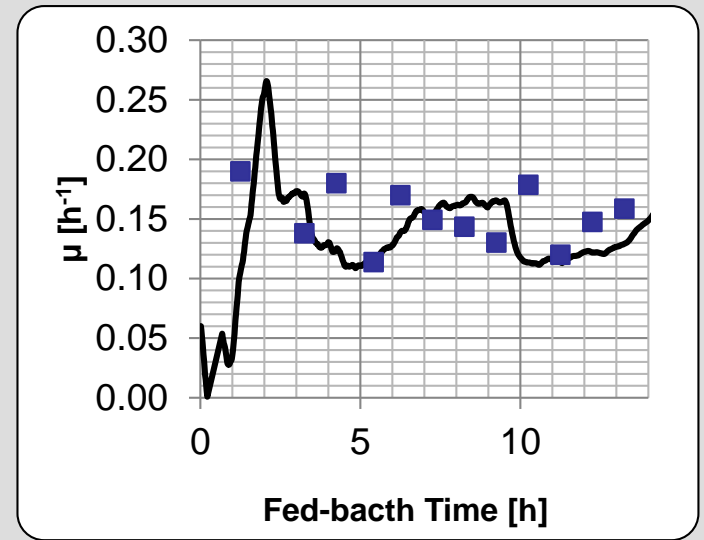
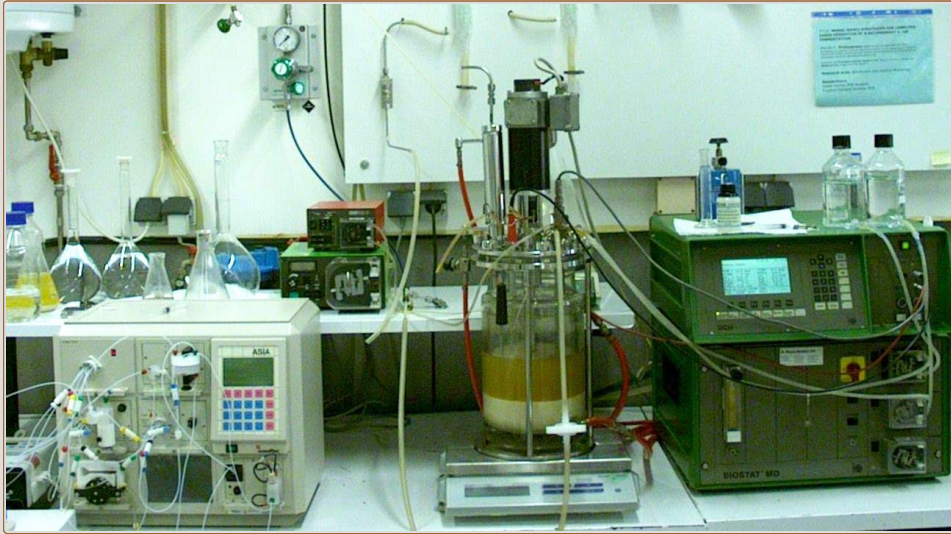
MODEL-BASED CONTROL AND OPTIMIZATION



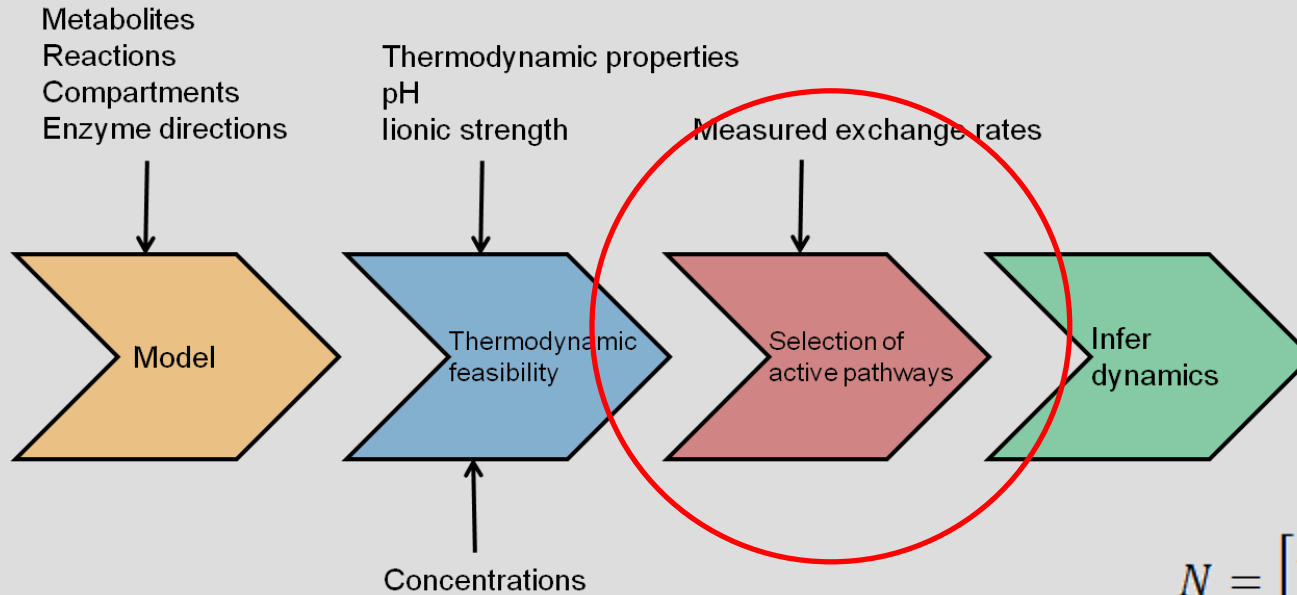
INTRODUCTION
 METABOLIC MODELS
 SIMULATION
 IN SILICO METABOLIC ENGINEERING
 PROCESS ENGINEERING
 CONCLUSIONS



PROCESS ENGINEERING MODEL-BASED CONTROL



PROCESS ENGINEERING LARGE-SCALE MODELS



Soons et al, JPC, 2011

Selection based on objective function:

$$J = RMSE + c_1 \cdot P + c_2 \cdot MS$$

↑
weighted sum of squared errors

↑
penalty for inefficiency of each elementary mode

←
penalty for model size

$$N = \begin{bmatrix} N_s \\ N_\xi \end{bmatrix}$$

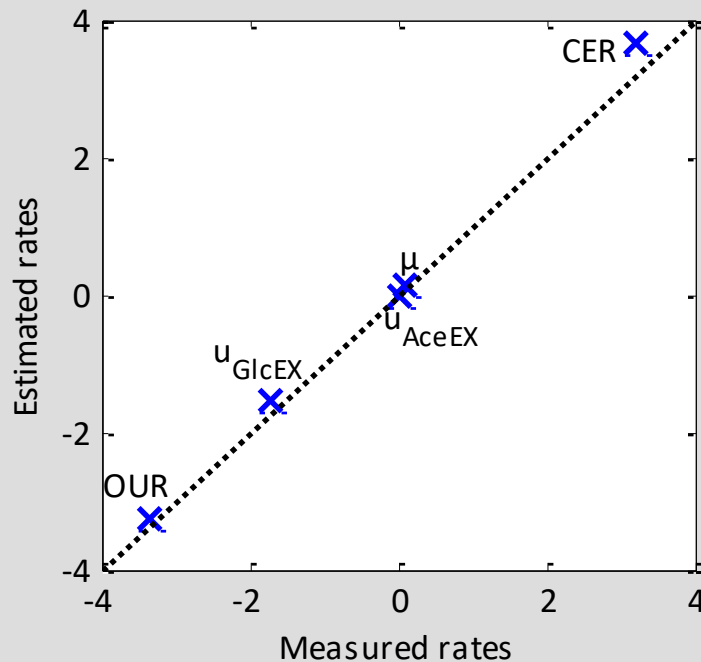
$$\frac{ds}{dt} \cong 0 \Rightarrow \begin{bmatrix} N_s \\ N_\xi \end{bmatrix} v - \begin{bmatrix} 0 \\ v_m \end{bmatrix} = 0$$

$$\frac{d\xi}{dt} = Kr(t) + u(t)$$

$$K = N_\xi \cdot E$$



- Reduction of the original model from 2706 to 3 elementary modes:
 - EM1 : $0.21 \text{ GlcEX} + 0.67 \text{ O}_2 \rightarrow 0.015 \text{ Biomass} + 0.67 \text{ CO}_2$
 - EM2 : $0.056 \text{ GlcEX} + 0.22 \text{ O}_2 \rightarrow 0.056 \text{ Acetate} + 0.22 \text{ CO}_2$
 - EM3 : $0.040 \text{ GlcEX} + 0.24 \text{ O}_2 \rightarrow 0.024 \text{ CO}_2$
- Good match with measured data from literature



- Initially: 2706 elementary modes

Soons et al, JPC, 2011



- So far, rational metabolic engineering design has only been performed with **stoichiometric models** and indicate only **knockout and gene additions**
- Nevertheless, it is already possible to **improve *in silico* the production of targeted compounds**
- Predictions are enhanced if **regulatory information** is added to the models
- Additional constrains maybe derived if **kinetic information** is available for part of the network
- Model reconstruction is far away from being made in a standard way
- The bridge between process engineering and optimization and large scale models is still not there...

ACKNOWLEDGMENTS





Funding Agencies

Portugal - FCT

EC – SysInBio CA

Dupont – Science & Engineering Grant

Dupont – R&D contract

MIT-Portugal Program

International Contributors

Kiran Patil – EMBL, Germany

Jens Nielsen – Chalmers – Sweden

Bruce Tidor – MIT

Silas Villas-Boas – Univ. of Auckland, New Zealand

Andreas Gombert – University of São Paulo, Brazil

Jean-Francois Tomb and Marcellinus Pont – Dupont, USA

Contributors – IBB – UMinho

Isabel Rocha(PI)

Anália Lourenço (PI)

Zita Soons

Sonia Carneiro

Çarla Portela

Óscar Dias

Daniel Machado

Paulo Maia

Pedro Tiago

Paulo Vilaça

Rafael Carreira

Rafael Costa

Hugo Costa

Simão Soares

Marisa Cunha

Contributors – CCTC– UMinho

Miguel Rocha (PI)

José Pedro Pinto