T1

Modeling integrated biochemical networks: methods, advances and perspectives

Nuno Tenazinha[a,c,d], **Susana Vinga**[a,b]

ntenazinha@kdbio.inesc-id.pt svinga@kdbio.inesc-id.pt

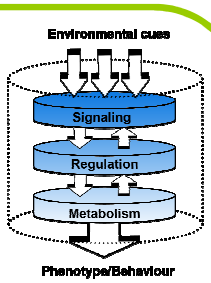
[a] *INESC-ID Instituto de Engenharia de Sistemas e Computadores: Investigação e Desenvolvimento*, Portugal
[b] *FCM/UNL Faculdade de Ciências Médicas – Universidade Nova de Lisboa*, Portugal
[c] *ITQB/UNL Instituto de Tecnologia Química e Biológica – Universidade Nova de Lisboa*, Portugal
[d] *PDBC/IGC PhD Program in Computational Biology – Instituto Gulbenkian de Ciência*, Portugal

1 Abstract

Signaling networks, gene regulation and metabolism have frequently been modeled independently. However, capturing their intertwinement is a key step for understanding how cellular systems develop integrated responses to their changing environment. We review current methods for modeling integrated biochemical networks, illustrating their potentials with successful case-studies in organisms such as *S. cerevisiae* and *E. coli*. Comparisons are also established regarding scalability with network size, required computational power and model predictive properties.

2 Introduction

Current understanding of cellular behaviors has been achieved from extensive research in Molecular and Systems Biology. The last decade advances in the 'omics' high-throughput technologies led to an unprecedented increase in the amount and quality of available biological information. It also fostered the emergence of Mathematical Modeling as an essential tool to understand natural phenomena.



Problems of integrative models

- Sheer complexity of the integrated systems, unknown interactions between signaling, regulatory and metabolic layers;
- Different time-scales of biological processes;
- Lack of sufficient data hinders the model building process and biases the modeler towards particular subsystems and problems;
- Choice of appropriate mathematical framework to capture the interconnected networks: level of detail vs. computational burden associated to large-scale models;

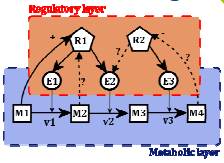
BUT... integration is required!!

Cells operate as an ensemble of different process layers to produce response phenotypes in their changing environment

3 Structural and Stoichiometric Modeling

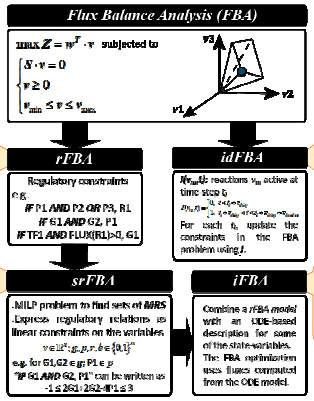
Network-based analysis

- Use Graph theory to learn unknown interactions among metabolites, enzymes, regulators, genes
- Reconstruction of genome-scale integrated networks



Constraints-based models

From stoichiometric data, *Flux Balance Analysis (FBA)* uses pseudo-steady state conditions and basic constraints to find optimal flux distributions that reflect phenotypic behaviors.



Regulated metabolism can be analyzed by including Boolean constraints on the bounded convex space of flux distributions

Search for compatible *Metabolic and Regulatory Steady States* (MRS) improves the model predictions

Specify active or inactive reactions by taking fast reactions at steady state and updating slower reactions from matrix *I*

ODEs can be embedded into FBA to improve model accuracy if information is available

BUT...
Lack of detail in the intracellular dynamics!

FBA-derived models do not fully comprise the flux-metabolite dependencies and this severely restricts their predictive properties

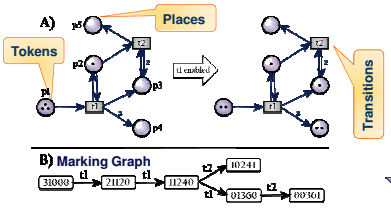
4 Kinetic Modeling

Discrete Logical models

- Discrete qualitative models - do not rely on parameter knowledge or estimation;
- Relationships between fluxes, gene expression status and metabolite concentrations expressed through logical functions;

Petri-Nets

- Regulatory effects embedded by adding variables accounting for presence or absence of regulators
- Transitions can fire with associated rates using the *Hybrid Functional Petri Net* framework



Hybrid Systems

- Semi-quantitative approach - couples continuous models with discrete state transitions;
- When **1)** Regulatory effects result in altered dynamic modes, **2)** different time-scales are implied **3)** approximating nonlinearities with piecewise-linear functions;

Continuous ODE-based models

- Sets of ordinary differential equations (ODEs) provide models closer to biological reality with quantitative predictive power;
- Require a wealth of *a priori* data on the mathematical form of the fluxes and on the system's parameters

State-variables: compounds' concentration

Stoichiometric Matrix

Parameters

$$\dot{X} = dX/dt = S \cdot v(X, \theta_{kin}, \theta_{env})$$

Fluxes

e.g.

$$v = V_m \frac{X^n}{K_m{}^n + X^n}$$

n = 1; Michaelis-Menten equation
n > 1; Hill equation

$$v = \gamma \prod_{i=1}^{N} X_i^{h}$$

Power-law approximation

Increasing level of detail

5 Summary and Perspectives

Reconstruction of integrated systems but lack the dynamic aspect

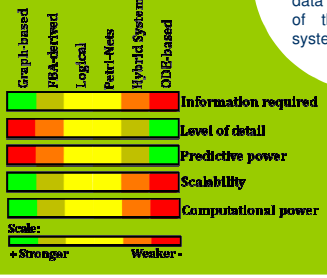
Stoichiometric approaches have proven successful in the **building of whole-genome models** of regulated metabolism

Deliver **qualitative predictions** with few *a priori* data requirements

Provide **quantitative predictions**, but require information on the system's parameters and are computationally more expensive

Are not fully explored but represent a **good compromise between semi-quantitative predictions and few *a priori* information needed**

Method	Application	References
STRUCTURAL AND STOICHIOMETRIC MODELING		
Network-based		
Graph-based	Reconstruction of the central carbon metabolism of *E. coli*;	(Yeang and Vingron 2006)
OT-based	Mimic the *E. coli* rFBA model by Covert and Palsson 2002;	(Kaleta, Centler et al. 2008)
Constraints-based		
rFBA	Regulated metabolic model of *E. coli* metabolism including glucose/acetate growth and glucose-lactose diauxie;	(Covert and Palsson 2002)
	Integrated genome-scale metabolic model of *E. coli*;	(Covert, Knight et al. 2004)
	Integrated genome-scale metabolic model of *S. cerevisiae*;	(Herrgard, Lee et al. 2006)
srFBA	Mimic the whole-genome *E. coli* model by Covert, Knight et al 2004;	(Shlomi, Eisenberg et al. 2007)
iFBA	Extension of the Covert and Palsson 2002 model including the phosphotransferase catabolite repression;	(Covert, Xiao et al. 2008)
idFBA	Model of the HOG pathway in *S. cerevisiae*;	(Lee, Gianchandani et al. 2008)
KINETIC MODELING		
Discrete		
Logical-based	Regulated carbon metabolism of *E. coli* subjected to three different carbon sources: glucose, glycerol and acetate;	(Asenjo, Ramirez et al. 2007)
Petri Nets	Tryptophan biosynthetic pathway in *E. coli*;	(Simao, Remy et al. 2005)
	Regulation of the urea cycle in the liver;	(Chen and Hofestadt 2003)
	Regulation of early human haematopoiesis;	(Troncale, Tahi et al. 2006)
Continuous		
ODE-based	Regulation of the *lac* operon involved in lactose uptake and metabolism;	(Babloyan.A and Sanglier 1972; Wong, Gladney et al. 1997; Santillan and Mackey 2004; van Hoek and Hogeweg 2006; van Hoek and Hogeweg 2007)
	Yeast response to osmotic shock (including phosphorelay, MAP kinase cascade, transcriptional regulation, carbon metabolism and glycerol production);	(Klipp, Nordlander et al. 2005)
	Glucose-Galactose metabolism and the phenomenon of catabolite repression in *S. cerevisiae*;	(Demir and Kurnaz 2006)
Hybrid systems		
	Initiation of sporulation in *B. subtilis*;	(De Jong, Geiselmann et al. 2004)
	Nutrient deprivation in bacteria;	(Belta, Finin et al. 2004)
	Regulation of the *lac* operon in lactose uptake and metabolism in *E. coli*;	(Halasz, Kumar et al. 2007)
	Xanthophyll cycle reaction in plants;	(Cho, Johansson et al. 2005)
	Genetic toggle switch associated to the SOS signaling pathway in *E. coli*;	(Wu and Voit 2009)



- **Integrated models capturing the interplay between signaling pathways, gene regulation and metabolism are now starting to develop and are essential for a better understanding of complex cellular behaviors**
- **The choice of modeling framework must rely on the desired level of abstraction on the system, on available data for validation but also on the model's aim**

References

Tenazinha, N. and Vinga, S. A survey on methods for modeling and analysis of integrated biochemical networks. *Submitted*.

Acknowledgments

The authors acknowledge financial support from the PhD Program in Computational Biology from Instituto Gulbenkian de Ciência (sponsored by FCG, Siemens SA and FCT) (SFRH /BD/33209/2007) and by project DynaMo (PTDC/EEA-ACR/ 69530/2006, FCT).