Metabolic network structure analysis: a case study on glycolysis in *Lactococcus lactis*

Susana Vinga[1,2], **Karl Thomaseth**[3], **João M. Lemos**[1,5], **Ana Rute Neves**[4], **Helena Santos**[4], **Ana T. Freitas**[1,5]

[1] *INESC-ID Instituto de Engenharia de Sistemas e Computadores: Investigação e Desenvolvimento,* Portugal
[2] *FCM/UNL Faculdade de Ciências Médicas – Universidade Nova de Lisboa,* Portugal
[3] *ISIB-CNR Istituto di Ingegneria Biomedica – Consiglio Nazionale delle Ricerche*, Italy
[4] *IST/UTL Instituto Superior Técnico – Universidade Técnica de Lisboa,* Portugal
[5] *ITQB/UNL Instituto de Tecnologia Química e Biológica – Universidade Nova de Lisboa*, Portugal

svinga@kdbio.inesc-id.pt

1 Abstract

The dynamic modeling of metabolic networks constitutes a major challenge in systems biology. The time evolution of metabolite concentration in cells is modeled by complex systems of non-linear differential equations with a large number of parameters. A case study of the stability analysis based on the Jacobian matrix of the model equations combined with singular value decomposition of output sensitivities is presented for glycolysis in *L. lactis*. This approach shows how a preliminary structural model can be reformulated in simplified form to substantially improve the parameter estimation task.

2 Introduction

The reverse engineering step of inferring the parameters and model structure from experimental time series data is still a major bottleneck to correctly identify the network dynamical behavior. There is currently no automatic and straightforward solution that guarantees convergence to a global optimum.

Local identifiability analysis evaluates linear dependencies among parameter sensitivities of model outputs. By performing the preliminary analysis of the structural model it is possible to identify a cluster of state variables with fast equilibration dynamics that can be lumped into a single state variable. This procedure eliminates practically unidentifiable fast modes and allows estimating a reduced parameter set that accurately reproduces, upon simulation, the original experimental time series.

3 Methods

➢ **Modeling procedure**

➢ **Network topology**

Graphs.
Nodes
pools/metabolites
Edges
flow/transport of material + process modulation

Biochemical systems theory
Systems of non-linear differential equations. Fluxes approximated by Taylor series in log space

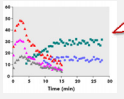
➢ **Dynamics**

Rate constants

$$\dot{X}_i = \sum_{k=1}^{m} \pm\gamma_{ik} \prod_{j=1}^{n} X_j^{f_{ijk}}$$

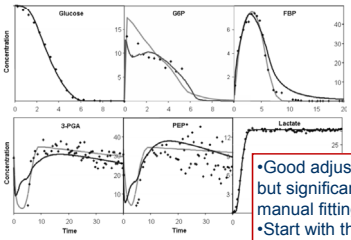
Kinetic orders

➢ **Data**



Multivariate time series of metabolite concentration (in vivo NMR)

➢ **Case study – glycolysis in *L. lactis***
Unfinished model - In Voit et al. (2006)



•Good adjustment but significant manual fitting
•Start with this solution

➢ **Structural analysis**

•**Global identifiability**: is it possible to separately and uniquely recover unknown parameters from error-free input-output data?
•**Local identifiability**: for almost every solution there is a neighborhood where no other solution exists. Related with sensitivity analysis: parameters that do not influence observed outputs are unidentifiable

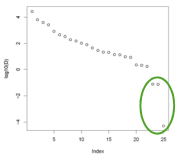
➢ **Local identifiability**

Sensitivity matrix

$$\delta\mathbf{y}(t) \approx \nabla_{\mathbf{p}}\mathbf{y}(t,\mathbf{p})\,\delta\mathbf{p}$$
$$\delta\mathbf{Y} \approx \mathbf{S}\,\delta\mathbf{p} \quad \{\partial\mathbf{Y}_i/\partial\mathbf{p}_j\}_{i,}$$

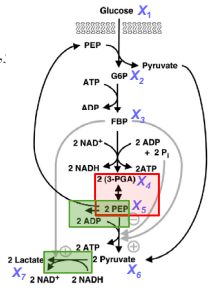
Singular value decomposition (SVD)

$$\mathbf{S} = \mathbf{U}\Sigma\mathbf{V}^T$$



Parameter variations are locally identifiable if **S** has full rank

over-parametrization?



➢ **Local stability**

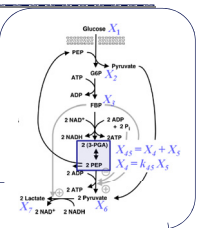
$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t),\mathbf{u}(t),\mathbf{p})$$
$$\mathbf{x}(0) = \mathbf{x}_0$$
$$\delta\mathbf{x}(t) = e^{\mathbf{A}\,t}\delta\mathbf{x}_0 = \mathbf{V}\,e^{\Lambda\,t}\mathbf{V}^{-1}\delta\mathbf{x}_0$$
$$= \sum_i c_i\,\mathbf{v}_i\,e^{\lambda_i\,t}$$

Eigenvectors $v_i \downarrow$ Eigenvalues $\lambda_i \rightarrow$

	-1506	-0.877	-0.394	-0.305	-0.0383	$-8.84\cdot10^{-6}$	0
1	0	0	0.362	0	0	0	0
2	0	0	-0.401	0	0.843	0	0
3	0	0.245	-0.032	0	0.038	0	0
4	-0.707	0.117	0.269	0	0.42	0.982	0
5	0.707	0.022	0.052	0	0.082	0.191	0
6	0	-0.962	-0.795	1	0.325	0	0
7	0	-0.002	0.001	-0.001	-0.009	-0.001	1

fast mode → fast convergence to equilibrium

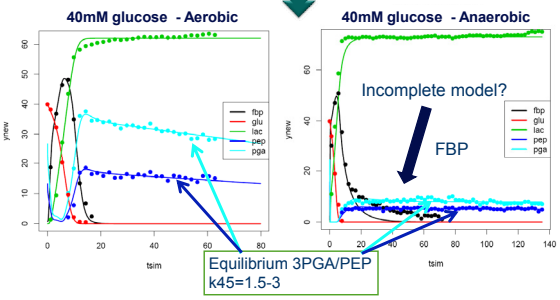
4 Results



$$\dot{X}_1 = -k(1+\alpha t^{\beta})X_1$$
$$\dot{X}_2 = \beta_1 X_1^{h_{11}} X_2^{h_{12}} X_5^{h_{25}} - \beta_2 X_2^{h_{22}} \mathrm{ATP}^{h_{2ATP}}$$
$$\dot{X}_3 = \beta_2 X_2^{h_{22}} \mathrm{ATP}^{h_{2ATP}} - \beta_3 X_3^{h_{33}} P_i^{h_{3P_i}} \mathrm{NAD}^{h_{3NAD}}$$
$$\dot{X}_{45} = 2\beta_3 X_3^{h_{33}} P_i^{h_{3P_i}} \mathrm{NAD}^{h_{3NAD}} - \beta_1 X_1^{h_{11}} X_2^{h_{12}} X_5^{h_{25}} - \beta_{51} X_3^{h_{513}} X_5^{h_{515}} P_i^{h_{51P_i}} - \beta_{52} X_5^{h_{525}}$$
$$\dot{X}_6 = \beta_1 X_1^{h_{11}} X_2^{h_{12}} X_5^{h_{25}} + \beta_{51} X_3^{h_{513}} X_5^{h_{515}} P_i^{h_{51P_i}} - \beta_{61} X_6^{h_{616}} X_3^{h_{613}} \mathrm{NAD}^{h_{61NAD}} - \beta_{62} X_6^{h_{626}}$$
$$\dot{X}_7 = \beta_{61} X_6^{h_{616}} X_3^{h_{613}} \mathrm{NAD}^{h_{61NAD}}$$

• **MODEL**
Infer parameters that better adjust experimental data using Pansym software (Thomaseth 2003) and non-linear least squares (in R)

• **SIMULATE**
Simulation with estimated parameters shows good adjustment to experimental data in both conditions



Time course for [1-^{13}C]glucose (40mM) consumption/product formation and concentrations of intracellular in non-growing suspensions of *L. lactis* strain MG1363.

5 Conclusions

- **Dynamic modeling of metabolic networks leads to complex optimization problems**
- **Structural analysis of networks improves model identification and parameter inference**
- **Priori addressing parameter sensitivity and identifiability issues can significantly improve the subsequent reverse engineering step.**
- **The application of established model building and analysis procedures can have a positive impact in the development of complex biological systems models**

References
• Neves AR et al. (2002) J. Biol. Chem. 277, 28088-28098.
• Voit EO et al. (2006) IEE Proc Syst Biol 153, 286-298.
• Thomaseth (2003) Comput Methods Programs Biomed 71(3), 189–201.
• Vinga et al. (2008) In Proc. of the 8th Portuguese Conference on Automatic Control (CONTROLO'2008). Portugal.

Acknowledgments
The authors acknowledge financial support by projects DynaMo (PTDC/EEA-ACR/ 69530/2006, FCT) and the Sixth Framework Programme of the European Union - Biosapiens project.