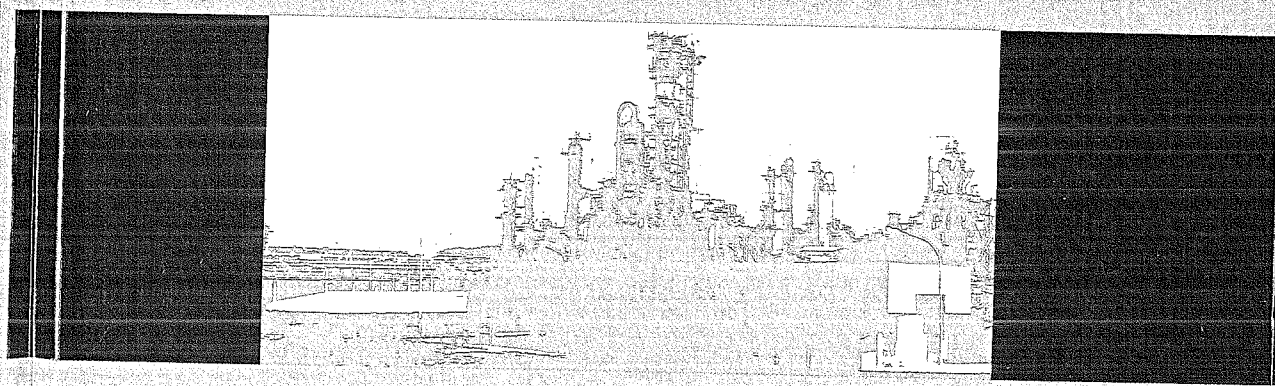
ELSEVIER

10TH INTERNATIONAL SYMPOSIUM ON PROCESS SYSTEMS ENGINEERING

PART B

Edited by
R. M. BRITO ALVES,
C. A. OLLER NASCIMENTO,
E. C. BISCAIA Jr



COMPUTER-AIDED CHEMICAL ENGINEERING, 27

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)

Design of Stable Large-Scale Metabolic Networks

Jimena Di Maggio, Aníbal M. Blanco, J. Alberto Bandoni, M. Soledad Diaz
Planta Piloto de Ingeniería Química (PLAPIQUI),
Universidad Nacional del Sur-CONICET,
Camino La Carrindanga Km 7, Bahía Blanca 8000, Argentina

Abstract

In this work we propose an eigenvalue optimization approach to ensure steady state stability of the Embden-Meyerhof-Parnas pathway, the pentose-phosphate pathway and the phosphotransferase system of *Escherichia coli*. The model consists of eighteen differential equations that represent dynamic mass balances for extracellular glucose and intracellular metabolites and thirty kinetic rate expressions. The nonlinear optimization problem including stability constraints has been solved with reduced space Successive Quadratic Programming techniques within program IPOPT (Waechter and Biegler et al., 2006). Numerical results provide useful insights on the stability properties of the studied kinetic model.

Keywords: Metabolic networks, Stability, Eigenvalue optimization.

1. Introduction

Metabolic networks design can be formulated as an optimization problem aimed at optimizing a given objective, for example the production of a certain metabolite, subject to mass balance equations that represent the network. Kinetic models allow the analysis of stability of the predicted states, which is of fundamental importance because biological systems may exhibit monotonic stable states, bistable switching threshold phenomena, oscillations and chaotic behavior. Due to nonlinear kinetics of the biochemical reactions and their coupling through common metabolites, biological systems may undergo drastic changes in their qualitative behavior when a variation on the enzyme level occurs. If no stability constraints are included in the formulation, the optimal operating point might be unstable, making the metabolic network vulnerable to external disturbances. In other words, in spite of the presence of modest disturbances an unstable network will reach physiological constrains and collapse. Several authors have addressed the analysis if biological systems of small to moderate size (Hatzimanikatis and Bailey, 1997, Haddad and Chellaboina, 2005). The design-for-stability problem, an important sub problem of the general design-for-operability problem, has also motivated many contributions from the process systems engineering community. Different strategies have been proposed to include stability considerations within the design problem (Chang and Sahinidis, 2004; Blanco et al., 2004).

In this work we propose an eigenvalue optimization approach (Blanco and Bandoni, 2007) to ensure steady state stability of the glycolysis, the pentose-phosphate pathway and the phosphotransferase system of *Escherichia coli* K-12 W3110 (Chassagnole et al., 2002). The nonlinear optimization problem, corresponding to steady state equations and stability constraints, has been solved with reduced space Successive Quadratic Programming techniques within program IPOPT (Waechter and Biegler, 2006). Optimization results provide an improved metabolic network for the maximization of

serine production in *Escherichia coli*, within the steady state stable region, based on a detailed kinetic model.

2. Optimization model description

In this work, we have studied an adaptation of the dynamic model for the Embden-Meyerhof-Parnas pathway, the pentose phosphate pathway and the phosphotransferase system, as shown in Fig. 1, of *Escherichia coli* K-12 W3110 (Chassagnole et al., 2002). It comprises eighteen differential equations that represent dynamic mass balances of extracellular glucose and intracellular metabolites, thirty kinetic rate expressions and it involves one hundred and sixteen parameters. Most influential parameters have been determined through previous work on global sensitivity analysis of the proposed model (Di Maggio et al., 2008a,b).

Equations (1) to (6) correspond to main mass balances on metabolites involved in the metabolic network shown in Fig. 1. A detailed description of the remaining twelve balances is given in Chassagnole et al. (2002).

$$\frac{dC_{glc}^{ext}}{dt} = D(C_{glc}^{a\lim} - C_{glc}^{ext}) + f^{pulso} - \frac{C_X r_{PTS}}{\rho_X} \tag{1}$$

$$\frac{dC_{g6p}}{dt} = r_{PTS} - r_{PGI} - r_{G6PDH} - r_{PGM} - \mu C_{g6p} \tag{2}$$

$$\frac{dC_{f6p}}{dt} = r_{PGI} - r_{PFK} + r_{TKb} + r_{TA} - 2r_{MurSynth} - \mu C_{f6p} \tag{3}$$

$$\frac{dC_{gap}}{dt} = r_{ALDO} + r_{TIS} - r_{GAPDH} + r_{TKa} + r_{TKb} - r_{TA} + r_{TrpSynth} - \mu C_{gap} \tag{4}$$

$$\frac{dC_{pep}}{dt} = r_{ENO} - r_{PK} - r_{PTS} - r_{PEPCxylase} - r_{DAHPS} - \mu C_{pep} \tag{5}$$

$$\frac{dC_{pyr}}{dt} = r_{PK} + r_{PTS} - r_{PDH} + r_{MetSynth} + r_{TrpSynth} - \mu C_{pyr} \tag{6}$$

Equations 7 to 13 show kinetic expressions for phosphotransferase system, glucose-6-phosphate isomerase, 6-phosphogluconate dehydrogenase, phosphoglycerate kinase and phosphoglycerate mutase, the Serine synthesis pathway and the Chorismate and Mureine synthesis pathway, respectively. The remaining rate equations involved in the present metabolic network can be found in Chassagnole et al. (2002).

$$r_{PTS} = \frac{r_{PTS}^{\max} c_{glc}^{extracelul\ lar} \frac{c_{pep}}{c_{pyr}}}{\left(K_{PTS,a1} + K_{PTS,a2}\frac{c_{pep}}{c_{pyr}} + K_{PTS,a3} c_{glc}^{extracelul\ lar} + c_{glc}^{extracelul\ lar}\frac{c_{pep}}{c_{pyr}}\right)\left(1 + \frac{c_{g6p}^{N_{PTS,g6p}}}{K_{PTS,g6p}}\right)} \tag{7}$$

$$r_{PGI} = \frac{r_{PGI}^{\max}\left(C_{g6p} - \frac{C_{f6p}}{K_{PGI,eq}}\right)}{K_{PGI,g6p}\left(1 + \frac{C_{f6p}}{K_{PGI,f6p}\left(1 + \frac{C_{6pg}}{K_{PGI,g6p,6pginh}}\right)} + \frac{C_{6pg}}{K_{PGI,g6p,6pginh}}\right) + C_{g6p}} \tag{8}$$

$$r_{PGDH} = \frac{r_{PGDH}^{\max} C_{6pg} C_{nadp}}{\left(C_{6pg} + K_{PGDH,6pg}\right)\left(C_{nadp} + K_{PGDH,nadp}\left(1 + \frac{C_{nadph}}{K_{PGDH,nadph,inh}}\right)\left(1 + \frac{C_{ATP}}{K_{PGDH,ATP,g6pinh}}\right)\right)} \tag{9}$$

$$r_{PGK} = \frac{r_{PGK}^{\max}\left(C_{adp} C_{pgp} - \frac{C_{atp} C_{3pg}}{K_{PGK,eq}}\right)}{\left(K_{PGK,adp}\left(1 + \frac{C_{atp}}{K_{PGK,atp}}\right) + C_{adp}\right)\left(K_{PGK,pgp}\left(1 + \frac{C_{3pg}}{K_{PGK,3pg}}\right) + C_{pgp}\right)} \tag{10}$$

: stable region, based on a

c model for the Embden-
nd the phosphotransferase
Chassagnole et al., 2002).
ynamic mass balances of
tic rate expressions and it
ial parameters have been
sis of the proposed model

etabolites involved in the
of the remaining twelve

(1)

(2)

(3)

(4)

(5)

(6)

erase system, glucose-6-
phoglycerate kinase and
d the Chorismate and
quations involved in the
002).

(7)

(8)

(9)

(10)

$$r_{PGluMu} = \frac{r_{PGluMu}^{max}\left(C_{3pg} - \frac{C_{2pg}}{K_{PGluMu,eq}}\right)}{K_{PGluMu,3pg}\left(1 + \frac{C_{2pg}}{K_{PGluMu,2pg}}\right) + C_{3pg}} \quad (11)$$

$$r_{SerSynth} = \frac{r_{SerSynth}^{max} C_{3pg}}{K_{SerSynth,3pg} + C_{3pg}} \quad (12)$$

$$r_{Synth1} = \frac{r_{Synth1}^{max} C_{pep}}{K_{Synth1,pep} + C_{pep}} \quad (13)$$

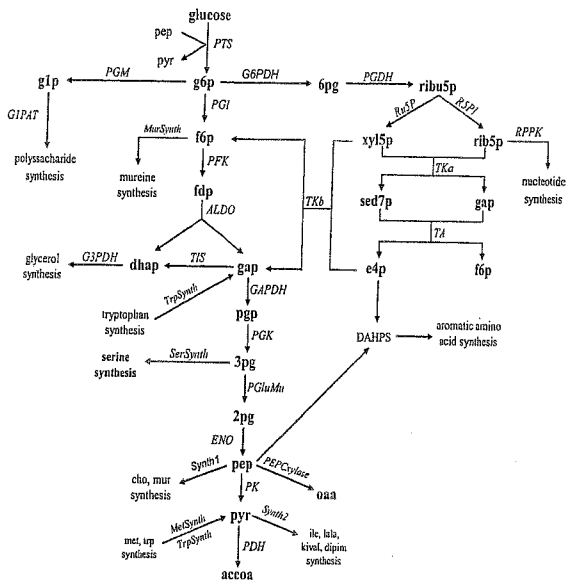


Figure 1. Metabolic network of glycolysis and pentose-phosphate pathway in *Escherichia coli*.

The maximum reaction rates, r_i^{max}, are associated to the corresponding enzyme concentration, so they could be tuned to maximize or minimize the production of a given metabolite. We have formulated an optimization problem in which the maximum reaction rates for certain pathways are the degrees of freedom. As changes in these parameters imply that the levels of enzymes will be modified, the total proteins concentration will vary and physiological changes that are not taken account in the dynamic model could be appear. To account for processes such as redistribution of limited mRNA contents and homeostasis (Mauch et al., 2001, Nikolaev et al., 2005), we have included equations (14) and (15). Eqn. (14) imposes that an increase in certain enzyme levels is compensated by a decrease in the remaining ones. Eqns. (15) ensure that gene expression rates of non-modulated enzymes be maintained at ratios equal to the ones at the reference steady state.

$$\frac{1}{M}\sum_{j=1}^{M}\frac{r_j^{max}}{r_j^{max,0}} = 1 \quad (14)$$

$$\frac{r_{j_1}^{max}}{r_{j_1}^{max,0}} = \ldots = \frac{r_{j_K}^{max}}{r_{j_K}^{max,0}} = \gamma \quad (15)$$

where $r_j^{max,0}$ is the maximum reaction rate at the reference state, M is the number of enzymes in the metabolic network, $j_1 \ldots j_K$ are the indices of non-modulated enzymes,

$K = M-L$ and L is the number of modulated enzymes (corresponding to maximum reaction rates being optimization variables).
Equations (14) and (15) can be re-written as (16) and (17), respectively, to be included in the optimization problem formulation.

$$\frac{r_{j_1}^{\max}}{r_{j_1}^{\max,0}} + \ldots + \frac{r_{j_L}^{\max}}{r_{j_L}^{\max,0}} + K\gamma = M \tag{16}$$

$$r_{j_S}^{\max} = \gamma r_{j_S}^{\max,0} \quad , S = 1,\ldots,K \tag{17}$$

3. Optimization under stability constraints

In order to assess asymptotic stability of dynamic systems, eigenvalue analysis is usually performed. For an asymptotically stable equilibrium point, the eigenvalues of the dynamic system Jacobian matrix lie on the left half of the complex plane. In an eigenvalue optimization problem (Blanco and Bandoni, 2007), the real parts of the eigenvalues of the Jacobian matrix of the dynamic system under study are forced to be strictly negative, ensuring that way asymptotic stability of the resulting equilibrium point. The eigenvalue optimization problem can be stated as follows:

$$\begin{aligned} & \min_{y} \Phi(y) \\ \text{s.t} \quad & A^T(y)P + PA(y) + I = 0 \\ & \det(P_i^{-1}) \geq \xi \qquad i=1,\ldots,n \\ & \xi > 0 \\ & h(y) = 0 \\ & g(y) \leq 0 \\ & y \in Y \end{aligned} \tag{18}$$

where $\boldsymbol{y}$ represents the optimization variables which comprises both, design and operating variables. $\Phi(y)$ is the objective function, $h(y)$ is the set of equality constraints and $g(y)$ is the set of inequality constraints. $A(y)$ is the Jacobian matrix of the dynamic system and $\boldsymbol{P}$ is a real symmetric matrix defined in (18) through the Lyapunov equation. $\mathrm{Det}(\boldsymbol{P}_i^{-1})$ stands for determinants of the principal minors of the inverse of $\boldsymbol{P}$, $\boldsymbol{I}$ is the identity matrix and ξ a user defined positive constant. For details on the derivation and solution strategy of (18) see Blanco and Bandoni, (2007).

4. Discussion of results

In this study, and based on previous work on global sensitivity analysis on the dynamic metabolic network to main parameters (Di Maggio et al., 2008a,b), we have selected four maximum reaction rates as design variables, corresponding to serine synthesis ($r_{SerSynth}^{max}$), glucose-6-phosphate isomerase (r_{PGI}^{max}), phosphotransferase system (r_{PTS}^{max}) and Chorismate and Mureine synthesis (r_{Synth1}^{max}), respectively. The design problem to maximize serine production ($\Phi(y)=r_{SerSynth}$), for the steady state kinetic model of the Embden-Meyerhof-Parnas pathway, the pentose phosphate pathway and the phosphotransferase system comprises mass balances for eighteen metabolites and thirty rate equations associated to thirty enzyme levels (partially shown as Eqns. (7) to (13), as well as twenty seven equations (K+1) standing for Eqns. (16) and (17).
When formulating the design problem under stability constraints, Eqn. (18), additional equality constraints (eighteen) and inequalities (eighteen) are included in the optimization problem formulation, standing for Lyapunov's equation and nonnegativity on the determinants of the principal minors of P, to ensure its positive definiteness,

ponding to maximum

ctively, to be included

(16)

(17)

igenvalue analysis is
nt, the eigenvalues of
complex plane. In an
the real parts of the
study are forced to be
resulting equilibrium
ws:

(18)

es both, design and
of equality constraints
natrix of the dynamic
e Lyapunov equation.
inverse of P, I is the
on the derivation and

alysis on the dynamic
b), we have selected
g to serine synthesis
ferase system (r_{PTS}^{max})
ne design problem to
kinetic model of the
pathway and the
netabolites and thirty
s Eqns. (7) to (13), as
17).
Eqn. (18), additional
re included in the
on and nonnegativity
positive definiteness,

respectively. Nonlinear optimization models have been implemented in a Fortran 90 environment and they have been solved with reduced space Successive Quadratic Programming techniques within program IPOPT (Waechter and Biegler, 2006). The design problem has been first solved without stability constraints (Case 1) and it has been extended to include the proposed equations to ensure a stable optimal network design (Case 2).

Table 1. Main optimization variable values.

Optimization variables	Nominal values	Case 1	Case 2
C_{glcext} (mM)	1,7222	0,9946	0,2599
C_{g6p} (mM)	3,4800	5,1859	4,2877
C_{f6p} (mM)	0,6000	0,8941	0,7393
C_{fdp} (mM)	0.2720	0,2211	0,3280
C_{gap} (mM)	0,2180	0,1818	0,2349
C_{dhap} (mM)	0,1670	0,1485	0,1838
C_{pgp} (mM)	0,0080	0,0035	0,0059
C_{3pg} (mM)	2,1300	0,8866	1,5158
C_{2pg} (mM)	0,3990	0,1656	0,2837
C_{pep} (mM)	2,6700	1,0491	1,7945
C_{pyr} (mM)	2,6700	3,1879	2,9868
C_{6pg} (mM)	0,8080	0,9927	0,8572
C_{ribu5p} (mM)	0,1110	0,1475	0,1385
C_{xyl5p} (mM)	0,1380	0,1713	0,1637
C_{sed7p} (mM)	0,2760	0,5147	0,3624
C_{rib5p} (mM)	0,3980	0,5089	0,4823
C_{e4p} (mM)	0,0980	0,1005	0,1110
C_{g1p} (mM)	0,6530	0,9661	0,7910
r_{PGI}^{max} (mM/sec)	495,870	442,316	495,842
$r_{SerSynth}^{max}$ (mM/sec)	0,0203	0,2000	0,0985
r_{Synth1}^{max} (mM/sec)	0,0148	0.0000	0.0000
Serine Production (mM/sec)	0,0138	0,0939	0,0594

Case 1: Optimization for maximization of Serine production

Numerical results show that serine production could be increased from 0,01381 mM/sec in a reference steady state (experimental) to 0,09396 mM/sec when the maximum serine synthesis reaction rate ($r_{SerSynth}^{max}$) is at its upper bound (=10* $r_{SerSynth}^{max,0}$). Main variable values are shown in the third column of Table 1. The eigenvalues of matrix $A(y)$ show that the metabolic network is stable, being the real part of the largest eigenvalue -3.7E-4. While stable, the system is close to critical stability and small changes in some parameters could easily lead to an unstable equilibrium. Taking into account the difficulty to modulate enzyme levels with precision, this implies that a relative deviation on the enzyme levels from their desired value may produce unstable behavior in the system. Therefore, the inclusion of stability constraints in the design problem becomes necessary.

Case 2: Optimization for maximization of Serine production under stability constraints

In order to modify the spectrum of matrix $A(y)$ different trials with parameter ξ were performed. For ξ= 1E-4, the largest eigenvalue in real part is -1.5E-3. In this case, maximum serine production can be increased 330% from its reference steady state value (0,01381 mM/sec) to 0,05936 mM/sec. It can be noted that some robustness of the

network regarding dynamic stability was achieved at the expense of a worsening in the adopted objective function. Furthermore, serine synthesis reaction rate ($r_{SerSynth}^{max}$) is not at its upper bound, as constraints on the determinants of the inverse of P become active. The fourth column in Table 1 shows main variable values in this case. In both Case 1 and 2, optimal values for the maximization of Serine production implies $r_{Synth1}^{max} = 0$. This corresponds to deletion of Chorismate and Mureine synthesis pathway.

5. Conclusions

The proposed approach in the formulation of a design problem for a large-scale kinetic model for a metabolic network including stability constraints, allows the determination of an improved network for serine production by ensuring its stability at the design level.

6. Acknowledgment

The authors gratefully acknowledge financial support from the National Research Council, Universidad Nacional del Sur and ANPCYT, Argentina.

References

A.M. Blanco and J.A. Bandoni, 2007, Eigenvalue Optimization Based Formulations for Dynamics and Control Problems, Chemical Engineering Process, 46, 1192-1199

A.M. Blanco, J.A. Bandoni, L. Biegler, 2004, Re-design of the Tennessee Eastman Challenge Process: An Eigenvalue Optimization Approach, Proceedings FOCAPD 2004, Princeton, 517-520

Y. Chang, N. Sahinidis, 2004, Optimization of metabolic pathways under stability considerations, Computers and Chemical Engineering, 29, 467-479

C. Chassagnole, N. Noisommit-Rizzi, J. Schmid, K. Mauch, M. Reuss, 2002, Dynamic Modeling of the Central Carbon Metabolism of *Escherichia coli*, Biotechnology and Bioengineering, 79, 1, 53-73

J. Di Maggio, J.C. Diaz Ricci, M.S. Diaz, 2008a, Global Sensitivity Analysis: Estimation of Sensitivity Indices in Metabolic Network Dynamic Models, submitted to Metabolic Engineering.

J. Di Maggio, J.C. Diaz Ricci, M.S. Diaz, 2008b, Ranking of most influential kinetic parameters in metabolic networks through Global Sensitivity Analysis, Engineering Conferences International-Metabolic Engineering VII, September 14-19, Puerto Vallarta, Mexico

W. Haddad and V. Chellaboina, 2005, Stability and dissipativity theory for nonnegative dynamical systems: a unified analysis framework for biological and physiological systems, Nonlinear Analysis-Real World Applications, 6, 33-65

V. Hatzimanikatis and J. Bailey, 1997, Studies on glycolysis.1.Multiple steady states in bacterial glycolysis, Chemical Engineering Science, 52, 2579-2588

K. Mauch, S. Buziol, J. Schmid, M. Reuss, 2001, Computer Aidded Design of Metabolic Networks, Chemical Process Control-6 Conference, Tucson, Arizona

E. Nikolaev, P. Pharkya, C. Maranas, A Armaou, 2005, Optimal selection of enzyme levels using large-scale kinetic models, Proceedings of 16th I.F.A.C. World Congress, Prague, Czech Republic

A. Waechter, L. Biegler, 2006, On the Implementation of an Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming, Mathematical Programming, 106, 1, 25-57