Universidade de Lisboa

Faculdade de Ciências

Departamento de Química e Bioquímica

Model discrimination in time-course kinetics: the glyoxalase pathway in *S. cerevisiae*

Nuno Filipe Gonçalves das Lages

Doutoramento em Bioquímica

(Especialidade: Bioquímica Teórica)

2010

Universidade de Lisboa

Faculdade de Ciências

Departamento de Química e Bioquímica

Model discrimination in time-course kinetics: the glyoxalase pathway in *S. cerevisiae*

Nuno Filipe Gonçalves das Lages

Doutoramento em Bioquímica

(Especialidade: Bioquímica Teórica)

Tese orientada pelos Professores Doutores

António Eduardo do Nascimento Ferreira

Carlos Alberto Alves Cordeiro

2010

De acordo com o disposto no artigo nº 40 do Regulamento de Estudos Pós-graduados da Universidade de Lisboa, Deliberação nº 961/2003, publicada no Diário da República – II Série nº 153 – 5 de Julho de 2003, foram incluídos nesta dissertação os resultados do seguinte artigo:

Hou J, Lages NF, Oldiges M, Vemuri GN. 2009. Metabolic impact of redox cofactor perturbations in *Saccharomyces cerevisiae*. *Metabolic Engineering*. **11**: 253-261.

No cumprimento do disposto na referida deliberação, esclarece-se serem da minha responsabilidade a execução das experiências que estiveram na base dos resultados apresentados (excepto quando referido em contrário), assim como a interpretação e discussão dos mesmos.

Agradecimentos/Acknowledgements

Para a realização deste trabalho contribuíram muitas pessoas de diversas formas às quais quero expressar os meus sinceros agradecimentos.

O primeiro agradecimento é para os meus orientadores, António e Carlos, que me acolheram, ensinaram e incentivaram desde o estágio da licenciatura. Para o António, que admiro desde a licenciatura pela incrível variedade de temas que consegue ensinar bem (só mesmo quem sabe), por tudo o que me ensinou sobre modelação, programação e não só, e pelo seu minucioso sentido crítico. Para o Carlos, por me levar para o laboratório e me motivar para o trabalho experimental; pela tranquilidade e confiança que proporciona a quem com ele trabalha e pela capacidade de apontar outras direcções. Só com os ensinamentos de ambos foi possível fazer este trabalho. A ambos, por todas as conversas científicas e não científicas e pela amizade, obrigado.

À Professora Ana agradeço a enorme generosidade com que partilha a sua imensa experiência académica e de vida com quem, como eu, tem o privilégio de trabalhar no seu Grupo de Enzimologia. Pelo optimismo, pela motivação, pelos ensinamentos e pela amizade, muito obrigado.

Estes anos não teriam sido como foram sem a amizade, a boa disposição e o espírito de entreajuda dos colegas do Grupo de Enzimologia. Um obrigado sentido para o Luís, um grande amigo de todas as horas; para o Ricardo Gomes, pelas muitas conversas sobre os mais diversos assuntos pelos quais partilhamos interesse; para o Hugo, um amigo e um colega sempre disponível; para a Marta pelas discussões bem-dispostas e pela amizade; para a Lídia pela alegria com que contagia todos no laboratório; para o Gonçalo pelo companheirismo e pelo entusiasmo. Ao Ricardo Pais e à Patrícia pela excelente companhia no nosso laboratório de Bioquímica Computacional, Ao Francisco e ao Alexandre pelas conversas interessantes que tivemos. Obrigado a todos por tornarem todos os dias de trabalho bons dias.

Quero também agradecer ao Goutham por me ter convidado para colaborar consigo na Dinamarca e pela orientação na parte do trabalho que aí foi realizada. Ensinou-me imenso e foi de uma disponibilidade irrepreensível. À incansável Jin, pela colaboração

frutuosa e pelo contagiante entusiasmo. À Professora Lisbeth Olsson por ter tornado a estadia possível e pelas discussões proveitosas. E não posso concluir o parágrafo sem uma palavra de apreço para a Elena, o Yury, o Giovanni e o Martin, sem os quais a estadia na Dinamarca não teria sido a mesma coisa.

Agradeço à Fundação para a Ciência e Tecnologia – Ministério da Ciência, Tecnologia e Ensino Superior a atribuição de uma bolsa de doutoramento (SFRH / BD / 21947 / 2005).

Fica também uma palavra para os grandes Tiago e Filipe, companheiros desde o princípio nesta vida de bioquímico, cuja amizade e apoio foram também imprescindíveis. A ambos um especial agradecimento pela partilha de experiências e reflexões ao longo dos anos.

Agradeço aos meus pais, Dionil e Adelina, e à minha irmã Ana Luísa pela enorme paciência e pelo apoio em todos os momentos que me permitiram continuar sempre a lutar por mais. À minha tia Lurdes pela sempre presente amizade, preocupação e incentivo. Aos meus tios, Palmira e José Manuel, e ao meu primo Pedro, que estão sempre a torcer por mim. E finalmente, aos meus avós, António, Francisca, Isaura e José, a quem dedico este trabalho, pelas referências que têm sido e por tudo o que me têm dado.

A todos o meu obrigado.

Table of contents

Abstract

The present work addresses the problem of model discrimination in enzyme kinetics. Frequently, more than one kinetic model is considered during the characterization of an enzymatic reaction or a metabolic pathway. The statistical selection of a model may be difficult if the candidate models fit the experimental data with very similar fitting scores. Since each model corresponds to a different possible mechanism of the studied process, model selection also reflects the choice of a particular mechanism. In addition, predictions given by models with equal fitting scores may be different.

The glyoxalase system is a metabolic pathway that has been studied using two alternative kinetic models. These models could not be discriminated despite extensive kinetic experiments and an alternative branched mechanism combining the two models has been proposed. This pathway is therefore ideal for model discrimination research in Biochemistry. The glyoxalase pathway comprises the enzymes glyoxalase I and glyoxalase II. Glyoxalase I catalyzes the isomerization of the hemithioacetal that forms from the condensation of methylglyoxal (a by-product of glycolysis) and glutathione to S-D-lactoylglutathione. Glyoxalase II catalyzes the hydrolysis of S-D-glutathione to D-lactate and glutathione. The methylglyoxal-glutathione hemithioacetal forms spontaneously without the presence of enzymes. Therefore the glyoxalase I reaction can be described either as irreversible single-substrate or as irreversible two-substrate, considering that the hemithioacetal forms before binding the enzyme or that it forms in the active centre of the enzyme after sequential binding of glutathione and methylglyoxal, respectively. The glyoxalase system is the most important catabolic pathway for methylglyoxal. Methylglyoxal is a toxic agent due to its ability to react with proteins and nucleic acid amine groups that leads to formation of advanced glycation end-products. Therefore the glyoxalase pathway was suggested to be a potential dug target for its cellular defensive role against methylglyoxal.

An introduction to the subjects developed through this dissertation is given in chapter 1, covering the state of art of research on the glyoxalase pathway and methylglyoxal metabolism and on relevant mathematical and computational methods for model analysis and discrimination.

In chapter 2 the glyoxalase system is investigated by analyzing the algebraic solutions of the rate equations describing the pathway at steady state. The two mentioned glyoxalase I kinetic models were used in this approach. It is observed that for the existence of a steady state a minimum amount of glutathione must be available; in addition, glyoxalase I and II activities must exceed thresholds higher than the flux of the pathway. It is shown that methylglyoxal steady-state concentration is not sensitive to variations of glyoxalase II activity but varies significantly with total glutathione concentration and methylglyoxal formation rate. Sensitivity to glyoxalase I activity depends on the kinetic model describing the enzyme: highly sensitivity if the two-substrate model is used but not so for the one-substrate model. The pathway seems to operate very far from the conditions of disruption of the physiological steady state to assure a very low methylglyoxal concentration and a fast regeneration and high concentration of free glutathione.

Time-course kinetic studies with purified yeast enzyme and yeast permeabilized cells are described in chapter 3. Akaike's information criterion and residual analysis are used to discuss the selection of the most appropriate kinetic model for glyoxalase I. Parameter least-square estimates for this study are obtained with a combination of the stochastic Differential Evolution with the deterministic downhill-simplex optimization algorithms. Although the two-substrate model performs slightly better for the purified enzyme data, the Akaike score differences for both data sets and the residual analysis for the permeabilized cell data are not conclusive.

A method developed to design optimized experimental conditions for model discrimination is explained in chapter 4. The method employs a multiobjective optimization algorithm (the Generalized Differential Evolution, generation 3) to search for the experimental conditions that maximize the divergence between the reaction time courses predicted by the models. The Kullback-Leibler distance is the measure of divergence employed. The combination of the chosen algorithm and divergence criterion is successful in finding solutions that result in very different predictions from the two models for glyoxalase I in the presence of glyoxalase II, proving to be useful for planning model discrimination experiments.

The importance of keeping a high free glutathione concentration seems to establish the properties of the glyoxalase pathway identified in chapter 2. Glutathione is also a key

antioxidant and its oxidized form is reduced through the glutathione reductase system at the expense of NADPH. Indeed, the pyridine nucleotides NADPH and NADH have crucial metabolic roles. NADH, formed mainly in catabolic reactions, is the substrate of the respiratory chain and therefore it ultimately supplies the synthesis of ATP. NADPH is the main reducing agent in biosynthetic pathways. In addition, the pyridine nucleotides are among the metabolites that participate in a larger number of reactions in the cell. Therefore it is important to understand the effects of concentration changes of these metabolites. In chapter 5 perturbations to pyridine nucleotide concentrations are studied in living yeast cells cultured in bioreactors. The results for five recombinant *S. cerevisiae* strains overexpressing a cytosolic NADH oxidase, a mitochondrial NADH oxidase, a cytosolic NADH kinase, a mitochondrial NADH kinase and a cytosolic soluble pyridine nucleotide transhydrogenase are discussed. Extracellular and intracellular metabolite measurements and a stoichiometric model are used to assess the consequences of such perturbations, unveiling how metabolism in intact cells adapts to different redox conditions. Strains with enhanced NADH oxidation in the cytosol show a lower glycerol production. On the other hand enhanced NADH consumption in the mitochondrion lowers ethanol production and enhances ATP synthesis efficiency.

The results presented here show that different kinetic models may fit experimental data equally well, making the selection of one model extremely. An original contribution is established to aid planning experiments for model discrimination. In addition, a broad characterization of the effects of perturbations to pyridine nucleotide metabolism is given, which is valuable to understand the complex response of yeast's metabolic network, with direct biotechnological application.

Keywords: glyoxalase pathway, time-course kinetics, model discrimination, redox metabolism, *Saccharomyces cerevisiae*

Resumo

Este trabalho aborda o problema da discriminação de modelos em cinética enzimática. Frequentemente mais do que um modelo cinético é considerado para descrever uma reacção enzimática ou uma via metabólica. A selecção estatística de um modelo pode ser difícil se o ajuste dos modelos candidatos aos resultados experimentais for muito semelhante. Como a cada modelo corresponde um mecanismo diferente do processo estudado, a selecção de um modelo reflecte também a escolha de um mecanismo particular. Além disso, as previsões dadas por modelos com ajustes semelhantes podem ser muito diferentes. Consequentemente, o estudo de um sistema bioquímico usando modelos diferentes pode conduzir a conclusões diferentes.

O sistema dos glioxalases é uma via metabólica que tem sido estudada com recurso a dois modelos cinéticos alternativos. Estes modelos não foram discriminados apesar de o sistema ter sido alvo de estudos cinéticos detalhados, tendo sido proposto um mecanismo ramificado que combina os dois modelos. Esta via é portanto ideal para investigação em discriminação de modelos em Bioquímica. A via dos glioxalases é constituída pelos enzimas glioxalase I e glioxalase II. O glioxalase I catalisa a isomerização do hemitioacetal formado por condensação de metilglioxal (um produto secundário do metabolismo glicolítico) com glutationo a S-D-lactoilglutationo. O glioxalase II catalisa a hidrólise do S-D-lactoilglutationo a D-lactato e glutationo. O hemitioacetal de glutationo e metilglioxal forma-se espontaneamente na ausência de enzimas. Por isso, a reacção do glioxalase I pode ser descrita como uma reacção a um substrato ou uma reacção a dois substratos, se se considerar que o hemitioacetal se forma antes de se ligar ao enzima ou que se forma no centro activo do enzima após ligação sequencial do glutationo e do metilglioxal, respectivamente. O sistema dos glioxalases é a via de catabolismo mais importante para o metilglioxal, um metabolito tóxico devido à capacidade de reagir com os grupos amina de proteínas e aminoácidos, levando à formação de produtos avançados de glicação. Por isso foi sugerido que o sistema dos glioxalases poderia ser um alvo terapêutico pelo seu papel na defesa celular contra o metilglioxal.

Os assuntos desenvolvidos ao longo da tese são introduzidos no capítulo 1, no qual se descreve o estado do conhecimento sobre a via dos glioxalases e o metilglioxal e sobre

métodos matemáticos e computacionais relevantes para a análise e discriminação de modelos.

No capítulo 2 o sistema dos glioxalases é investigado analisando as soluções algébricas das equações de velocidade que descrevem a via em estado estacionário. Ambos os modelos mencionados para o glioxalase I são usados nesta abordagem. Para que se estabeleça um estado estacionário tem de estar disponível uma quantidade mínima de glutationo; por outro lado, as actividades do glioxalase I e do glioxalase II têm de exceder limiares superiores ao fluxo da via. Mostra-se que a concentração de metilglioxal em estado estacionário não é sensível a variações de actividade do glioxalase II mas varia significativamente com a concentração total de glutationo e a velocidade de formação de metilglioxal. A sensibilidade relativamente a variações da actividade do glioxalase I depende do modelo cinético utilizado para descrever o enzima: a sensibilidade é elevada quando se considera o modelo de dois substratos mas não quando se considera o modelo de um substrato. A via funciona em condições muito afastadas da ruptura do estado estacionário para assegurar uma concentração muito reduzida de metilglioxal, a regeneração rápida do glutationo e uma elevada concentração de glutationo livre.

No capítulo 3 são descritos estudos cinéticos de curvas de progressão com enzimas purificados e células permeabilizadas de levedura. São utilizados o critério de informação de Akaike e análise de resíduos para discutir a selecção do modelo mais apropriado para o glioxalase I. As estimativas dos parâmetros para este estudo são obtidas com uma combinação do algoritmo estocástico de optimização Evolução Diferencial com o algoritmo de optimização determinístico *downhill-simplex*, usando o critério dos mínimos quadrados. Embora a análise de resíduos das curvas de progressão obtidas com enzima purificado seja favorável ao modelo de dois substratos, o critério de informação de Akaike (aplicado às curvas de progressão obtidas com enzimas purificados e células permeabilizadas) e a análise de resíduos das curvas de progressão obtidas com células permeabilizadas não são conclusivos.

No capítulo 4 é apresentado um método desenvolvido para planear experiências optimizadas para discriminação de modelos. O método usa um algoritmo de optimização multiobjectivo (a Evolução Diferencial Generalizada, geração 3) para encontrar as condições experimentais que maximizam a divergência entre as curvas de

progressão previstas pelos modelos. A medida de divergência utilizada é a distância de Kullback-Leibler. A combinação do algoritmo e do critério de divergência escolhida é capaz de encontrar soluções que resultam em previsões muito diferentes dos dois modelos para o glioxalase I em presença do glioxalase II, provando ser úteis no planeamento de experiências para discriminação de modelos.

A importância da manutenção de uma concentração de glutationo livre elevada parece estar na base das propriedades da via dos glioxalases identificadas no capítulo 2. O glutationo é também um antioxidante chave e a sua forma oxidada é reduzida pelo sistema glutationo redutase, que depende de NADPH. De facto, os nucleótidos de piridina NADH e NADPH desempenham papéis muito importantes no metabolismo. O NADH está envolvido sobretudo em vias de catabolismo e é o substrato da cadeia respiratória, sendo em última análise o fornecedor da energia necessária para a síntese de ATP. O NADPH é o principal agente redutor em vias biossintéticas. Os nucleótidos de piridina estão ainda entre os metabolitos que participam em maior número de reacções na célula. Portanto, é importante compreender os efeitos de variações das concentrações destes metabolitos. No capítulo 5 são estudadas perturbações às concentrações dos nucleótidos de piridina em células de levedura cultivadas em biorreactor. São discutidos os resultados para cinco estirpes recombinantes de *S. cerevisiae* que sobreexpressam um NADH oxidase citosólico, um NADH oxidase mitocondrial, um NADH cinase citosólico, um NADH cinase mitocondrial ou um trans-hidrogenase de nucleótidos de piridina. As consequências destas perturbações são avaliadas por medição de metabolitos intra- e extracelulares e através de um modelo estequimétrico, revelando como o metabolismo em células intactas se adapta a diferentes condições redox. As estirpes com oxidação de NADH aumentada no citosol apresentam maior produção de glicerol. Por outro lado estirpes com maior consumo de NADH no mitocôndrio apresentam menor produção de etanol e maior eficiência na síntese de ATP.

Os resultados aqui apresentados mostram que modelos cinéticos diferentes podem resultar em bons ajustes aos mesmos dados, tornando a selecção de um dos modelos extremamente difícil. É dada uma contribuição importante para o planeamento de experiências com vista à discriminação de modelos. Foi também feita uma caracterização ampla dos efeitos de perturbações aos níveis dos nucleótidos de piridina,

a qual é valiosa para compreender a resposta complexa do metabolismo da levedura, com aplicação biotecnológica directa.

List of symbols and abbreviations

2PG	2-phosphoglycerate
3PG	3-phosphoglycerate
3PHP	3-phosphohydroxypyruvate
3PS	3-phosphoserine
ADP	Adenosine triphosphate
AGE	Advanced glycation end-product
AIC	Akaike's information criterion
AKG	α-ketoglutarate
AMP	Adenosine monophosphate
AOX	Yeast strain overexpressing the *AOX1* gene
AOX1	Gene encoding an alternative mitochondrial NADH oxidase
ATP	Adenosine triphosphate
BST	Biochemical Systems Theory
C-bal	Carbon balance
Cit	Citrate
CoA	Coenzyme A
CR	Crossover probability in Differential Evolution
CTP	Cytidine triphosphate
D	Dilution rate
DE	Differential Evolution
DHIV	Dihydroxyisovalerate
DHAP	Dohydroxyacetone phosphate
E4P	Eritrose 4-phosphate
EMP	Embden-Meyerhof-Parnas pathway
F6P	Fructose 6-phosphate
G3P	Glyceraldehyde 3-phosphate
G6P	Glucose 6-phosphate
GAP	Glyceraldehyde phosphate
G_{max}	Maximum number of generations in Differential Evolution
GMP	Guanosine monophosphate
GSH	Glutathione (reduced) steady-state concentration
GTP	Guanosine triphosphate
HTA	methylglyoxal-glutathione hemithioacetal steady-state concentration
Icit	Isocitrate
$I(f, g)$	Kullback-Leibler distance of model g from model f
IMP	Inositol monophosphate
k_1	Hemithioacetal formation kinetic constant
k_{-1}	Hemithioacetal dissociation kinetic constant
k_{cat1}	Catalytic constant for glyoxalase I
k_{cat2}	Catalytic constant for glyoxalase II
K_{m1}	Michaelis constant of glyoxalase I for the methylglyoxal-glutathione hemithioacetal
K_{m11}	Michaelis constant of glyoxalase I for glutathione
K_{m12}	Michaelis constant of glyoxalase I for methylglyoxal
K_{m2A}	Michaelis constant of glyoxalase II in model 1
K_{m2B}	Michaelis constant of glyoxalase II in model 3
FBP	Fructose bisphosphate
KIle	Ketoisoleucine
KIV	Ketoisovalerate
MAL8	Gene encoding maltase
MATa	Yeast mating type a
mATP	Maintenance on ATP
MCA	Metabolic control analysis
MGO	Methylglyoxal (2-ketopropanal) steady-state concentration
GDE	Generalized Differential Evolution

NAD^+	Nicotinamide and adenine dinucleotide (oxidized)
NADH	Nicotinamide and adenine dinucleotide (reduced)
$NADP^+$	Nicotinamide and adenine dinucleotide phosphate (oxidized)
NADPH	Nicotinamide and adenine dinucleotide phosphate(reduced)
NOX	Yeast strain overexpressing the *nox* gene
NP	Population number in Differential Evolution
OAS	O-acetilserine
ODE	Ordinary differential equation
PEP	Phosphoenolpyruvate
pntAB	Gene encoding a membrane-bound pyridine nucleotide transhydrogenase
PO	ATP synthesized per atom of oxygen used
PP_i	Pyrophosphate
R5P	Ribose 5-phosphate
r_{ATP}	Specific ATP formation rate
R-bal	Redox balance
r_{CO_2}	Specific carbon dioxide production rate
r_{eth}	Specific ethanol production rate
r_{glu}	Specific glucose uptake rate
r_{NADH}	Specific NADH formation rate
r_{NADPH}	Specific NADPH formation rate
r_{O_2}	Specific oxygen uptake rate
RSS	Residual sum of squares
Ru5P	Ribulose 5-phosphate
S7P	Sedoheptulose 7-phosphate
SDLG	S-D-lactoylglutathione steady-state concentration
SH_{Total}	Total glutathione (glutathionyl moiety)
SUC2	Gene encoding invertase (beta-D-fructofuranoside fructohydrolase EC. 3.2.1.26)
TCA	Tricarboxylic acid (cycle)
TTP	Thymidine triphosphate
udhA	Gene encoding a soluble pyridine nucleotide transhydrogenase
Ura3-52	Mutated URA3 gene (which expresses orotidine 5'-phosphate decarboxylase)
UTP	Uridine triphosphate
t	time
v_1	Glyoxalase I reaction rate
v_2	Glyoxalase II reaction rate
V_{1A}	Activity of glyoxalase I in model 1
V_{1B}	Activity of glyoxalase I in model 2
V_{2A}	Activity of glyoxalase II in model 1
V_{2B}	Activity of glyoxalase II in model 2
v_f	Methylglyoxal-glutathione hemithioacetal formation rate
v_r	Methylglyoxal-glutathione hemithioacetal dissociation rate
v_{in}	Methylglyoxal formation rate
Xy5P	Xylulose 5-phosphate
Y_{ox}	Yield of biomass per unit of oxygen
Y_{sx}	Yield of biomass per unit of glucose
Y_{xATP}	ATP required per unit of biomass
Y_{sx}^{true}	True yield coefficient of biomass per unit of glucose
Y_{ox}^{true}	True yield coefficient of biomass per unit of oxygen
μ	Specific growth rate

1 Introduction

1.1 The glyoxalase pathway

The ubiquitous glyoxalase pathway had the attention of many biochemists since its discovery in the early 20th century (Dakin and Dudley 1913a; Neuberg 1913). Despite the extensive study of these enzymes, the actual functional relevance of the pathway is still not clear. The pathway was independently discovered in 1913 by Neuberg (Neuberg 1913) and Dakin and Dudley (Dakin and Dudley 1913a), who concluded that it catalyzed the conversion of 2-ketoaldehydes to 2-hydroxycarboxylic acids, among which methylglyoxal (2-ketopropanal), and named it glyoxalase (Dakin and Dudley 1913b).

Neuberg suggested that methylglyoxal and the glyoxalase pathway were part of glycolysis (Neuberg and Kobel 1934). In 1932 reduced glutathione, which was known not to be necessary for glycoslysis, was found to be an essential cofactor of the pathway (Lohmann 1932) and in 1951 D-lactate was identified as the final product, instead of L-lactate, which resulted from muscle glycolytic metabolism (Racker 1951). These two discoveries led to the rejection of the theory that the glyoxalase system was part of glycolysis. Racker also showed in the same work that the pathway consisted not of one enzyme but two, which he named glyoxalase I [lactoylglutathione lyase, EC 4.4.1.5] and glyoxalase II [hydroxyacylglutathione hydrolase, EC 3.1.2.6] (Racker 1951).

Curiosity about the biochemical role of the pathway didn't fade out mostly due to its presence in diverse organisms (Hopkins and Morgan 1945) and it eventually was acknowledged to be nearly ubiquitous (Thornalley 1993). Also the publication of the retine-promine theory by Szent-Gyorgyi, which stated that the glyoxalase pathway and methylglyoxal had a role in cell division and cancer development (Szent-Gyorgyi 1965), brought the attention of more researchers to the subject and as a consequence many important insights were achieved about the metabolism and reactivity of methylglyoxal (Thornalley 2008). However, although dividing cells show a higher glyoxalase activity (Principato *et al.* 1982), a clear evidence linking the two events was

never found and studies with yeast deletion mutants showed that lack of the glyoxalase system is not lethal and doesn't halt cell proliferation (Inoue and Kimura 1996; Bito *et al.* 1997; Ponces Freire *et al.* 2003). A major boost in the interest on methylglyoxal metabolism in general also came from the discovery of its involvement in pathogenic events in several diseases, namely diabetes mellitus, atherosclerosis and, later, some neurodegenerative disorders (Thornalley 1993; 1996).

1.1.1 Biochemistry of methylglyoxal

1.1.1.1 Methylglyoxal metabolism

According to current knowledge, methylglyoxal is the main substrate of the glyoxalase pathway (Mannervik 2008) and this is, in yeast, along with aldose reductase [alditol:NAD(P)$^+$ 1-oxidoreductase, EC.1.1.1.21.] (Vander Jagt *et al.* 1992), its most important catabolic route (Thornalley 2003; Gomes *et al.* 2005a).

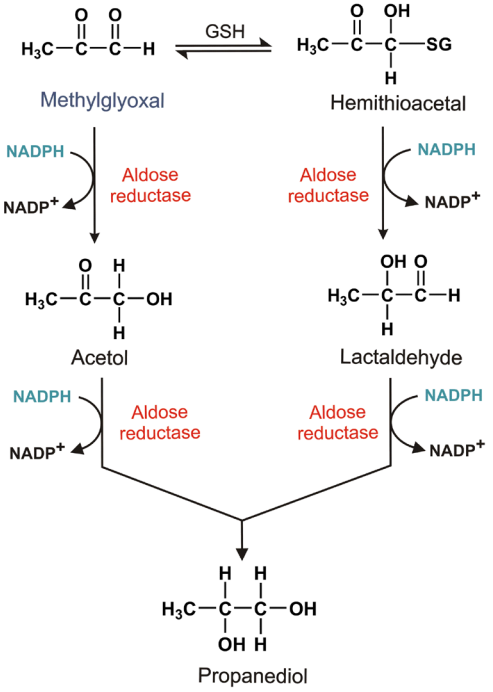


Figure 1.1 Aldose reductase-catalyzed catabolism of methylglyoxal. The enzyme catalyzes directly the reduction of methylglyoxal or of its GSH hemithioacetal using NADPH as reducing agent (Gomes 2007).

Methylglyoxal is extensively hydrated in aqueous solution and it is not known whether glyoxalase I is able to catalyze its dehydration, necessary for the formation of the hemithioacetal (Mannervik 2008). A few other enzymes can however catabolize methylglyoxal: α-oxoaldehyde dehydrogenase [2-oxoaldehyde:NAD(P)$^+$ 2-oxidoreductase, EC. 1.2.1.23] (Monder 1967), aldehyde dehydrogenase [aldehyde:NAD$^+$ oxidoreductase, EC. 1.2.1.3] (Izaguirre *et al.* 1998), aldose reductase [alditol:NAD(P)$^+$ 1-oxidoreductase, EC. 1.1.1.21], methylglyoxal reductase [D-lactaldehyde:NAD$^+$ oxidoreductase, EC. 1.1.1.78] (Ray and Ray 1984) (figure 1.1) and pyruvate dehydrogenase [pyruvate: dihydrolipoyllysine-residue acetyltransferase-lipoyllysine 2-oxidoreductase (decarboxylating, acceptor-acetylating), EC. 1.2.4.1)] (Baggetto and Lehninger 1987).

Dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (G3P), intermediates of glycolysis (Meyerhof and Lohmann 1934; Richard 1985; 1991; 1993b), exist in equilibrium with a phosphorylated 1,2-enediolic intermediate. The elimination of the phosphoryl group of this intermediate has a low energy barrier and leads to methylglyoxal formation (Richard 1993b). This is believed to be the most important source of methylglyoxal formation in eukaryotic cells (Thornalley 2008). The triose phosphate isomerase enzyme catalyzes the interconversion of DHAP and G3P through this intermediate, being very efficient in stabilizing it avoiding methylglyoxal formation (Richard 1993b). Nevertheless some of the intermediate leaks from the enzyme active centre and methylglyoxal is formed (Richard 1991), although at a much lower rate than the isomerization of the triose phosphates (Webb *et al.* 1977; Iyengar and Rose 1981) – figure 1.2.

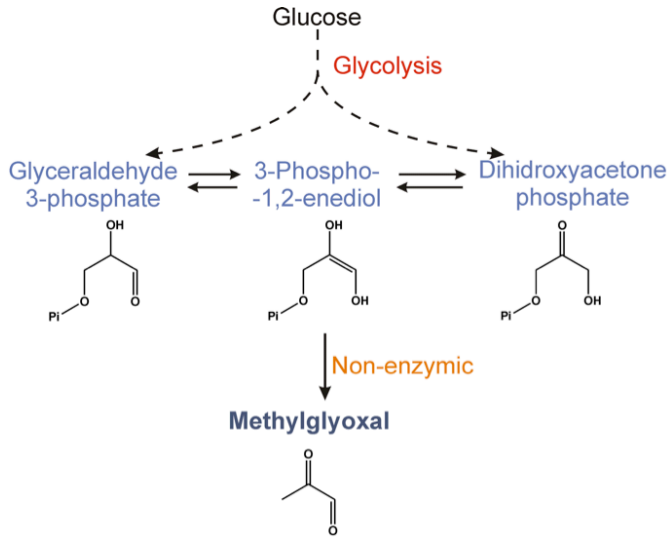


Figure1.2 Main methylglyoxal formation reaction: spontaneous elimination of the phosphate group of the enediolic intermediate in equilibrium with the glycolytic triose phosphates.

Though methylglyoxal formation rate depends on the type of cell and the physiological conditions (Thornalley 2008), it is correlated with the glycolytic flux, supporting the importance of the contribution of the triose phosphate isomerase reaction (Gomes 2007) – in yeast for example, which has a high glycolytic activity, methylglyoxal represents about 0.3% of the observed flux (Martins *et al.* 2001a; Gomes 2007).

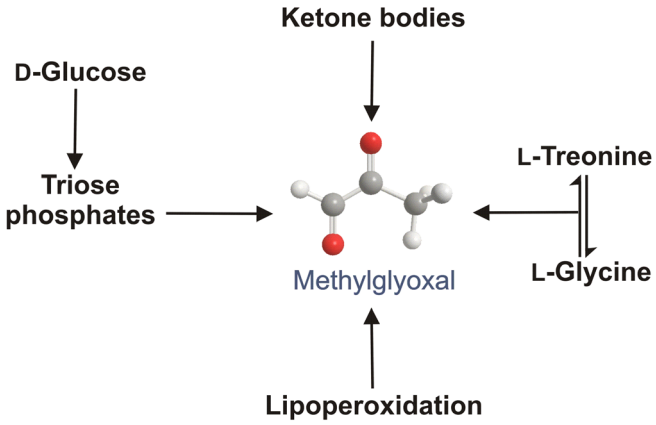


Figure 1.3 Known methylglyoxal formation pathways (Gomes 2007).

Some bacteria have an enzyme that specifically catalyzes the formation of methylglyoxal, methylglyoxal synthase [glycerone-phosphate phosphate-lyase (methylglyoxal-forming), EC.4.2.3.3] (Cooper and Anderson 1970; Hopper and Cooper 1971). In the case of these organisms this is the major methylglyoxal formation source.

Methylglyoxal also forms in other reactions (figure 1.3), mainly in pathological conditions related to diabetes mellitus, where ketoacidosis occurs due to glucose deprivation (Turk *et al.* 2006). One of the main catabolic pathways of L-threonine is its conversion to glycine and acetyl-CoA at the expense of a free CoA molecule (Lages *et al.* 2008). In situations where CoA level is low (as in ketoacidosis, where most CoA is in the form of acetyl-CoA) aminoacetone and carbon dioxide are formed instead of glycine and acetyl-CoA. Aminoacetone is then converted to methylglyoxal by the semicarbazide-sensitive amine oxidase enzyme [amine:oxygen oxidoreductase, EC.1.4.3.6.] (Tressel *et al.* 1986; Lyles and Chalmers 1992). Other reactions of formation of methylglyoxal are the enzymatic oxidation of acetoacetate by

myeloperoxidase (hydrogen-peroxide oxidoreductase, EC. 1.11.1.7) (Aleksandrovskii 1992) and oxidation of acetone by cytochrome P450 IIE1 [reduced-flavoprotein:oxygen oxidoreductase, EC 1.14.14.1] (Casazza *et al.* 1984; Koop and Casazza 1985).

1.1.1.2 Methylglyoxal and glycation in pathological states

After being discarded as a likely intermediate of glycolysis and later as a key player in the process of cell division, methylglyoxal became interesting as a glycation agent. Glycation is the chemical non-enzymatic process of reaction between carbonyl and amine groups of biological molecules such as proteins, nucleic acids and aminated phospholipids, eventually yielding stable advanced glycation end-products (AGEs) (Bucala and Cerami 1992; Vlassara *et al.* 1994; Westwood and Thornalley 1997). Interestingly, methylglyoxal is both an aldehyde and a ketone, having the carbonyls in adjacent carbon atoms. Interaction between the two carbonyls causes the aldehyde C=O bond to be more polarized and therefore the corresponding carbon atom to be more electrophilic and reactive (Szent-Gyorgyi 1965; Abdulnur 1976; Jencks 1987; Ventura and Cubas 1992).

Nucleic acid glycation by methylglyoxal is mutagenic – methylglyoxal reacts mainly with guanosine residues (Shapiro and Hachmann 1966; Shapiro *et al.* 1969; Krymkiewicz 1973). The reaction is faster with RNA and single-stranded DNA (Krymkiewicz 1973) but native methylglyoxal-glycated DNA was already found *in vivo* as well (Bucala *et al.* 1984; Seidel and Pischetsrieder 1998; Schneider *et al.* 2006). There is evidence that methylglyoxal may cause interstrand cross-linking (Rahman *et al.* 1990), strand breaks (Rahman *et al.* 1990; Pischetsrieder *et al.* 1999) and increased mutation rate (Cajelli *et al.* 1987; Migliore *et al.* 1990; Lee *et al.* 1995; Pischetsrieder *et al.* 1999).

Methylglyoxal has been shown to react in physiological conditions with aminated protein amino acid residues, namely arginine and lysine leading to the irreversible modification of their side chains (figure 1.4). Since the N-terminal of proteins is usually modified, it is not considered a main target of glycation (Driessen *et al.* 1985). The

reaction of methylglyoxal and lysine yields N^{ε}-carboxyethyl-lysine (CEL) (Ahmed *et al.* 1997) or protein cross-links in the form of methylglyoxal-lysine dimmers (Sell and Monnier 1989; Fu *et al.* 1994; Nagaraj *et al.* 1996). Arginine residues appear glycated in a larger extent than lysine (Lo *et al.* 1994; Oya *et al.* 1999). Similarly to lysine, a methylglyoxal-induced lysine-arginine cross-link has also been identified (Lederer and Klaiber 1999). Besides the dimer, arginine glycation by methylglyoxal also leads to formation of hydroimidazolones (Ahmed *et al.* 2002; Ahmed and Thornalley 2002) (of which three structural isomers have been identified), tetrahydropyrimidine and argpyrimidine (Oya *et al.* 1999). Glycation may lead to conformational changes, causing protein malfunction (Luthra and Balasubramanian 1993; Raabe *et al.* 1996; Seidler and Seibel 2000; Bouma *et al.* 2003; Bakhti *et al.* 2007; Gomes *et al.* 2008).

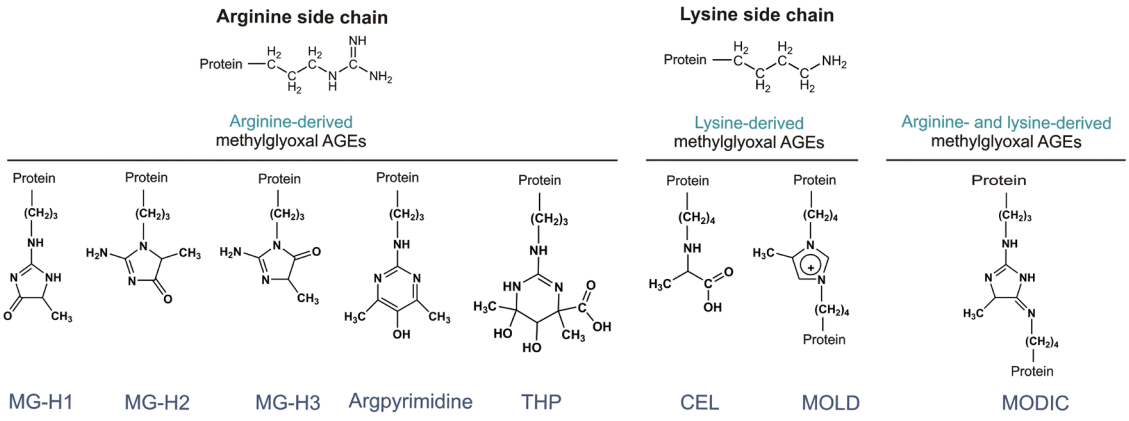


Figure1.4 Protein-methylglyoxal advanced glycation end products (adapted from Gomes 2007).

As would be expectable, methylglyoxal also reacts reversibly with protein sulfhydryl groups in cysteine residue side chains, forming hemithioacetals (Lo *et al.* 1994). Little is known about the importance of these reactions, nevertheless there are some reports on the effect of such modification in certain proteins. The Yap1 transcription factor, related to the oxidative stress response in budding yeast and analogous of mammalian AP-1, concentrates in the cell nucleus and activates the transcription of the target gene following exposure to methylglyoxal but it fails to do so if all of its cysteine residues are substituted (Maeta *et al.* 2004).

Methylglyoxal AGEs have been found to be related with characteristic clinical complications of pathologies like diabetes (McLellan *et al.* 1994), atherosclerosis (Kume *et al.* 1995), dialysis-related amyloidosis (Miyata *et al.* 1993; Miyata *et al.* 1994a; Miyata *et al.* 1994b), and neurodegenerative diseases such as Alzheimer's (Vitek *et al.* 1994; Yan *et al.* 1994; Du Yan *et al.* 1997) and Parkinson's (Castellani *et al.* 1996; Munch *et al.* 2000). They also accumulate in low turn-over proteins as collagens and eye lens proteins, being related to the ageing process (Nagaraj *et al.* 1996). In diabetes AGE accumulation is enhanced due to periodic insulin-administration-related increase of the glycolytic flux (McLellan *et al.* 1994). Methylglyoxal-lysine dimmers were found in higher level in the serum of diabetic patients (Nagaraj *et al.* 1996). Argpyrimidine was found in the intima and media of small artery walls in the kidneys of diabetic patients (Oya *et al.* 1999) and in human cataractous lenses (Padayatti *et al.* 2001). Moreover, the overexpression of glyoxalase I prevents the formation of hyperglycemia-induced AGE in bovine endothelial cells. (Shinohara *et al.* 1998). In pathologies other than diabetes, AGE formation is not related to abnormal glycaemia but to carbonylic stress, i.e. an increase of the concentration of carbonylic AGE-precursors like methylglyoxal, glycoxidation and lipoxidation products (Baynes and Thorpe 1999), either by increase of their formation or decrease of their catabolic activity. In amyloidotic diseases, including Parkinson's and Alzheimer's, AGEs were found in the corresponding characteristic amyloid fibrils (Miyata *et al.* 1993; Miyata *et al.* 1994a; Miyata *et al.* 1994b; Vitek *et al.* 1994; Yan *et al.* 1994; Castellani *et al.* 1996; Du Yan *et al.* 1997; Munch *et al.* 2000). Little is still known about the actual contribution of methylglyoxal AGEs to these pathological phenotypes and the underlying mechanisms. Another contribution of protein glycation by methylglyoxal to disease development may be the enhancement of cell oxidative stress as highly reactive radical species are formed in the process (Yim *et al.* 1995; Lee *et al.* 1998).

A study with a mammalian fibrosarcoma cell line pointed to phosphorylation of glyoxalase I and the formation of a specific methylglyoxal-protein adduct as an event in cell death associated to the tumor necrosis factor (Van Herreweghe *et al.* 2002). The authors suggest that the phosphorylated form of glyoxalase I may have a catalytic role in the formation of this adduct. In this case, the phosphorylated glyoxalase I is not a detoxifying pathway but an enhancer of methylglyoxal toxicity. On the other hand,

another study points glyoxalase I as having a role in prevention of apoptosis in human leukemic cells (Sakamoto *et al.* 2000). In addition, it is noteworthy that inhibition of glyoxalase I seems to be effective as a drug strategy against cancer (Creighton *et al.* 2003) and that it is one of the few proteins found to be uniquely overexpressed in invasive tumors when compared to low-malignancy tumors (Jones *et al.* 2002). Recently a positive correlation was found between glyoxalase I activity and anxiety, but little research has been done on the subject so far and the actual behavioral importance of the glyoxalase system and methylglyoxal is yet to be explored (Hovatta *et al.* 2005; Thornalley 2006; Landgraf *et al.* 2007; Mannervik 2008).

Glyoxalase II seems to be involved in the development of important pathological states as well. A GLO2 yeast deletion strain (lacking cytosolic glyoxalase II) is susceptible to Huntington fragment toxicity (associated to Huntington's disease) and a GLO4 yeast deletion strain (lacking mitochondrial glyoxalase II) shows sensitivity to α-synuclein toxicity (associated to Parkinson's disease). Interestingly, the cytosolic form of glyoxalase II was also shown to have a role in the prevention of methylglyoxal-induced apoptosis (Xu and Chen 2006).

1.1.2 Enzymology of the glyoxalase pathway

The major role of the glyoxalase pathway is widely accepted to be methylglyoxal detoxication, although it is also known to catabolize other 2-ketoaldehydes that may be formed in the metabolism of xenobiotic substances and in reactions involving epoxides, diols and amines (Mannervik 2008). The existence of the glyoxalase system in mammals and yeast is known since 1913 (Dakin and Dudley 1913a; Neuberg 1913), but the first kinetic characterization (for the yeast enzymes) was published only in 1957 (Kermack and Matheson 1957). In this study it was observed that the initial rate of the yeast glyoxalase I reaction was faster if started with addition of the enzyme simultaneously with glutathione compared to addition of enzyme to a mixture of glutathione, methylglyoxal and the glutathione-methylglyoxal hemithioacetal (obtained

from addition of the same initial amounts of glutathione and methylglyoxal). Due to the outcome of this experiment these authors believed that free glutathione and methylglyoxal were the actual substrates of the enzyme and they supported a two-substrate ternary-complex-mechanism Michaelis-Menten rate equation for the reaction.

$$v = \frac{V_1 \, GSH \, MGO}{K + K_{m12} \, GSH + K_{m11} \, MGO + GSH \, MGO} \qquad 1.1$$

However, in 1961 a one-substrate Michaelis-Menten equation was proposed considering the substrate of the reaction to be the glutathione-methylglyoxal hemithioacetal. As pointed by the authors, this model implied the unusual occurrence of a non-enzymatic reaction preceding the enzymatic one (Cliffe and Waley 1961).

$$v = \frac{V_1 \, HTA}{K_{m1} + HTA} \qquad 1.2$$

In this study of the yeast enzyme, the difference of initial rate (for different MGO/GSH ratios and equal hemithioacetal concentration) observed by Kermack and Matheson, which supported the two-substrate mechanism, was not observed. Also, Cliffe and Waley observed that (in the equilibrium) for the same hemithioacetal concentration, different GSH/MGO concentration ratios did not correspond to different glyoxalase I reaction rates, except for high GSH, which they explained by inhibition of the enzyme by GSH.

A first characterization of a mammalian glyoxalase I was done in 1972 for the porcine erythrocyte enzyme (Mannervik *et al.* 1972). In this study, Mannervik and colleagues considered the substrate of the reaction to be the hemithioacetal and they stated upon a detailed statistical analysis that the best rate equation to describe the reaction is the one accounting for excess substrate inhibition.

$$v = \frac{V_1 \, HTA}{K_{m1} + HTA + K \, HTA^2} \qquad 1.3$$

In the same work they also observe – and not for the first time (Koivusalo and Uotila ; Davis and Williams 1966) – that glyoxalase I activity is dependent on the presence of a divalent metal ion, such as Mg^{2+} (Mannervik *et al.* 1972). Later more conclusive studies

showed that human glyoxalase I has two catalytic sites with a Zn^{2+} ion each (Cameron *et al.* 1997). The enzyme from budding yeast seems to use Zn^{2+} as well (Frickel *et al.* 2001). However, since glyoxalase I shows significant activity when zinc is replaced by other divalent ions, the possibility that other ions occur *in vivo* in this enzyme in a situation of zinc shortage cannot be ruled out (Mannervik 2008).

In the early 1970's a statistical procedure for enzyme kinetic model discrimination was presented (Bartfai and Mannervik 1972) and applied to glyoxalase I from porcine erythrocytes (Mannervik *et al.* 1973) and yeast (Bartfai *et al.* 1973). The authors concluded that the only model able to accurately account for the kinetic data from the reaction, by meeting every discrimination criteria they used, was a random pathway joining the one- and two-substrate mechanisms.

$$v = \frac{V_A\ HTA + V_B\ HTA\ MGO}{K_1 + K_2\ HTA\ MGO + K_3\ GSH + K_4\ MGO + HTA} \qquad 1.4$$

Mannervik and co-workers stated that their study did support the coexistence of both mechanisms at least at low concentrations of glutathione and hemithioacetal. Mannervik and co-workers' method was based on the following ordered criteria: (A) the model cannot be overdetermined (if it can be linearized), i.e. the parameters can be estimated by linear regression of the linearized model from the observable variables. (B) The estimates obtained with the linear regression are used as preliminary results for a non-linear regression; if the model can not be linearized, primary estimates of the parameters can be chosen by trial and error within constraints set by physical limits or other considerations; models that do not converge or show to be very difficult to converge in the non-linear regression are regarded as defective and rejected if an alternative models exists which converges easily. (C) Parameter estimates must to be significantly different from zero, as established by the results of t-tests. (D) The residuals of the model must have an approximately 0-mean normal distribution and the expected value of their variance must approximate the residual sum of squares (RSS). (E) The RSS of the candidate models which were not rejected by the previous criteria are compared using a *F*-test for the null hypothesis $RSS_{model\ A} = RSS_{model\ B}$; if these values are significantly different, the model with smaller RSS is chosen (Bartfai and Mannervik 1972).

Although it is still uncertain which mechanism, one- or two-substrate, contributed more to the glyoxalase I-catalyzed reaction, some results providing information about the structure and the reaction mechanism were obtained. The human enzyme is a homodimer with two binding sites comprising residues from both subunits. Each subunit contains two domains quite similar to each other. The Zn^{2+} ion is bound to two residues from each subunit (glutamine 33 and glutamate 99 from one chain, and histidine 126 and glutamate 172 from the other chain) and two molecules of water which are displaced by the binding substrate (Marmstal *et al.* 1979; Cameron *et al.* 1997; Mannervik 2008). The glyoxalase I enzymes from budding yeast and from the malaria pathogen *Plasmodium falciparum* are also zinc-dependent and are single polypeptide chains joining two identical sequences homologous to the human monomer. These observations point to a double consecutive gene duplication and a fusion event in evolution (Frickel *et al.* 2001; Deponte *et al.* 2007). The reaction mechanism involves a base-catalyzed proton transfer (being glutamate 172 the likely base) from C-1 to C-2 of the hemithioacetal to form an enediol intermediate, followed by ketonization to the thioester product (Thornalley 1990; Ridderstrom *et al.* 1998; Thornalley 2003). A steady-state kinetic study of glyoxalase I from *P. falciparum* has shown that the two catalytic centers are kinetically different and exhibit cooperative behavior (Deponte *et al.* 2007).

The expression of the gene coding glyoxalase I (GLO1) is up-regulated in budding yeast when the cell is exposed to methylglyoxal (Maeta *et al.* 2005). The enzyme is also involved in the response to osmotic stress: in yeast cells grown in glycerol-containing medium extra glucose is consumed to synthesize glycerol to balance the osmotic pressure (Inoue *et al.* 1998). This higher glucose consumption leads to increased formation of methylglyoxal and expression of GLO1 (Inoue *et al.* 1998). In both cases the increase of expression of GLO1 is mediated by the high osmolarity glycerol (HOG) mitogen-activated protein kinase (MAPK) pathway (Inoue *et al.* 1998; Maeta *et al.* 2005).

While glyoxalase I has only been confirmed to be absent only in African trypanosomes (Wendler *et al.* 2009), glyoxalase II seems to be absent in more types of cell, namely human spleen, mouse heart and rat skeletal muscle (Jerzykowski *et al.* 1978). Also a

mitochondrial isoform of glyoxalase II was found besides the cytosolic one (Talesa *et al.* 1988; 1989; Talesa *et al.* 1990; Bito *et al.* 1997; Maiti *et al.* 1997; Cordell *et al.* 2004). In mammals the two isoenzymes are synthesized from the same gene but have different translation initiation points (Cordell *et al.* 2004) while in yeast two distinct genes exist (GLO4, YOR040W for the mitochondrial form and GLO2, YDR272W for the cytosolic form) (Bito *et al.* 1997). GLO4 has the distinctive feature of being expressed only in glycerol-containing medium (Bito *et al.* 1997). Remarkably, glyoxalase I is not present in rat and yeast mitochondria (Talesa *et al.* 1988; 1989; Bito *et al.* 1997), which raises questions about the real role of glyoxalase II in mitochondria. Nonetheless S-D-lactoylglutathione was found in the organelle so hydrolysis of this metabolite when it diffuses to the mitochondrion is one of the likely functions of the mitochondrial glyoxalase II (Scire *et al.* 2000).

Interestingly, glyoxalase II is not stereoselective: it catalyzes the hydrolysis of both D and L forms of S-lactoylglutathione. This means that the exclusive formation of the D isomer of lactate from methylglyoxal through the glyoxalase pathway is due to selectivity in the reaction of glyoxalase I, which has S-D-lactoylglutathione as the sole product (Mannervik 2008).

Glyoxalase II is a two-domain monomeric protein. One of the domains is a four-layer β-sheet and the other is mostly a fold of α-helix. The single active site contains two metal ions, one zinc and another bivalent ion likely to be an iron or zinc (Maiti *et al.* 1997; Cameron *et al.* 1999; Mannervik 2008). As for glyoxalase I, the identity of the ions are difficult to determine and it cannot be excluded that more than one combination of ions are used in the native protein (Mannervik 2008). One of the metals is coordinated by histidines 54, 56 and 110 while the other binds to aspartate 58, histidine 173 and histidine 59. Aspartate 134 and a hydroxyl coordinate both ions (Cameron *et al.* 1999; Mannervik 2008). Glyoxalase II appears not to be structurally or evolutionarily related to glyoxalase I but to the metal-dependent β-lactamases and it has been possible to engineer this protein to obtain β-lactamase specific catalytic activity (Park *et al.* 2006; Mannervik 2008). The hydroxyl in the active site of glyoxalase II was proposed to have a central role in the catalytic mechanism of the enzyme, by attacking the electrophilic thioester carbon of S-D-lactoylglutathione, with consequent cleavage of the glutathionyl

moiety (Cameron *et al.* 1999; Mannervik 2008). From a kinetic point of view, the hydrolysis of S-D-glutathione by glyoxalase II is well described by a one-substrate Michaelis-Menten equation (Guha *et al.* 1988; Talesa *et al.* 1990; Ridderstrom *et al.* 1996; Bito *et al.* 1999; Martins *et al.* 1999; Martins *et al.* 2001b; Freire *et al.* 2003; Gomes *et al.* 2005b; Sousa Silva *et al.* 2005a). This is consistent with the existence of a single binding site and the thermodynamics of the process, as the reaction is associated to a -8 kcal/mol Gibbs free energy variation, comparable to that of the hydrolysis of adenosine triphosphate (Cameron *et al.* 1999; Mannervik 2008).

1.2 Sensitivity analysis of metabolic pathways

Sensitivity analysis is the branch of mathematics and engineering concerned with the response of a system to permanent changes of either the numerical values of constant variables or the values of the parameters (Voit 2000). At the cellular or biochemical level a change of an independent variable is usually due to an environmental alteration while the change of a system parameter relates to an alteration of the metabolic or genetic structure of the organism. The term *sensitivity* refers to the response of a system variable to an infinitesimal change in the value of parameter of the system. For the change due to alteration of an independent variable the expression *logarithmic gain* is more often used (Heinrich and Schuster 1996; Voit 2000). Sensitivities may be expressed as regular derivates $\partial a/\partial p$ but usually they are defined in the normalized form (equation 1.5), giving the relative change of the variable a due to an infinitesimal relative change of the independent variable or parameter p.

$$\frac{\partial a / a}{\partial p / p} = \frac{\partial \ln a}{\partial \ln p} \qquad 1.5$$

Sensitivity analysis has been employed in biological studies in the frameworks of *Biochemical Systems Theory* and *Metabolic Control Analysis*. The thermodynamic and kinetic characterization of the reactions involved in the different metabolic pathways allowed the qualitative prediction on how the flux and concentrations of the intermediates might change due to a change in an independent variable or in a kinetic parameter. However, no integrated quantitative methodology existed for this purpose (Fell 1996; Voit 2000).

In a series of articles in the late 1960's and early 70's Michael Savageau created a novel mathematical theory to model biochemical pathway kinetics based on power-law differential equations. Within this Biochemical Systems Theory (BST), he used logarithmic gains and sensitivities to understand the dynamics of biochemical systems. BST not only filled the goal to be a framework to quantitatively predict the systemic changes caused by perturbations, it was also used to understand the evolutionary pressures for the appearance of regulatory mechanisms in biochemical networks such as

feedback and feedforward control (Savageau 1969a; b; 1970; 1971; 1972; 1974; 1976; Voit 2000).

Metabolic Control Analysis (MCA) is a theoretical framework originally developed with the aim of quantifying the contribution of the individual steps to the global flux (Heinrich and Rapoport 1973; 1974; Heinrich and Schuster 1996). Initially it was limited to the study of steady states and infinitesimal changes of variables and parameters. MCA challenged the traditional biochemical view which assumed the existence of rate-limiting steps controlling the flux of pathways and ultimately gave a crusial contribution to shift biochemical research into to a several-step-distributed-control perspective. In the framework of MCA *flux control coefficients* and *concentration control coefficients* were defined to quantify the infinitesimal change of flux or intermediate concentration, respectively, caused by a change in the in the concentrations of enzymes (Kacser and Burns 1973) or activities of individual steps (Heinrich and Rapoport 1973; 1974; Heinrich and Schuster 1996) – table 1.1. These coefficients have a very important role in MCA, as they provide a link between the effects of effectors on the reaction rates of the individual enzymes and the overall response of a whole network to these effectors.

Table 1.1 Flux and concentration control coefficients from metabolic control analysis (Kacser and Burns 1973; Heinrich and Schuster 1996)

Coefficient		Mathematical definition
Flux control coefficients	$C_{V_k}^{J_j}$	$\frac{V_k}{J_j}\frac{\partial J_j}{\partial V_k} = \frac{\partial \ln J_j}{\partial \ln V_k}$
Concentration control coefficients	$C_{V_k}^{S_j}$	$\frac{V_k}{S_i}\frac{\partial S_i}{\partial V_k} = \frac{\partial \ln S_i}{\partial \ln V_k}$

1.3 The Problem of Model Selection in Enzyme Kinetics

The focus of research in Biochemistry has shifted from the characterization of isolated biochemical components, namely enzymes, to that of living cells and organisms. The contemporary biochemist aims to understand the processes underlying living systems from the interaction of the components for which mathematical modeling is a central tool (Bruggeman and Westerhoff 2007). This approach requires having a reliable mathematical description of the components from the start; otherwise the error-propagation associated to merging mathematical descriptions of the smaller components would eventually render the system-level model inconsistent and useless. This is true even for small networks such as the glyoxalase pathway; as shown in chapter 2, the outputs of two different kinetic models, with parameters estimated from the same biochemical data, are different.

For this reason Enzymology is today implicitly at the centre of the development of this 'Systems Biochemistry'. In the process of characterization of an enzymatic pathway, often there is more than one possible kinetic model that fits well the data gathered from experiments. Statistics and Information Theory provide frameworks to discriminate the equation set which describes better the reactions in study.

Another important outcome of the selection of the most adequate kinetic model is that if the models correspond to specific catalytic mechanisms (or regulatory events) the selection points to the most-likely-occurring mechanism.

A necessary step prior to the employment of a procedure to select between candidate models is the estimation of the parameters of the models, addressed in the next section.

1.4 Strategies for parameter estimation

Estimation of kinetic parameters has been an important problem in Enzymology at least since the Michaelis-Menten equation was proposed. Being a non-linear equation, the estimation of its parameters was not a trivial task to begin with. While the reductionist approach of characterizing one enzyme at a time prevailed, this problem was usually tackled through the regression of initial-rate experimental data to linearized forms of the equation or eventually, for better accuracy, by non-linear regression to the hyperbolic expression directly using the result of the regression to the linearized forms as initial estimate (Cornish-Bowden 2005). In addition, biochemical kinetic models describing time-dependent variation of metabolic intermediates are built as systems of non-linear ODEs using non-linear rate equations. The kinetic parameters of these models may be estimated for each reaction rate equation in isolation, through initial-rate experiments. Alternatively, the concentrations of model variables may be measured along time, and this time-dependent data (time-course data) may be used to estimate the parameters of ODEs of the model. Time-course experiments should be more informative than initial-rate experiments, since more data is acquired in a single and eventually less time-consuming experiment (figure 1.5).

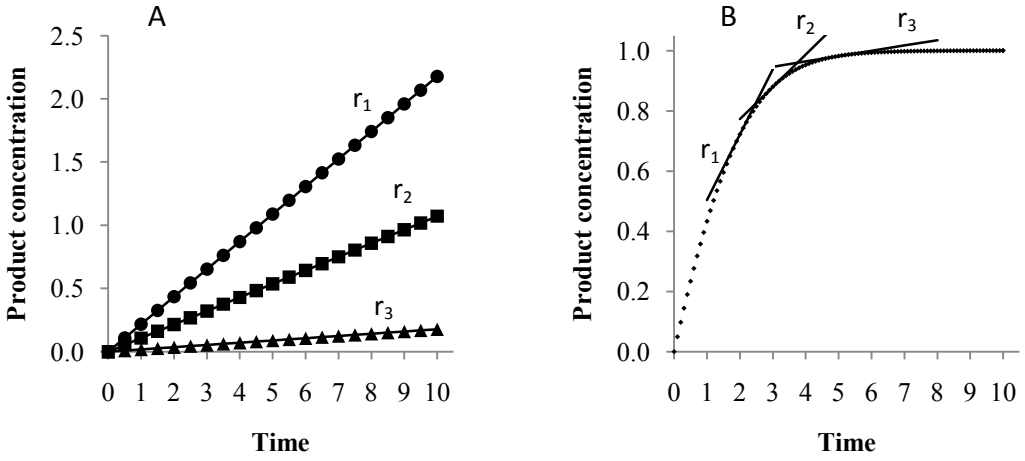


Figure 1.5 Comparison between (A) initial-rate and (B) time-course experiments. In A, three initial product formation rates need to be measured separately (data represented as circles, squares and triangles) to acquire three data points while in B the acquisition of a single time course (dots) contains the information given by those three initial-rate measurements (solid lines) and more. Time and product concentration are represented in the same arbitrary units in A and B. r_1, r_2 and r_3 represent reaction rates. The presented simulations were obtained for an irreversible single-substrate Michaelis-Menten reaction with parameters V and K_m set to 1 and initial substrate and product concentrations of 1 and 0, respectively (arbitrary units). Concentration and rates were computed with the software PLAS (Ferreira 2000).

On the other hand, focus is now on understanding cells and organisms as complex systems composed of interacting molecular components instead of studying the molecular components (namely enzymes) in isolation. Time-course data is invariably the outcome of such approach – mRNA micro-arrays, metabolome analysis with chromatographic, mass spectrometric or NMR spectroscopy are examples of methods that can be used to monitor the dynamics of biochemical systems, therefore biological systems need to be modeled from time-course data using non-linear functions.

In optimization, the function which global optimum is being searched for is named the *objective function* – an example of a commonly used objective function is the residual sum of squares (Nocedal and Wright 1999). When a non-linear model is fitted to a data set, the objective function of the optimization is also non-linear. Non-linear functions are more difficult do deal with in regression problems when compared to linear ones because they often are multimodal, i.e. they have several local minima (Polisetty *et al.* 2006; Rodriguez-Fernandez *et al.* 2006). Deterministic regression algorithms are very susceptible to multimodality, as they will find the closest local minimum to the given initial guess (Nocedal and Wright 1999). Therefore these algorithms are not the most appropriate to perform non-linear fittings to time-course data.

Stochastic optimization algorithms are a proved viable alternative to perform regressions with multimodal objective functions (Price 1999; Ashlock 2006). These algorithms introduce randomness in the search, increasing the probability of finding the global minimum (Ashlock 2006). A number of stochastic algorithms have been described which are able to solve non-linear estimation problems. Some were inspired by observations or theories from Biology, most notably ant-colony (inspired by the movements of ants) (Dorigo and Di Caro 1999), immune system-inspired methods (Dasgupta 1999) and a variety of evolutionary algorithms (based on the Darwinian theory of evolution) (Ashlock 2006). Evolutionary algorithms, in particular, have in common some basic features:

1. they start with the generation and evaluation of an initial set, S_1, of random candidate solutions;
2. a new set of solutions, S_2, is generated by applying changes (which vary with from algorithm to algorithm) to selected members of S_1;

3. the solutions in S_2 are evaluated and compared with the solutions in S_1; in each comparison, if the solution from S_2 approaches more the optimum, it replaces the solution (the one it was compared to) in S_1; otherwise the solution from S_1 stays in S_1;
4. steps 2 and 3 repeat iteratively until a termination criterion is met, i.e. the optimal solution is found (Price 1999).

In the present work, an evolutionary algorithm named Differential Evolution, which is described in detail in chapter 2, is the chosen algorithm for estimation of enzyme kinetic parameters. Differential Evolution, though having a simple implementation, has the ability to converge rapidly to the global optimum for objective functions defined in continuous domains (Storn and Price 1997; Price 1999). The solution found by the Differential Evolution is given as input to the deterministic downhill simplex algorithm (Nelder and Mead 1965) to improve the accuracy of the estimation.

1.5 Multiobjective optimization

Fitting different ODE models to time-course experimental data may result in immediate selection of a model if its performance (measured by its objective function value and eventually other criteria such as the residual distribution) is clearly better than the remaining candidates. However often more than one model fits well the experimental data and statistical discrimination of the most adequate model to describe the system is not trivial. In this case, solving the model discrimination problem may require planning an experiment optimally designed for that specific purpose.

In this dissertation a method is proposed to fulfill this goal. The idea behind the suggested procedure is to maximize the divergence between the outputs of each pair of candidate models for the experimentally observable variables. In these conditions it should be easier to discriminate among the candidate models by comparing the outputs of the models with the experimental data obtained for the same conditions.

In statistics a commonly-used measure of divergence between probability distribution functions is the Kullback-Leibler distance (Kullback and Leibler 1951). The Kullback-Leibler distance, defined as

$$I(f,g) = \int f(x \mid \theta_f) \ln \frac{f(x \mid \theta_f)}{g(x \mid \theta_g)} dx \qquad 1.8$$

measures the divergence of a probability distribution function g with respect to a probability distribution function f (Kullback and Leibler 1951; Burnham and Anderson 1998). It is important to note at this point that $I(f, g)$ is different from $I(g, f)$. The Kullback-Leibler distance was generalized to the space of positive functions as (Amisaki and Eguchi 1995)

$$I(f,g) = \int f(x \mid \theta_f) \left(\ln \frac{f(x \mid \theta_f)}{g(x \mid \theta_g)} + \frac{g(x \mid \theta_g)}{f(x \mid \theta_f)} - 1 \right) dx \qquad 1.9$$

and it can be used for positive functions such as the concentrations predicted by biochemical ODE models. Therefore the Kullback-Leibler distance can be used as objective function to find conditions to discriminate between models.

It is important that in an experiment planned for model discrimination every model as equal probability to be selected *a priori*. Therefore the procedure that maximizes the divergence between the outputs of the models must fulfill this requirement. The way to assure this is to maximize simultaneously the divergence between every possible pair of models (in both directions, if the Kullback-Leibler distance is used). Therefore multiple objective functions need to be simultaneously optimized. This problem can be solved by using a multiobjective optimization algorithm to simultaneously maximize the divergence between each two pairs of models (figure 1.6).

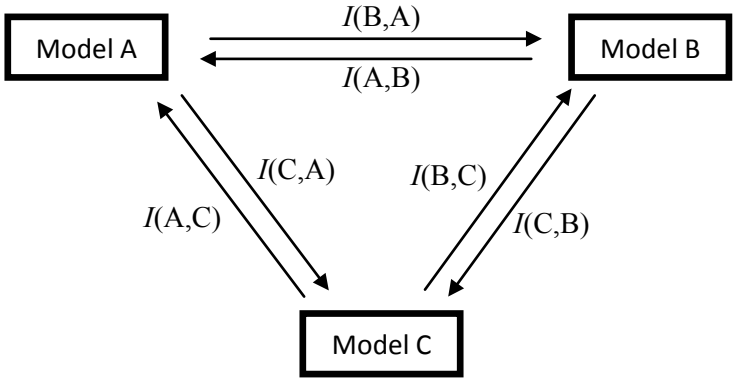


Figure 1.6 The Kullback-Leibler distance between pairs of candidate models, for an example set of 3 candidate models. Conditions are searched for that maximize the Kullback-Leibler distance (*I*) between any two models, in both directions. At Pareto-optimal solutions one of the *I*(i, j) cannot be increased without decreasing another *I*(x, y).

In opposition to single-objective optimizations, where a single optimal solution is sought, for multiobjective optimization problems several solutions of equivalent quality may exist, each with different trade-offs regarding the contribution to each objective. The set of the best (optimal) solutions for a multiobjective numerical optimization problem is named the Pareto front, after Vilfredo Pareto, who generalized this concept of optimality (Pareto 1896) originally proposed by Francis Ysidro Edgeworth (Edgeworth 1881; Coello Coello *et al.* 2007). The optimal solutions are those for which to improve the score of one objective necessarily means decreasing the score of another

objective(s). In practice optimal solutions are very hard to determine so the optimization process usually gives a set which approximates the Pareto front (Kukkonen and Lampinen 2005).

Evolutionary algorithms are naturally suited for multiobjective numerical optimizations since the generation of sets of possible solutions allows finding several optimal solutions in a single run (Coello Coello *et al.* 2007). Also, they are generally less susceptible than other stochastic algorithms to get stuck at good approximations of the Pareto set and not improve the quality of the solutions (Abraham *et al.* 2005). In an evolutionary algorithm the successive selection of new or old candidate solutions to form the next generation may also be directly used as a way to approach the Pareto front through the dominance criterion (Kukkonen and Lampinen 2005). A candidate solution dominates another if the former is as good or better than the latter for every objective function and better for at least one objective function (Kukkonen and Lampinen 2005). More formally:

Definition A

A solution u dominates a solution v, i.e. $v \prec u$ if and only if for the vector of objective functions $f = (f_1, f_2, ..., f_i, ...)$

$$\forall i : f_i(v) \leq f_i(u) \ \wedge \ \exists i : f_i(v) < f_i(u)$$

.

The best solutions found by the optimization procedure are non-dominated since there is not a dominance relationship among them and they dominate eventual remaining candidate solutions of the population.

The details of how multiobjective differential evolution is used according to the Pareto-dominance criterion to optimize experimental setups for model selection may be found on chapter 4, where the glyoxalase pathway is used as an application.

1.6 Oxidation-reduction metabolism in yeast

Pyridine nucleotides are the metabolites that contribute more for the redox potential in living cells (Nelson and Cox 2005). Indeed, appropriate NADH/NAD^+ and NADPH/$NADP^+$ ratios are fundamental to keep the catabolic and anabolic pathways, respectively, at fluxes that meet the requirements for energy and building blocks for cell growth and division. On the other hand, NADPH is involved in several detoxifying mechanisms that protect cells from damage caused by toxic metabolic by-products or other chemical agents. Methylglyoxal is such a case, as NADPH is the co-substrate of aldose reductase, the enzyme shown, along with the glyoxalase pathway, to be its most important detoxifying means, at least in yeast (Gomes *et al.* 2005b). In addition, NADPH is the reducing substrate used by glutathione reductase (glutathione:$NADP^+$ oxidoreductase, EC. 1.8.1.7) to recycle oxidized glutathione to its reduced form (Witte *et al.* 2005; Tan *et al.* 2009). Glutathione is vital for methylglyoxal catabolism through the glyoxalase pathway and generally serves as antioxidant agent, remarkably as the substrate for the glutathione peroxidase reaction, which reduces hydrogen peroxide to water (Meister and Anderson 1983; Meister 1994). NADPH is also the substrate of thioredoxin reductase (thioredoxin:$NADP^+$ oxidoreductase, EC. 1.8.1.9), the enzyme that reduces the antioxidant thiolic thioredoxin proteins, which also have a detoxifying role of oxidant agents and ensure correct sulphur assimilation and deoxynucleotide triphosphate synthesis in yeast cells (Tan *et al.* 2009).

In this context, a contribution to the understanding of the cell-wide effect of perturbations to the levels of pyridine nucleotides is included in this dissertation as a complement to the study of the glyoxalase pathway. Not only NADPH is of paramount importance in the catabolism of methylglyoxal and other toxic agents, but methylglyoxal formation rate correlates with the glycolytic flux, which depends strongly on the NADH/NAD^+ ratio, as the oxidation of the intermediates of glycolysis is coupled to the reduction of NAD^+ to NADH (Vemuri *et al.* 2007).

1.6.1 Yeast metabolic engineering as a tool in Biochemistry

The budding yeast *Saccharomyces cerevisiae* is a most useful biological model for biochemical studies. It is a unicellular eukaryote that shares most features with higher organism cells – indeed metabolic and signaling pathways are well conserved from yeast to mammals (Mustacchi *et al.* 2006). For this reason, protein synthesis and post-translational modification pathways are very similar as are the mechanisms of metabolic regulation. Its genome is fully sequenced and it's among the most extensively annotated, a wide range of well-known and versatile molecular biology tools are available, related microbiology techniques are straightforward and its cell-division rate is compatible with laboratorial time (Mustacchi *et al.* 2006). In addition, it is not pathogenic and it can be easily cultivated in bioreactor which is a great advantage to perform well-controlled studies (Nielsen *et al.* 2003). These advantages make yeast preferable even to mammal cell lines for many biochemical studies.

Metabolic engineering can be generally defined as *the direct improvement of product formation or cellular properties through modification of specific biochemical reactions or the introduction of new ones with the use of recombinant DNA technology* (Stephanopoulos *et al.* 1998). This field has an important technological relevance since it allows the optimization of the production of a desired industrial product. However, it is also extremely useful as a tool in fundamental Biochemistry since the impact of intended alterations introduced in cells allow researchers to draw important information about the operation of biochemical pathways.

In this dissertation metabolic engineering strategies are pursued to study the cell-wide effect of perturbing the NADH/NAD^+, NADPH/$NADP^+$, ATP/ADP ratios in pH- and temperature-controlled bioreactor yeast cultivations.

1.6.2 Biochemistry of pyridine nucleotides

Genome sequencing and the assignment of functional roles to genes made possible the development of metabolic genome-scale models. It was shown that metabolic networks are scale-free, meaning that the higher the number of metabolites it is more likely that one of those metabolites is involved in less reactions, following a power law distribution (Jeong *et al.* 2000) – figure 1.7.

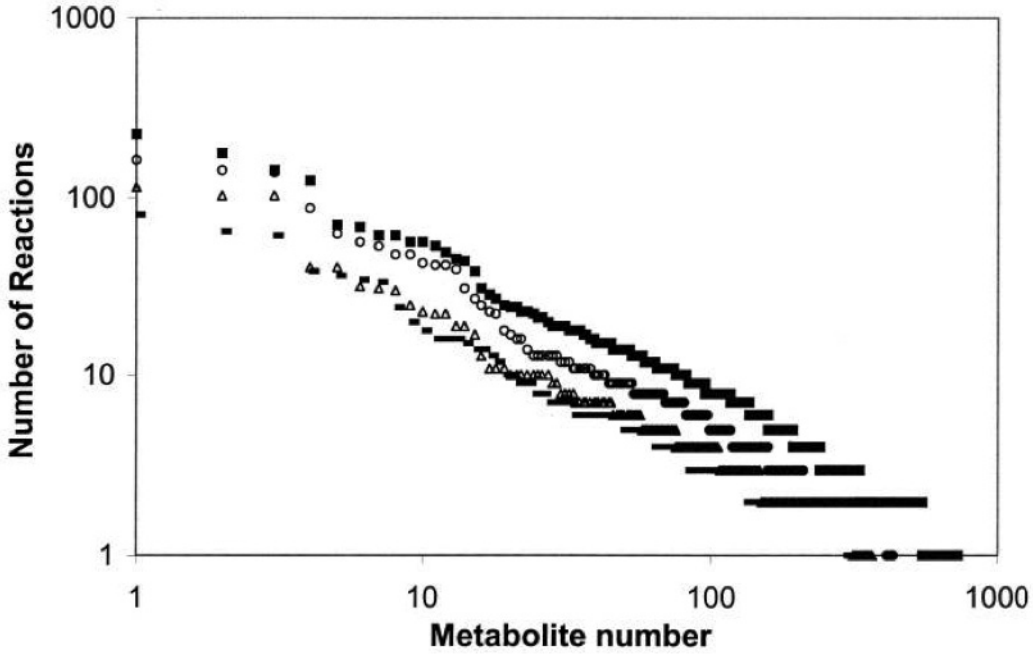


Figure 1.7 Number of reactions that metabolites are involved in versus the number of metabolites, illustrating the metabolic scale-free networks of four organisms, including *S. cerevisiae*. ° *E. coli*, ▪ *S. cerevisiae*, ▵ *H. influenzae*, ▬ *H. pylori*. Adapted from Nielsen 2003b.

The most connected metabolites of the metabolic network have been identified – table 1.2. Due to their high connectivity, variations of their level in the cell are expected to have a significant widespread effect in cellular metabolism. For this reason, it is important to understand how changes in the level of these metabolites affect the fluxes through metabolic pathways and the formation of products. Insight on this kind of response of the metabolic network to perturbations is important to understand the details of how metabolism works. It has been shown that the pairs ATP/ADP, NADPH/$NADP^+$

and NADH/NAD^+, which were known to be involved in many reactions are indeed among the most connected intermediates of metabolism (Forster *et al.* 2003).

Table 1.2 Most connected metabolites of the metabolic network of *S. Cerevisiae* (Nielsen 2003a). The species directly perturbed in the study presented in this dissertation are highlighted with bold characters.

Metabolite	Number of reactions
H^+	229
ATP	**188**
ADP	**146**
P_i	131
CO_2	90
$NADP^+$	**86**
PP_i	81
NADPH	**78**
NAD^+	**78**
Glutamic acid	68
NADH	**65**
NH_3	56

NADH/NAD^+ and NADPH/$NADP^+$ are water-soluble molecules that undergo reduction or oxidation in multiple metabolic processes. Usually they are not tightly bound to enzymes and function as co-substrates of oxidation-reduction reactions. NADPH/$NADP^+$ and NADH/NAD^+ are metabolites that largely contribute for the maintenance of a redox potential suitable to sustain a regular cell metabolism and cell growth (Nelson and Cox 2005).

$$NAD^+ + 2e^- + H^+ \rightarrow NADH \qquad E'° = -0.320$$

$$NADP^+ + 2e^- + H^+ \rightarrow NADPH \qquad E'° = -0.324$$

Equation 1.10 Standard reduction potentials for the redox pairs $NADH/NAD^+$ and $NADPH/NADP^+$ at pH 7 and 25° C.

Despite the structural resemblance, NAD^+ and $NADP^+$ have clearly distinct metabolic roles and, although exceptions exist, usually enzymes are highly specific to only one of the two (Nelson and Cox 2005). NADH is net-generated in glycolysis in the cytosol and in the citrate cycle in the mitochondria (Nelson and Cox 2005). It is mainly used as reducing agent in the respiratory electron-transport chain in the mitochondria, being the ultimate purpose of the entire process the synthesis of ATP (Nelson and Cox 2005). On the other hand, the most important source of NADPH is the pentose phosphate pathway and it is essentially used as reducing agent in biosynthetic pathways (Nelson and Cox 2005). This commitment of each reducing substrate either to an almost exclusive catabolic or anabolic function allows a differential regulation to meet both the energetic and synthetic needs of the cell. (Sauer *et al.* 2004).

The two pyridine nucleotides are essential for a number of signaling pathways, being precursors of messenger molecules (Pollak *et al.* 2007). In addition, NAD^+ is the substrate for the covalent modification of specific enzymes (Pollak *et al.* 2007).

Separated pools of $NADPH/NADP^+$ and $NADH/NAD^+$ also exist in the cytosol and in the mitochondria since these metabolites cannot diffuse through the mitochondrial membrane (von Jagow and Klingenberg 1970). Appropriate ratios of the reduced and oxidized forms of both nucleotides must be kept, otherwise metabolic processes, cell survival and proliferation might be jeopardized. This means that independent oxidation mechanisms exist for NADPH and NADH in the cytosol and in the mitochondria and both compartments have different specific enzymes for this purpose. Two cytosolic enzymes bound to the external surface of the yeast mitochondrial membrane, encoded by genes NDE1 and NDE2, oxidize NADH feeding the mitochondrial respiratory chain (Luttik *et al.* 1998). The same task is performed on the inner side of the mitochondrial membrane by the enzyme encoded by the NDI1 gene (Marres *et al.* 1991). The glycerol

synthesis pathway also plays an important role in NADH oxidation at the reaction catalyzed by the glycerol 3-phosphate dehydrogenases encoded by genes GPD1 and GPD2. Indeed these enzymes seem to be very important in keeping an appropriate $NADH/NAD^+$ ratio by oxidizing excess NADH, resulting in increased glycerol formation (Larsson *et al.* 1998). They are also part of the glycerol 3-phosphate shuttle, which provides reducing equivalents to the electron transport chain through oxidation of glycerol 3-phosphate by the mitochondrial glycerol 3-phosphate dehydrogenase, a FAD-dependent enzyme encoded in the GUT2 gene bound to the outer side of the inner mitochondrial membrane (Larsson *et al.* 1998). The importance of the malate-aspartate and ethanol-acetaldehyde shuttles, relevant in NADH oxidation in other eukaryotes, has been investigated but the relative contribution of these pathways in NADH delivery to the electron-transport chain in the budding yeast remains to be clarified (Nissen *et al.* 1997; Small and McAlister-Henn 1998; Rigoulet *et al.* 2004).

1.6.3 Pyridine nucleotide metabolism in other microorganisms

Some microorganisms have additional enzyme systems involved in the metabolism of the pyridine nucleotides, which give them the ability to adapt to otherwise very unfavorable or even lethal conditions. The pyridine nucleotide transhydrogenase and NADH oxidases are examples of such enzymes.

1.6.3.1 Pyridine nucleotide transhydrogenase

A pyridine nucleotide transhydrogenase, i.e. an enzyme that catalyzes the reaction

$$NADH + NADP^+ \rightleftharpoons NAD^+ + NADPH \qquad 1.11$$

is found in *Escherichia coli* (Anderlund *et al.* 1999) and other organisms (Nissen *et al.* 2001). In *E. coli* there are two isoforms of transhydrogenase, one soluble in the cytosol (encoded by the *udhA* gene (Boonstra *et al.* 1999)) and another bound to the plasma membrane (encoded by the *pntA* and *pntB* genes (Clarke *et al.* 1986)). The membrane-bound transhydrogenase is one of the main suppliers of NADPH in *E. coli* during standard aerobic batch growth on glucose while the soluble isoenzyme is essential when the carbon substrate is oxidized through the pentose phosphate pathway alone. Furthermore, the *pntA* and *udhA* genes are down-regulated when NADPH production is increased or decreased, respectively (Sauer *et al.* 2004). So the transhydrogenases buffer the concentrations of NADH and NADPH allowing the cell to cope well with extreme variations of anabolic and catabolic demands (Sauer *et al.* 2004).

1.6.3.2 NADH oxidases

Metabolism of pyridine nucleotides is involved in the virulence of some microbial species. Many organisms, among which *Streptococcus pneumoniae*, carry a gene encoding a soluble NADH oxidase (*nox*). The deletion of this gene significantly attenuates the virulence of these bacteria and limits its ability to grow in aerobic conditions (Yu *et al.* 2001). In addition, decrease of the expression of the *nox* gene results in decrease of the natural competence of *S. pneumoniae* cells (Auzat *et al.* 1999). Other organisms such as the pathogenic fungus *Histoplasma capsulatum* have a branched respiratory electron transport chain, with cyanide-sensitive and cyanide-insensitive oxygen-consuming branches. The cyanide insensitive branch comprises an alternative NADH oxidase. The *AOX1* gene encoding this enzyme was shown to be up-regulated when *H. capsulatum* cells are exposed to hydrogen peroxide-mediated oxidative stress or electron transport enzymes inhibitors antimycin A and cyanide, suggesting that the alternative oxidase is important for the survival of the organism in these conditions and perhaps for its virulence (Johnson *et al.* 2003).

1.6.4 Cell redox potential and glycolytic metabolism

For yeast growing in low-glucose-concentration media, glucose is fully oxidized to carbon dioxide through glycolysis, the pyruvate dehydrogenase complex and the citrate cycle (Vemuri *et al.* 2007). NAD^+ is the electron acceptor in the oxidizing reactions of these pathways (Nelson and Cox 2005). In high-glucose-concentration media, cells are able to uptake more glucose than they can oxidize through the respiratory pathway (Vemuri *et al.* 2007). A phenomenon known as Crabtree effect (Crabtree 1928) occurs when glucose concentration is such that its uptake rate by yeast cells is higher than its oxidation through the respiratory pathway. In this situation, glucose is also oxidized through the ethanol-fermentative pathway.

Concomitantly with the Crabtree effect, a set of regulatory events triggered by high glucose concentrations takes place. Several genes encoding proteins involved in respiration, alternative carbon source (such as galactose, sucrose and maltose) uptake and catabolism are repressed. On the other hand, the expression of glucose transporter- and glycolytic enzyme-encoding genes is induced (Entian and Schüller 1997; Carlson 1999). Even if glucose repression is not the cause of the Crabtree effect, it should give a significant contribution to increase it. Also, since the conversion of glucose to ethanol generates less ATP than the respiratory pathway, glucose uptake and the glycolytic flux increase even more to meet the demand of biomass production, enhancing the Crabtree effect. A recent study has shown that a yeast strain overexpressing the *AOX1* gene from *H. capsulatum*, which is a strain with increased NADH-oxidizing activity of the electron transport chain, requires higher glucose concentrations to display the Crabtree effect. This result suggests that the Crabtree effect is due not to saturation of the pyruvate dehydrogenase complex or repression of respiration by glucose but to a saturation of the respiratory metabolism by NADH (Vemuri *et al.* 2007). In the same work it was evident that glycerol synthesis has a crucial role in the oxidation of excess NADH, allowing maintaining an appropriate $NADH/NAD^+$ ratio for the glycolytic flux to handle higher glucose availability.

These metabolic features are thought to be advantages from an evolutionary point of view as yeast is able to uptake the glucose at a very high rate, which varies linearly with

the concentration in the medium (Stephanopoulos *et al.* 1998). Otherwise the available glucose might be consumed by competing organisms. The production of ethanol and the decrease of medium pH also prevent the growth of competing microorganisms. The breakdown of glucose to ethanol is a redox-neutral pathway so it doesn't oxidize excess NADH, which is necessary to keep or increase the glycolytic flux. Therefore, part of the carbon from glucose is directed to the synthesis of glycerol, a NADH-consuming pathway, at the triose phosphate branch point (Nelson and Cox 2005; Vemuri *et al.* 2007).

2 Steady-state symbolic analysis of the glyoxalase pathway

2.1 Abstract

Two kinetic models have been used to describe the glyoxalase pathway: one in which methylglyoxal and glutathione are sequentially-binding substrates of glyoxalase I and another where their hemithioacetal is considered to be the substrate of this enzyme. The steady-state concentration of methylglyoxal and its sensitivity were derived algebraically for both of these models and were expressed as functions of glyoxalase I and II activities, total glutathione concentration and methylglyoxal formation rate. It is shown that due to the regeneration of glutathione, thresholds higher than methylglyoxal influx have to be met by the pathway enzyme activities to avoid steady state disruption and substrate accumulation. Total glutathione concentration must also meet a threshold to ensure the existence of a steady state. Under physiological conditions, the pathway operates far from these thresholds. Sensitivity analysis gives a rationale to justify these operating conditions: the sensitivity of methylglyoxal concentration to pathway input and thiol concentration is supra-linear. Therefore, the system seems to operate under high activity values and high thiol concentration to ensure the minimum possible absolute value for the sensitivities of this toxic compound. It also is shown that the predicted effect of glyoxalase I inhibition on the concentration of methylglyoxal depends heavily on the kinetic mechanism of this enzyme as shown for the two candidate models considered. In case of the two-substrate model a significant sensitivity of methylglyoxal concentration to the activity of this enzyme suggests that glyoxalase I should be an enzyme to be considered as a drug target in the context of broader therapeutic strategies.

2.2 Introduction

The glyoxalase pathway was suggested to be a drug target in many different pathological conditions, ranging from cancer to diabetes (Thornalley 1993) and, recently, trypanosomatid-related diseases such as African-sleeping sickness (Irsch and Krauth-Siegel 2004; Vickers *et al.* 2004; Sousa Silva *et al.* 2005b; Trincao *et al.* 2006; Silva *et al.* 2008). Despite having been discovered in 1913 (Dakin and Dudley 1913a; Neuberg 1913) and extensively studied since then, this pathway's actual biological relevance is still not well understood (Mannervik 2008). The main known role of the glyoxalase enzymes seems to be the degradation of toxic α-ketoaldehydes (Mannervik 2008). One of these compounds in particular, methylglyoxal, is a ubiquitous by-product of glycolysis formed by spontaneous elimination of the phosphate group during the interconversion of dihydroxyacetone phosphate and glyceraldehydes 3-phosphate (Meyerhof and Lohmann 1934; Richard 1991; 1993a) (figure 2.1).

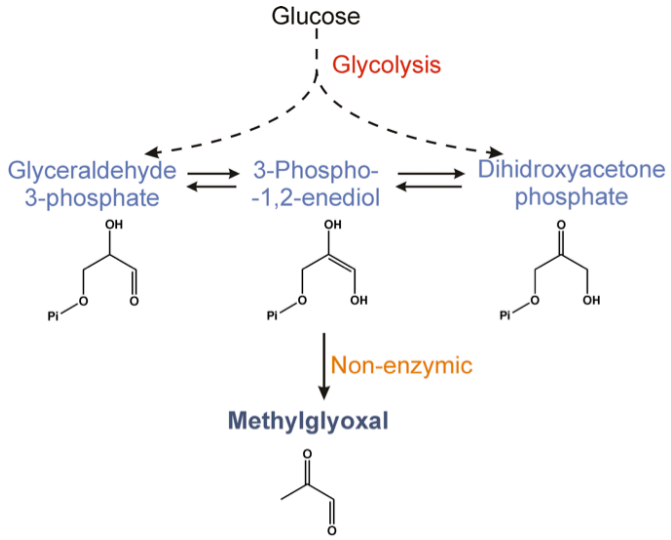


Figure 2.1 Main methylglyoxal formation reaction: spontaneous elimination of the phosphate group of the enediolic intermediate in equilibrium with the glycolytic triose phosphates.

Methylglyoxal is known to form adducts with proteins which are associated to clinical complications of diabetes and neurodegenerative diseases in a process known as glycation (Nyhlin *et al.* 2000; Dukic-Stefanovic *et al.* 2001; Sousa *et al.* 2001; Chen *et al.* 2004; Gomes *et al.* 2005a; Ahmed and Thornalley 2007). While in the latter cases methylglyoxal is an undesired inevitability, it was regarded as a possible chemical agent

against tumor cells (Kang *et al.* 1996) and pathogenic unicellular eukaryotes such as trypanosomatids (Irsch and Krauth-Siegel 2004; Vickers *et al.* 2004), which have in common a remarkably high glycolytic activity in some stages of their life cycle (Sousa Silva *et al.* 2005a; Thornalley 2008). The glyoxalase pathway is considered the most promising target for rational drug design to have an effect on methylglyoxal concentration *in vivo* and is regarded as especially promising in the context of trypanosomatid-associated diseases due to the enzyme peculiar specificity in these organisms. Glutathione is replaced by N^1,N^8-bis-glutathionyl spermidine (commonly named trypanothione), a spermidine backbone substituted with two glutathionyl groups, one in each terminal amine (Irsch and Krauth-Siegel 2004; Vickers *et al.* 2004) as cofactor of the glyoxalase pathway and other glutathione-dependent systems in trypanosomatids.

Kinetic studies of the glyoxalase system revealed that it operates through either one of two different mechanisms (Mannervik *et al.* 1973; Mannervik *et al.* 1974; Freire *et al.* 2003; Gomes *et al.* 2005b; Sousa Silva *et al.* 2005b). The first mechanism comprises the spontaneous non-enzymatic formation of a hemithioacetal of methylglyoxal and glutathione (or trypanothione) which subsequently is the substrate of the first enzyme of the pathway, glyoxalase I, with single-substrate irreversible Michaelis-Menten kinetics. In the second mechanism, glutathione and methylglyoxal bind sequentially to glyoxalase I and form the hemithioacetal in the active centre of the enzyme, which catalyzes its isomerization as a double-substrate irreversible Michaelis-Menten reaction. It remains to be proved the extent to which each model contributes to the global reaction but several studies favor the mechanism where the hemithioacetal binds glyoxalase I instead of methylglyoxal and glutathione (Mannervik *et al.* 1973; Mannervik *et al.* 1974; Sellin and Mannervik 1983). For glyoxalase II a single-substrate irreversible Michaelis-Menten model is widely accepted (Guha *et al.* 1988; Talesa *et al.* 1990; Ridderstrom *et al.* 1996; Bito *et al.* 1999; Martins *et al.* 1999; Martins *et al.* 2001b; Freire *et al.* 2003; Gomes *et al.* 2005b; Sousa Silva *et al.* 2005a).

The present study aims at understanding if the inhibition of glyoxalase I and II is a viable to increase methylglyoxal concentration in a magnitude that justifies the use of these enzymes as drug targets. Mathematical conditions for the existence of a steady state are derived for the two opposing glyoxalase I kinetic models to test whether steady state disruption is possible. In addition, methylglyoxal steady-state concentration

sensitivities to enzyme activities, methylglyoxal formation rate and total glutathione concentration are algebraically derived to evaluate the effect of changing these parameters. Implications for the control of the pathway and its usefulness in therapeutic strategies are discussed.

2.3 Materials and methods

2.3.1 Glyoxalase pathway kinetic models

Two kinetic models of the glyoxalase pathway were studied: one in which the methylglyoxal-glutathione hemithioacetal is the substrate of glyoxalase I (model 1 – figure 2.2 A) and another in which methylglyoxal and glutathione are direct substrates of this enzyme (model 2 – figure 2.2 B). So, in model 2 methylglyoxal binds to glyoxalase I after glutathione and the hemithioacetal is formed in the enzyme's active centre. To evaluate the influence of the hemithioacetal-formation reaction on the behavior of the system with a two-substrate mechanism a third model was studied, which comprehends the same rate equations as model 2 except hemithioacetal non-enzymatic formation and breakdown (model 3).

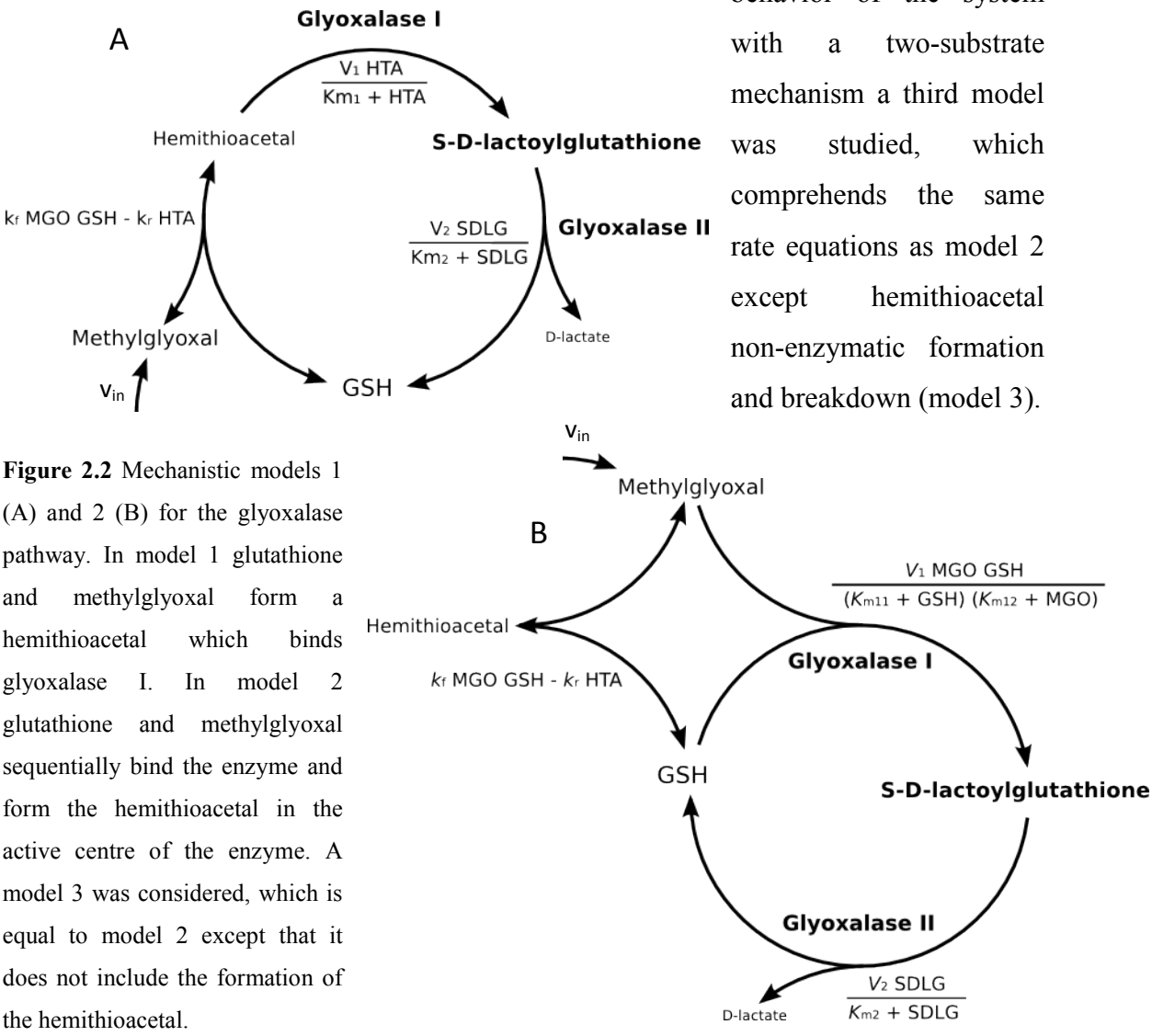


Figure 2.2 Mechanistic models 1 (A) and 2 (B) for the glyoxalase pathway. In model 1 glutathione and methylglyoxal form a hemithioacetal which binds glyoxalase I. In model 2 glutathione and methylglyoxal sequentially bind the enzyme and form the hemithioacetal in the active centre of the enzyme. A model 3 was considered, which is equal to model 2 except that it does not include the formation of the hemithioacetal.

In the two models methylglyoxal is considered to be formed at a constant rate (v_{in}) (Gomes *et al.* 2005b). In models 1 and 2 it reacts reversibly with the sulfhydryl group of glutathione forming a hemithioacetal. In model 1 this hemithioacetal is isomerized to S-D-lactoylglutathione or S-D-lactoyltrypanothione by glyoxalase I. In model 2 hemithioacetal formation is a dead-end branch and methylglyoxal and glutathione direct substrates of glyoxalase 1. S-D-lactoylglutathione is hydrolyzed by glyoxalase II, forming D-lactate and regenerating glutathione. S-D-lactoylglutathione hydrolysis is equal for the three models. In accordance with previous experimental studies, hemithioacetal formation is a reversible mass-action reaction (Creighton *et al.* 1988) and glyoxalase I and II follow irreversible Michaelis-Menten kinetics (Freire *et al.* 2003; Gomes *et al.* 2005b). Michaelis constants and reference enzyme activities were estimated as described in section 2.3.2 and kinetic constants for hemithioacetal non-enzymatic formation and breakdown are from the literature (Creighton *et al.* 1988; Ponces Freire *et al.* 2003; Gomes *et al.* 2005b). These reference values are in table 1.

2.3.2 Estimation of the kinetic parameters of glyoxalase I and II

The kinetic parameters of glyoxalase I and II were estimated from the same S-D-lactoylglutathione time-course data, instead of using parameters from the literature, to assure an objective comparison of both models. S-timator, an in-house-developed software by António E. N. Ferreira, was used to estimate enzyme kinetic parameters with the Differential Evolution algorithm using the simple least-squares criterion.

Time-course kinetic assays

For the time-course assays *Saccharomyces cerevisiae* cells were harvested at the end of exponential growth phase and incubated at 30 °C for 30 min in an orbital-shaker (Infors, Bottmingen, Switzerland) at 160 rpm for permeabilization with digitonin (Calbiochem) 0.01 % (m/V) in MES (Sigma) 0.05 M pH 6.5, as described (Cordeiro and Freire 1995). S-D-lactoylglutathione absorbance at 240 nm was measured in an Agilent 8453

spectrophotometer with magnetic stirring at a controlled temperature of 30°C. A volume of permeabilized cells suspension corresponding to a total of 50 µg of protein was used, measured using the Bradford reagent (Biorad) and a bovine serum albumin (Sigma) calibration curve. The assays were done in a reaction volume of 3 mL using KH_2PO_4 (Sigma) 70 mM pH 6.5 as buffer. Assays started with addition of methylglyoxal after pre-incubation of permeabilized cells and glutathione (Sigma). Two sets of glutathione and methylglyoxal initial concentrations were used: 1.00 and 0.33 mM respectively in the first assay and 0.66 mM for both in the second. Methylglyoxal was prepared fresh by heat-acid hydrolysis of methylglyoxal-1,1-dimethylacetal (Sigma) (Kellum *et al.* 1978). Estimated activities were normalized to the median *S. cerevisiae* cell volume of 36.6 μm^3 (Tyson *et al.* 1979).

Differential evolution

Most enzyme kinetic models, including the two considered here, are built with non-linear ODEs, which leads to the existence of multiple local minima (multimodality) of the objective function for the parameter estimation (the residual sum of squares (Polisetty *et al.* 2006; Rodriguez-Fernandez *et al.* 2006). Stochastic algorithms, namely evolutionary algorithms, perform better than the deterministic for multimodal non-linear objective functions (Price 1999; Moles *et al.* 2003; Ashlock 2006).

As the name suggests, evolutionary algorithms are elegantly inspired by the biological theory of evolution (Ashlock 2006). Invariably these algorithms initialize a set (population) of potential solutions randomly in the search space. The objective function is evaluated for this set of solutions and a new set of solutions is somehow generated from the previous one with the introduction of variations (*crossover* and *mutation*). The new solutions are then compared to the first ones and those with scores large enough according to a selection criterion are selected. The selected solutions have the chance to *reproduce* and a new set of solutions is generated. The algorithm repeats this cycle of genetic operations on the solutions, selection and reproduction until a stop criterion is met (Ashlock 2006). Differential Evolution (DE), the optimization procedure chosen for this work, is an evolutionary algorithm, which despite its simplicity has proven to be robust and remarkably effective in converging fast to the actual global minimum in problems with real and multimodal objective functions (Price 1999).

The DE algorithm was originally designed to fulfill four requirements (Storn and Price 1997): (i) to be able to handle non-differentiable, non-linear and multimodal cost functions; (ii) to be easily parallelizable; (iii) to be easy to use by having few robust easy-to-choose control variables to drive the minimization and (iv) to consistently reach the optimal solution in several consecutive tests. The authors claim to have satisfied the four goals. DE, as implemented in S-timator, initializes by generating a user-defined number of random possible solution vectors, *NP*, within a space defined by constraints. The probability for the generation of the candidate solutions is uniform in the parameter space. Each initial solution vector is then compared with a mutated vector. In DE, *mutation* is achieved by adding a weighted difference of two random vectors to a third random vector (all different from the target vector to be compared with). The weighting factor *F* (equal to or higher than 0) may be set according to the specific needs of the problem. In S-timator it is set to 0.5. The new mutated vector is built from parameters resulting from mutation and parameters from the target vector in a process called *cross-over*, which depends on a probability parameter, *CR*. *CR* is the probability of a parameter in the mutated vector to be a result of mutation while 1 - *CR* is the probability that the parameter is the same as in the target vector. In S-timator *CR* is set to 0.7. This mutated vector is compared with the corresponding target vector and the one yielding the lowest objective function value is *selected* to be part of the next generation. After evaluation of the objective function for the initial set of solutions, this initial population enters successive cycles of mutation, recombination and selection named generations. The optimization stops when the maximum number of generations set by the user, G_{max}, is reached or another stopping criterion is met (figure 2.3).

Input: D, G_{max}, $NP \geq 4$, $F \in [0, 1+]$, $CR \in [0, 1]$ and $\vec{x}^{(lo)}$, $\vec{x}^{(hi)}$

Initialize P_G: $(\vec{x}_{1,G}, \vec{x}_{2,G}, \ldots, \vec{x}_{NP,G})$

$G = 0$

$\forall\, i \leq NP \wedge \forall\, j \leq D$: $x_{j,i,G=0} = x_j^{(lo)} + rand_j[0,1] \cdot \left(x_j^{(hi)} - x_j^{(lo)}\right)$

While $G < G_{max}$ or termination criteria not met:

$\forall\, i \leq NP$:

Mutate and recombine

Select $\vec{x}_{i,G+1} = \vec{u}_{i,G+1}$ if $f(\vec{x}_{i,G+1}) \leq f(\vec{u}_{i,G+1})$ or $\vec{x}_{i,G+1} = \vec{x}_{i,G}$ otherwise

Figure 2.3 Pseudo-code for the differential evolution algorithm. D is the number of parameters to be estimated, G_{max} is the allowed number of generations, NP is the size of the population of solutions, F is a scaling factor and CR is the probability of mutation (Price 1999).

Several mutation schemes have been proposed for use in DE (Price 1999). For the estimation of the kinetic parameters of the glyoxalase pathway, DE/rand/1/bin was used, meaning *differential evolution*/*randomly* chosen donor vectors for mutation/*number* of pairs of vectors contributing to mutation/discretely recombined with the parameters of the target vector using a *binomial* distribution (figure 2.4). This recombination scheme is the simplest proposed for use with differential evolution and has the advantage of keeping the population of candidates well distributed in the search space while converging to the optimal solution. In S-timator, the optimal solutions found by DE are given as input to the downhill-simplex deterministic algorithm (Nelder and Mead 1965) to improve the accuracy of the estimates.

DE/rand/1/bin:

$r_1, r_2, r_3 \in (1, 2, \ldots, NP)$ randomly chosen with $r_1 \neq r_2 \neq r_3 \neq i$

$j_{rand} \in (1, 2, \ldots, D)$ randomly chosen once for each i

$\forall\, j \leq D$, $u_{j,i,G+1}$:

$u_{j,i,G+1} = x_{j,r_3,G} + F \cdot \left(x_{j,r_1G} - x_{j,r_2G}\right)$ if $rand_j[0,1] < CR \vee j = j_{rand}$ or $x_{j,i,G}$ otherwise

Figure 2.4 DE/rand/1/bin recombination strategy (Price 1999).

Computational implementation

S-timator was developed in Python, making use of the numpy (numpy.scipy.org) and scipy (www.scipy.org) modules for numeric calculations, sympy (code.google.com/p/sympy) for symbolic computations, wxPython (www.wxpython.org) to build the user interface and matplotlib (matplotlib.sourceforge.net) to display graphical results. ODE integration is done with the *odeint* function from the scipy.integrate._odepack module, which uses the Adams/Backward Differentiation Formulae method (Hindmarsh 1983; Hindmarsh and Petzold 1983). The *fmin* function of the scipy.optimize module is used as implementation of the downhill-simplex algorithm (Nelder and Mead 1965) .

2.3.3 Symbolic derivation of steady-state conditions and sensitivity functions

At steady state, net variation of metabolite concentrations is zero, therefore equations 2.1.1 to 2.1.4 apply in case of model 1, equations 2.2.1 to 2.2.3 apply in case of model 2 and equations 2.3.1 to 2.3.3 apply for model 3. Equations 2.1.1, 2.2.1 and 2.3.1 stand for the conservation of total glutathione – the sum (SH_{Total}) of the concentrations of the glutathionyl groups existing in hemithioacetal (HTA) and S-D-lactoylglutathione (SDLG) besides free glutathione (GSH), is constant in models 1 and 2 and the sum of free glutathione and S-D-lactoylglutathione is constant in model 3.

$$SH_{Total} = GSH + HTA + SDLG \quad 2.1.1$$

$$v_{in} - v_f + v_r = 0 \quad 2.1.2$$

$$v_f - v_r - v_1 = 0 \quad 2.1.3$$

$$v_1 - v_2 = 0 \quad 2.1.4$$

$$SH_{Total} = GSH + SDLG + HTA \quad 2.2.1$$

$$v_{in} - v_f + v_r - v_1 = 0 \quad 2.2.2$$

$$v_f - v_r = 0 \quad 2.2.3$$

$$v_1 - v_2 = 0 \quad 2.2.4$$

$$SH_{Total} = GSH + SDLG \quad 2.3.1$$

$$v_{in} - v_1 = 0 \quad 2.3.2$$

$$v_1 - v_2 = 0 \quad 2.3.3$$

The systems of equations 2.1.1 to 2.1.4, 2.2.1 to 2.2.3 and 2.3.1 to 2.3.3 were algebraically solved for metabolite concentrations with the function *Solve* in Mathematica 6 from Wolfram Research, Inc. (Champaign, IL, USA) and steady-state concentrations were found as functions of the parameters of the system (equations 2.4.1 to 2.4.4 for model 1, equations 2.5.1.1 to 2.5.4.2 for model 2 and equations 2.6.1 to 2.6.4 for model 3).

Sensitivity algebraic expressions were also computed as functions of the parameters and intermediate steady-state concentrations by the methods developed by Reder, 1988. The simple derivatives of the steady-state concentrations were calculated with respect to the parameters using the equation

$$\mathbf{D_\mu \sigma = \Gamma\, D_\mu v}$$

where $\boldsymbol{\mu}$ represents the vector of parameters, $\boldsymbol{\sigma}$ represents the vector of steady-state concentrations, $\mathbf{v}$ represents the vector of reaction rates, $\mathbf{D_\mu \sigma}$ represents the matrix of the derivatives of the steady-state concentrations with respect to the parameters and $\mathbf{D_\mu v}$ represents the matrix of the derivatives of the reaction rates with respect to the parameters. The matrix $\boldsymbol{\Gamma}$ is calculated according to the equation

$$\mathbf{\Gamma = - L\, (N_R\, D_x v\, L)^{-1}\, N_R}.$$

These equations may be deduced from the differential equation system defining the internal state of the pathway

$$\frac{d\mathbf{x}}{dt} = \mathbf{N}\,\mathbf{v}$$

where **x** and **v** represent the intermediate concentration vector and the reaction rate vector, respectively. From the *n*-ranked **N** matrix, the *reduced matrix* of **N**, $\mathbf{N_R}$, and the *link matrix* **L**, can be determined:

$$\mathbf{N} = \mathbf{L}\,\mathbf{N_R}.$$

where **L** is a matrix of *m* lines and *n* columns. The first *n* lines of **L** are the *n*-dimensional identity matrix (Reder 1988).

By normalizing the simple derivatives computed using this method to the logarithmic space, sensitivity equations 2.7.1 to 2.7.4 were obtained for model 1 and equations 2.8.1 to 2.8.3 for model 2 (sensitivity to methylglyoxal influx is not shown for model 2).

2.4 Results

2.4.1 Time-course-based kinetic characterization of the glyoxalase pathway

Two models have been considered to describe the kinetic behavior of the glyoxalase pathway (figure 2.2). To analyze the outcome of using each model when studying the control of the pathway, the kinetic parameters of the enzymes were estimated from the same data for both models. S-D-lactoylglutathione time-course data was acquired using permeabilized yeast cells, a method which keeps in-cell protein concentrations and preserves native protein-protein interactions. The rate equations of the kinetic models are indicated in figure 2.2 and the estimates of the kinetic parameters of models 1 and 2 are outlined in table 2.1.

Table 2.1 Estimates of the kinetic parameters of models 1 and 2, reference total glutathione concentration and reference methylglyoxal formation rate in *S. cerevisiae*

Parameter	Model 1	Model 2
K_{m1}	0.214 mM	-
K_{m2A}	0.0286 mM	-
K_{m11}	-	2.569 mM
K_{m12}	-	1.202 mM
K_{m2B}	-	0.0347 mM
V_{IA}	4.569 mM/min	-
V_{IIA}	0.633 mM/min	-
V_{IB}	-	25.727 mM/min
V_{IIB}	-	0.658 mM/min
v_{in}	2.41×10^{-3} mM/min *	
SH_{Total}	4.00 mM *	

*From Gomes *et al.* 2005b

Both model 1 and model 2 fit very well the experimental data and the corresponding curves nearly overlap (figure 2.5).

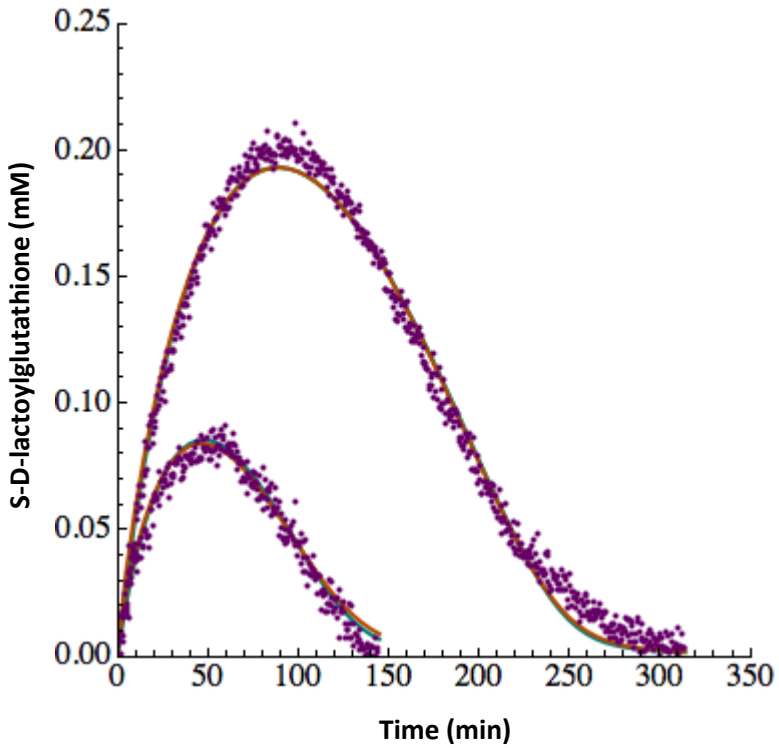


Figure 2.5 Experimental data (magenta) and fitting of models 1 (cyan) and 2 (orange). Glutathione and methylglyoxal initial concentrations were 1.00 and 0.33 mM for the experimental data set 1 and 0.66 mM for both substrates.

2.4.2 Steady-state analysis of the glyoxalase pathway

The systems of algebraic equations that characterize the glyoxalase pathway at steady state were solved for models 1 and 2. Equations 2.4.1 to 2.4.4 were found to define the concentrations of the intermediates of the pathway at steady state for model 1. For model 2 two possible steady states were found defined by equations 2.5.1.1, 2.5.2.1, 2.5.3 and 2.5.4.1 and 2.5.1.2, 2.5.2.2, 2.5.3 and 2.5.4.2. Another model was then considered, model 3, which is equal to model 2 but does not account for the formation of hemithioacetal. For model 3 only one steady state was found (equations 2.6.1 to

2.6.4), meaning that the existence of two possible steady states in model 2 is a consequence of the formation of hemithioacetal. The detailed mathematical expressions defining the symbols used here may be found in appendix A.

$$MGO_1 = \frac{v_{in}\ E_{2A}\ A_1}{k_1\ E_{1A}\ E_{2A}\ SH_{Total} - k_1\ v_{in}\ B_1} \quad 2.4.1$$

$$HTA_1 = \frac{K_{m1}\ v_{in}}{E_{1A}} \quad 2.4.2$$

$$SDLG_1 = \frac{K_{m2A}\ v_{in}}{E_{2A}} \quad 2.4.3$$

$$GSH_1 = SH_{Total} - \frac{v_{in}\ B_1}{E_{1A}\ E_{2A}} \quad 2.4.4$$

$$MGO_{21} = \frac{A_2 + B_2\sqrt{C_2}}{k_f\ v_{in}\left(D_2 - \sqrt{C_2}\right)} \quad 2.5.1.1$$

$$MGO_{22} = \frac{A_2 - B_2\sqrt{C_2}}{k_f\ v_{in}\left(D_2 + \sqrt{C_2}\right)} \quad 2.5.1.2$$

$$HTA_{21} = \frac{\delta_2 + \sqrt{\gamma_2}}{2\,E_{2B}\,Q_2} \quad 2.5.2.1$$

$$HTA_{22} = \frac{\delta_2 - \sqrt{\gamma_2}}{2\,E_{2B}\,Q_2} \quad 2.5.2.2$$

$$SDLG_2 = \frac{K_{m2B}\ v_{in}}{E_{2B}} \quad 2.5.3$$

$$GSH_{21} = \frac{\varphi_2 - \sqrt{\gamma_2}}{2\,E_{2B}\,Q_2} \quad 2.5.4.1$$

$$GSH_{22} = \frac{\varphi_2 + \sqrt{\gamma_2}}{2\,E_{2B}\,Q_2} \quad 2.5.4.2$$

$$MGO_3 = \frac{K_{m12}\, v_{in}\, A_3}{v_{in}\, C_3 - E_{1B}\, E_{2A}\, SH_{Total}} \quad 2.6.1$$

$$HTA_3 = \frac{k_1\, Km_{12}\, v_{in}\, D_3\, G_3}{E_{2B}\, L_3} \quad 2.6.2$$

$$SDLG_3 = \frac{K_{m2B}\, v_{in}}{E_{2B}} \quad 2.6.3$$

$$GSH_3 = \frac{k_{-1}\, D_3\, M_3}{E_{2B}\, L_3} \quad 2.6.4$$

For the reference parameters outlined in table 2.1, the steady-state concentrations predicted by models 1 and 2 are the ones in table 2.2. The steady state was confirmed to be stable for model 1 as the eigenvalues of the jacobian matrix of corresponding ODE system are real and negative (table 2.2) (Heinrich and Schuster 1996). For model 2, one of the states comprises a very high methylglyoxal concentration and almost complete depletion of glutathione and for the other state the opposite happens as there's a high concentration of free glutathione and a low concentration of methylglyoxal (table 2.2). Stability analysis of model 2 steady states revealed that the state which accounts for a low concentration of methylglyoxal is stable (table 2.2). However, the high-methylglyoxal-concentration steady state the stability analysis gave an inconclusive result since one of the eigenvalues of the jacobian matrix was 0 (table 2.2).

Table 2.2 Reference intermediate steady-state concentrations and corresponding jacobian matrix eigenvalues for the reference parameters outlined in table 2.1.

Intermediate	Reference steady-state concentrations (mM)		
	Model 1	Model 2	
		State 1	State 2
Methylglyoxal	1.86×10^{-3}	1.85×10^{-4}	4.94×10^{4}
Hemithioacetal	1.13×10^{-4}	2.49×10^{-4}	3.9996
S-D-lactoylglutathione	1.09×10^{-4}	1.27×10^{-4}	1.27×10^{-4}
Glutathione	3.9997	3.9996	2.41×10^{-4}
	Jacobian matrix eigenvalues		
	-22.44 -21.98 -1.30	-18.84 -14.49 -0.91	-18.84 -1.01 0

Model 2 was then further investigated with numerical simulations to understand the behavior of the system near the ambiguous steady state and it was found to be. Perturbation of this steady state with a slight increase of methylglyoxal concentration leads the system away from the steady state and to the building-up of methylglyoxal concentration over time. On the other hand, even a slight decrease of methylglyoxal concentration leads to the stable steady state. The low-methylglyoxal steady state is stable far beyond the *in vivo* concentration, which is in the μM range (Ohmori *et al.* 1987; Thornalley 1990; McLellan and Thornalley 1992; Cordeiro 1996). The value of methylglyoxal concentration in the stable steady state in model 2 is 10 fold lower than concentration of that compound in the steady state of model 1, suggesting that the mechanism described by model 2 might have been favored during evolution to lower intracellular methylglyoxal exposure.

2.4.3 Sensitivity analysis of the glyoxalase system

The dependency of methylglyoxal steady-state concentration on the activities of glyoxalase I and II was evaluated to investigate the difference of sensitivity when considering models 1 and 2 and to probe the potential of the inhibition or activation of these enzymes. An analogous analysis was done for the total glutathione concentration and methylglyoxal formation rate. Normalized derivatives of methylglyoxal steady-state concentration were also derived as functions of intermediate steady-state concentrations and parameters – equations 2.7.1 to 2.7.4 for model 1 and 2.8.1 to 2.8.3 for model 2 (sensitivity to methylglyoxal formation rate is not shown because it is very complex and no interpretation could be drawn from the corresponding equation). Figure 2.6 shows methylglyoxal steady-state concentration and methylglyoxal sensitivity as functions of the parameters of the system.

$$\frac{\partial \ln \mathrm{MGO}}{\partial \ln \mathrm{SH_{Total}}} = -\frac{\mathrm{SH_{Total}}}{\mathrm{GSH}} \qquad 2.7.1$$

$$\frac{\partial \ln \mathrm{MGO}}{\partial \ln \mathrm{V_{1A}}} = -\frac{\mathrm{V_{1A}}}{\mathrm{v_{in}}\ \mathrm{K_{m1}}}\left(\frac{\mathrm{v_r}}{\mathrm{v_f}}\mathrm{HTA} + \frac{\mathrm{HTA}^2}{\mathrm{GSH}}\right) \qquad 2.7.2$$

$$\frac{\partial \ln \mathrm{MGO}}{\partial \ln \mathrm{V_{2A}}} = -\frac{\mathrm{V_{2A}}\ \mathrm{SDLG}^2}{\mathrm{v_{in}}\ \mathrm{K_{m2A}}\ \mathrm{GSH}} \qquad 2.7.3$$

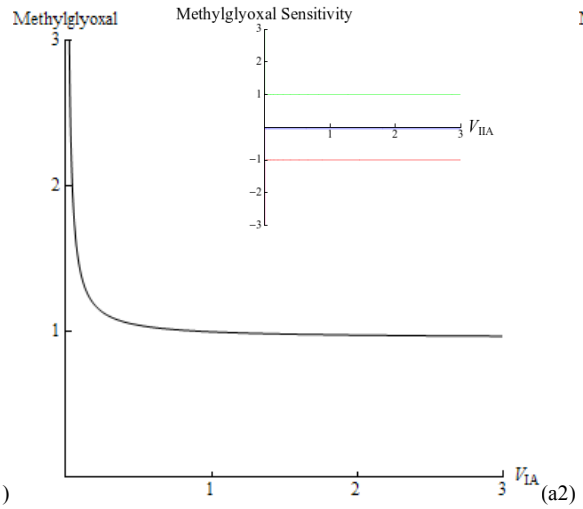
$$\frac{\partial \ln \mathrm{MGO}}{\partial \ln \mathrm{v_{in}}} = 1 + \frac{\mathrm{HTA} + \mathrm{SDLG}}{\mathrm{GSH}} + \frac{\mathrm{HTA}^2}{\mathrm{K_{m1}}\ \mathrm{GSH}} + \frac{\mathrm{SDLG}^2}{\mathrm{K_{m2A}}\ \mathrm{GSH}} + \frac{\mathrm{v_r}\ \mathrm{HTA}}{\mathrm{v_f}\ \mathrm{K_{m1}}} \qquad 2.7.4$$

$$\frac{\partial \ln MGO}{\partial \ln SH_{Total}} = \frac{SH_{Total}\, K_{m11}\, K_{m12}\, k_{-1} + SH_{Total}\, K_{m11}\, k_{-1}\, MGO}{k_1\, K_{m12}\, GSH^2\, MGO - k_1\, K_{m11}\, GSH\, MGO^2} \quad 2.8.1$$

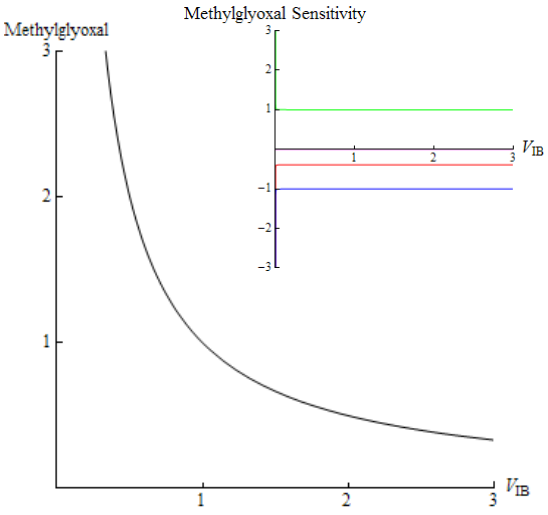
$$\frac{\partial \ln MGO}{\partial \ln V_{1B}} = -\frac{K_{m11}\, K_{m12}\, k_{-1} + K_{m12}\, k_{-1}\, GSH + K_{m11}\, G_2\, MGO + k_1\, K_{m12}\, GSH\, MGO + k_{-1}\, GSH\, MGO + k_1\, K_{m11}\, MGO^2 + k_1\, GSH\, MGO^2}{K_{m11}\, K_{m12}\, k_{-1} + K_{m12}\, k_{-1}\, GSH + k_1\, K_{m12}\, GSH\, MGO - k_1\, K_{m11}\, MGO^2} \quad 2.8.2$$

$$\frac{\partial \ln MGO}{\partial \ln V_{2B}} = -\frac{K_{m11}\, K_{m12}\, K_{m2}\, k_{-1}\, SDLG + K_{m11}\, K_{m2}\, k_{-1}\, MGO\, SDLG + K_{m11}\, K_{m12}\, k_{-1}\, SDLG^2 + K_{m11}\, k_{-1}\, MGO\, SDLG^2}{K_{m11}\, K_{m12}\, K_{m2}\, k_{-1}\, GSH + K_{m12}\, K_{m2}\, k_{-1}\, GSH^2 + k_1\, K_{m12}\, K_{m2}\, GSH^2\, MGO - k_1\, K_{m11}\, K_{m2}\, GSH\, MGO^2} \quad 2.8.3$$

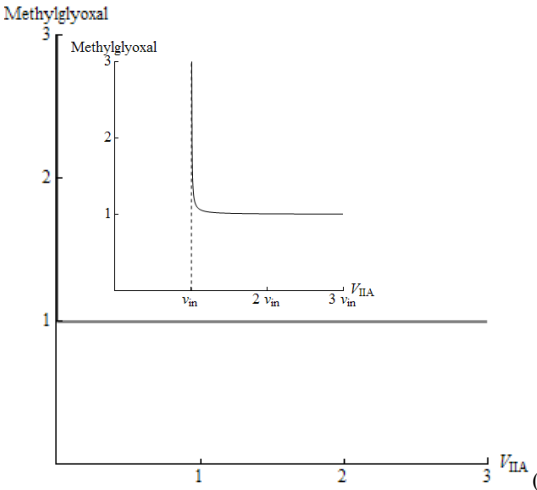
2. Steady-state symbolic analysis of the glyoxalase pathway



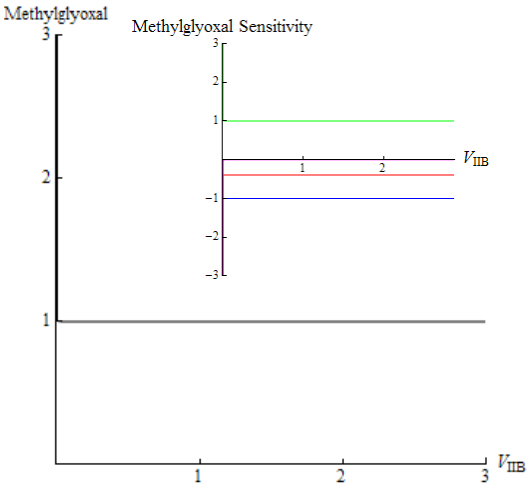
(a1)



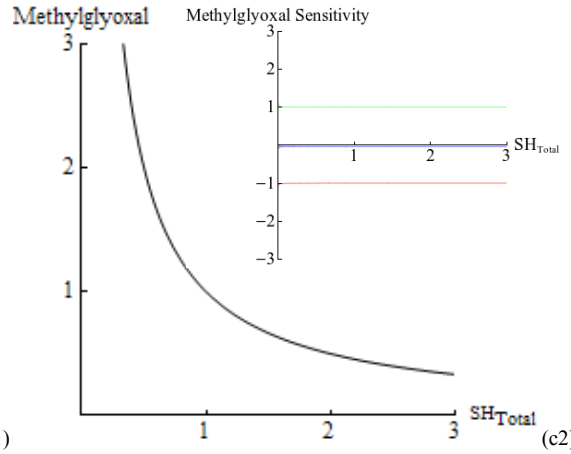
(a2)



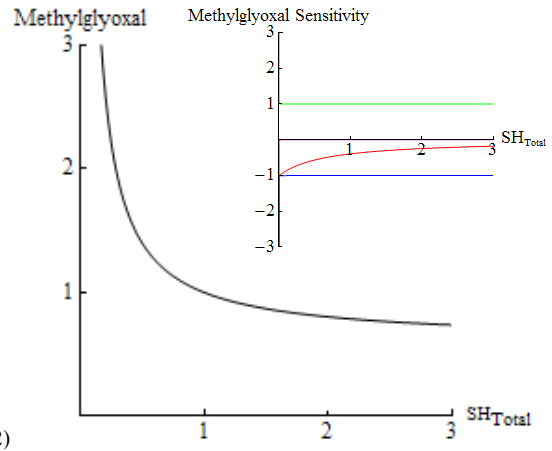
(b1)



(b2)



(c1)



(c2)

(continues on next page)

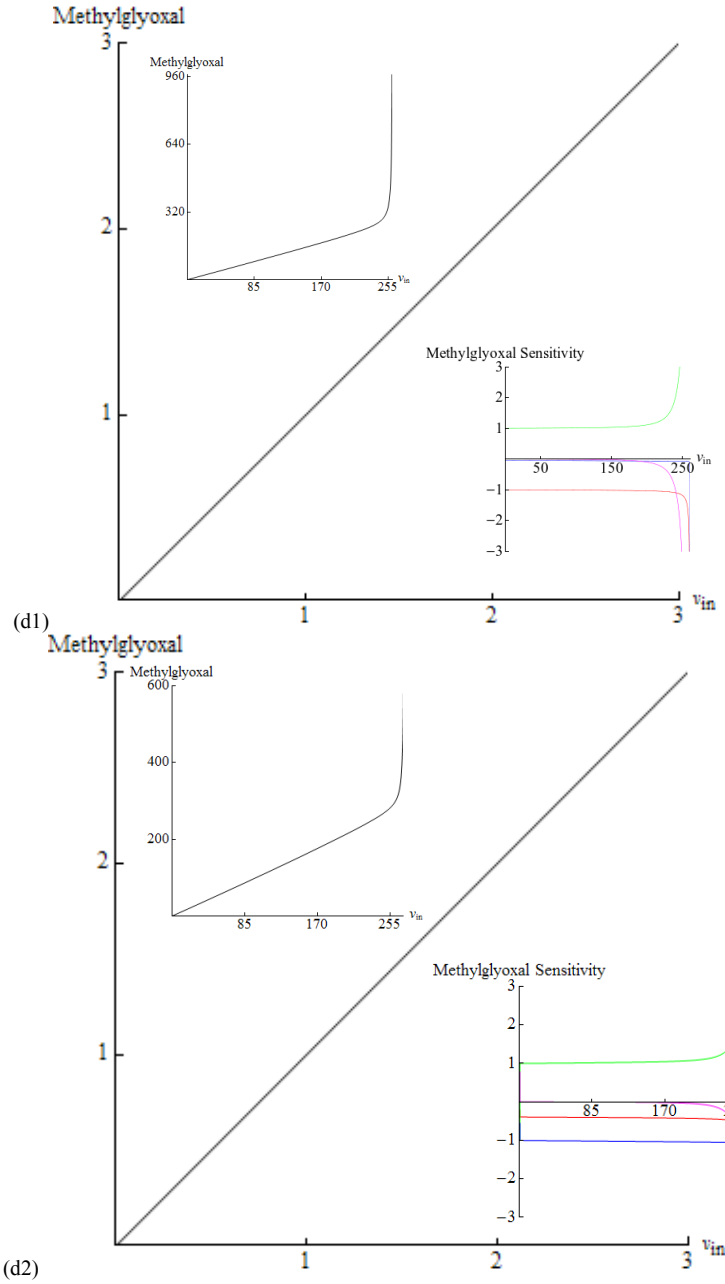


Figure 2.6 Dependence of methylglyoxal steady-state concentration on glyoxalase I activity for (a1) model 1 (with inserts representing the behavior near the lower limit for the achievement of a steady state and the normalized derivatives as functions of glyoxalase I activity) and (a2) model 2 (with the insert representing the normalized derivatives as functions of glyoxalase I activity); dependence of methylglyoxal steady-state concentration on glyoxalase II activity for (b1) model 1 (with inserts representing the behavior near the lower limit for the achievement of a steady state and the normalized derivatives as functions of glyoxalase II activity) and (b2) model 2 (with the insert representing the normalized derivatives as functions of glyoxalase II activity); dependence of methylglyoxal steady-state concentration on total glutathione concentration for (c1) model 1 and (c2) model 2 (with the normalized derivatives as inserts); dependence of methylglyoxal steady-state concentration on methylglyoxal formation rate for (d1) model 1 and (d2) model 2 (with inserts representing the behavior near the upper limit for the achievement of a steady state and the normalized derivatives as functions of methylglyoxal formation rate. The red lines represent methylglyoxal sensitivity to total glutathione changes, the green lines represent sensitivity to methylglyoxal influx and the blue and pink lines the sensitivity to glyoxalase I and II activities, respectively. The horizontal axes of the plots are normalized to the reference values for the parameters (table 2.1) and the concentration vertical axes are normalized to the reference methylglyoxal steady-state concentrations of each model (table 2.2).

According to model 1 MGO is very sensitive to total SH_{Total}. Indeed, it is approximately inversely proportional to this parameter within a wide variation range, including the reference state. MGO sensitivity to SH_{Total} changes is inversely proportional to GSH (equation 2.7.1), which is to say that changes in GSH when its value is low correspond to a larger variation of MGO. In the reference state, almost all glutathione is in the free form, so this sensitivity is approximately -1. The sensitivity to SH_{Total} becomes supralinear for extreme conditions at which this parameter has very low values (figures 2.6 c1 and c2) and the same happens for very high v_{in} (figures 2.6 d1 and d2). In model 2, MGO sensitivity to SH_{Total} is lower (-0.39) than for model 1. This is explained by the dependence on MGO in addition to GSH and by the squared GSH term in the denominator, which makes lower sensitivities expectable for high GSH.

Glyoxalase I and II activities have negligible effects on MGO in model 1 (-0.078 sensitivity for glyoxalase I and -6.2×10^{-6} for glyoxalase II). According to equation 2.7.2 this is explained by the very low HTA/K_{m1} and HTA/GSH ratios. The same effect is observed for methylglyoxal sensitivities to glyoxalase II activity and methylglyoxal formation rate (equations 2.7.3 and 2.7.4): in both cases, SDLG/GSH, HTA/GSH and the ratios of HTA and SDLG to the respective K_m's are very small, corresponding to a smaller sensitivity of MGO.

In model 2 sensitivity of MGO to glyoxalase I activity is related to GSH and MGO (equation 2.8.2). An important difference with respect to model 1 is that for model 2 MGO is very sensitive to glyoxalase I activity (-1.0) because the non-enzymatic reaction of methylglyoxal is a side pathway. However sensitivity to glyoxalase II activity is negligible (-2.4×10^{-5}) as in model 1. Sensitivity to glyoxalase II is higher for larger SDLG and smaller GSH. Since GSH is very high, this sensitivity is very low.

The sensitivity of MGO to methylglyoxal formation rate is the same in models 1 and 2 which is approximately 1. Sensitivity to v_{in} is larger for higher HTA and SDLG and smaller GSH. For model 1 equation 2.7.4 shows that the sensitivity to v_{in} is always higher than 1.

2.4.4 The system attains a steady state only if enzyme activities and total glutathione meet threshold values

The sensitivity of methylglyoxal steady state concentrations to each parameter increases severely when the parameter approaches critical values corresponding to graphic vertical asymptotes (plots a1, b1, c1 and d1 in figure 4). These values were algebraically derived for model 1 and found to be given by the right-hand members of inequalities 2.9 to 2.11, which indicate necessary conditions for the attainment of a steady state in the pathway.

$$SH_{Total} > \frac{K_{m1}\, v_{in}}{E_{1A}} + \frac{Km_{2A}\, v_{in}}{E_{2A}} \qquad 2.9$$

$$V_{1A} > v_{in}\, \frac{\tau_1 + K_{m1}\, E_{2A}}{\tau_1} \qquad 2.10$$

$$V_{2A} > v_{in}\, \frac{\psi_1 + K_{m2A}\, E_{1A}}{\psi_1} \qquad 2.11$$

Inequality 2.9 indicates that SH_{Total} must be higher than the sum of HTA and SDLG. This condition is stronger than the trivial condition which requires total cofactor concentration to be greater than zero. It means that free glutathione must exist in the system for methylglyoxal to be degraded. Therefore, this expression defines the catalytic amount of glutathione for the system to achieve a steady state. Inequalities 2.10 and 2.11 define critical values for enzyme activities. It is remarkable that it is not sufficient for these parameters to be higher than methylglyoxal influx. They actually have to exceed a value higher than methylglyoxal formation rate since this parameter multiplies by a factor greater than 1 in the right-hand member of both inequalities. Therefore, in this pathway, enzyme limiting rates cannot just be higher than or equal to the value of the flux of the pathway to attain a steady state.

Outside the space of parameters defined by inequalities 2.9 to 2.11 a steady state cannot be reached – methylglyoxal accumulates over time. Critical values of total glutathione concentration and glyoxalase I and II activities (given by inequalities 4.1, 4.3 and 4.4)

were found to be much lower than reference state (18,000 fold, 1,800 fold and 260 fold, respectively). The critical methylglyoxal formation rate is much higher than the reference (260 fold). This means that it is extremely difficult to for this system to reach conditions in which a steady state is not achieved and the cell cannot avoid the accumulation of methylglyoxal.

Threshold values also exist when model 2 is considered, however the meaning of the mathematical expressions is not trivial and there is not a correspondence with asymptotes.

2.5 Discussion

Two kinetic models for glyoxalase I have been previously considered: a one-substrate irreversible Michaelis-Menten equation where the methylglyoxal-glutathione hemithioacetal is the substrate and a two-substrate irreversible Michaelis-Menten equation where glutathione and methylglyoxal are the substrates that bind sequentially to the enzyme (Ponces Freire *et al.* 2003; Gomes *et al.* 2005b). Previous steady-state kinetic studies pointed to a more important contribution of the single-substrate hypothesis (Mannervik *et al.* 1973; Mannervik *et al.* 1974; Sellin and Mannervik 1983). Using reference parameters estimated from S-D-lactoylglutathione time-course data acquired until complete depletion of the methylglyoxal, the present work evaluates properties of the system derived from the two models and shows consequences of using these two models.

The present work is based on an algebraic derivation of methylglyoxal concentration instead of numerical simulations based on the ODE models. This methodology has important limitations as in generali it is not possible to derivate algebraically the steady-state concentrations for more complex pathways. Nevertheless this approach allowed a much more extensive (and faster) steady-state analysis and provided more complete information about the pathway. The mathematical definition of the necessary minimum amount of glutathione to achieve a steady state was derived and low-bound thresholds for enzyme activity, surprisingly higher than substrate influx, were found. As these thresholds are very low compared to the estimated in-cell enzyme activities and glutathione concentration no matter the model considered, cells seems to be well protected from reaching a situation in which the system wouldn't reach a steady state and methylglyoxal would accumulate continuously. These findings are especially interesting because they seem to be consequence of glutathione being both substrate and product of the same pathway. In that case, a threshold should be expected as well for any pathway where a cofactor is regenerated since free cofactor must be available to prevent accumulation of substrates; therefore its total concentration must exceed the sum of the free form and the intermediates of the pathway in between. Cases of other

such pathways are glycolysis and glycerol oxidation which comprise a turbo design where ATP is not just regenerated but net-produced (Haanstra *et al.* 2008).

Methylglyoxal sensitivities toward glyoxalase I and II activities and methylglyoxal input flux are concentration control coefficients in the framework of metabolic control analysis (Heinrich and Schuster 1996). For model 1, the control coefficients of the non-enzymatic formation and dissociation of hemithioacetal are -1 and 0, respectively, and the control coefficient of the methylglyoxal formation step is 1, so the enzyme-catalyzed reactions have null concentration control coefficients, satisfying the summation theorem of metabolic control analysis (Heinrich and Schuster 1996). On the other hand, in model 2 the methylglyoxal formation and the glyoxalase I reactions share the control of the pathway, having control coefficients of 1 and -1, respectively.

The growth curve of yeast glyoxalase-knock-out mutants is not very different from the wild type strain (Ponces Freire *et al.* 2003). Also, decreasing glutathione synthesis has a greater effect (Ponces Freire *et al.* 2003). If the glyoxalase system is to be targeted with the goal of disrupting the physiological steady state to achieve methylglyoxal accumulation, for instance in trypanosomatid cells, then the treatments must cause extreme variations of enzyme activities, which is an objective unlikely to be accomplished. Also in agreement with previous studies (Gomes *et al.* 2005b; Sousa Silva *et al.* 2005b), methylglyoxal concentration isn't significantly sensitive to variations of the activities of glyoxalase I or glyoxalase II (considering model 1). It is sensitive however to total glutathione concentration and methylglyoxal formation rate, so targeting glutathione metabolism seems to be a much better option as it is expected to have effect not only on methylglyoxal level but also on many other processes in which glutathione has a major role, such as oxidative stress defence (Jones 2008). In the case of trypanosomatids glutathione's role is generally played by trypanothione so targeting trypanothione-metabolism-exclusive enzymes appears to be an option as this is not expected to have effect on the human host (Muller *et al.* 2003). The glyoxalase pathway operates at high activities keeping virtually all intracellular glutathione in the free form. This is an important condition to ensure that the pathway is efficient in keeping a low methylglyoxal concentration: since the sensitivity of methylglyoxal concentration to total glutathione is always larger than -1 and inversely proportional to free glutathione,

it is very important that free glutathione is very high to keep this sensitivity as low as possible, i.e. approximately -1.

Drug discovery research has shifted from a single- to a multi-drug-target paradigm (Hornberg *et al.* 2007; Yang *et al.* 2008). Much of the effort to find single-effective drug targets has proved worthless (Sams-Dodd 2005). The design of biochemical networks is generally robust to inhibition or activation of a single enzyme (Kitano 2007). It is very unlikely that a significant outcome can be achieved on a biochemical network if only one of its nodes is targeted (Hopkins 2008). Solution for these problems lies in concertedly targeting a set of significant nodes of the biochemical network to achieve the desired effect. System-level mathematical models are crucial tools to identify this set of nodes and to predict the effect of putative therapeutic strategies. In this framework, the glyoxalase pathway may play an important role, if its inhibition is considered in the context of a broader strategy. Also the present work shows that the sensitivity of methylglyoxal concentration varies significantly with the assumed catalytic model for the glyoxalase system. This work is an example of how the predicted effect of changing a node of the metabolic network may depend extensively on the accurate characterization of the enzymes.

3 Model selection for the glyoxalase pathway using time-course kinetics

3.1 Abstract

Time-course S-D-lactoylglutathione kinetic data was used to discriminate among the two kinetic models proposed for glyoxalase II with the Akaike information criterion and residual analysis. The statistical analysis is based on residual square sum estimates of the kinetic parameters of the two models obtained with the Differential Evolution stochastic optimization algorithm implemented in the software S-timator. Despite the rich experimental data and the use of an algorithm that guarantees high-quality parameter estimates a discrimination of the models could not be accomplished with the Akaike scores, which were nearly identical for both models and different for permeabilized cells and purified enzyme data. The residual analysis points to a better performance of the two-substrate model. However, globally the discrimination was not conclusive, suggesting that an optimal experimental design method should be used to plan specific experiments for the purpose of model discrimination.

3.2 Introduction

Time-course measurement of concentrations of proteins, metabolites and RNA has become a vital strategy to understand the behavior of living systems. Mathematical modeling is an outstanding tool to extract information from such data. Models built from mechanistic interpretations of the biochemical processes are a convenient approach to study their properties and quantitatively predict their dynamics. Ordinary-differential-equation (ODE) systems are perhaps the most widely used framework to model the dynamics of biochemical pathways. However, there may be several plausible mechanistic interpretations for biochemical processes, each translated as a different ODE model. In these situations the selection of a particular mathematical model also has a correspondence with the selection of one of the possible descriptions of the studied process.

The choice of appropriate models to describe the kinetics of enzymatic reactions is also a major concern in the context of Systems Biology. Systems Biology uses mathematical models to study how biological systems behavior arises from the interaction of its components. Cell- and organism-level systems may be modeled employing ODE systems expressing the rates of the individual steps of the biochemical pathways. If these rate equations are not carefully chosen large deviations to the real behavior of the modeled system may occur. Even if such a deviation is not observed, wrong conclusions may be drawn from the analysis of the model, as the sensitivity analysis in the preceding chapter has shown.

Selecting an adequate ODE system to describe a biological network is not a trivial task. In accordance with the Principle of Parsimony, a model should be as simple as possible while capturing as much information as possible about the behavior of the real system. This means that an optimal trade-off between fitting to experimental data and mathematical structure simplicity should be sought while building a model of a biochemical system. Models should be chosen so that over-fitting does not occur, i.e. they do not contain superfluous parameters which do not have effect on the model output, and on the other hand they must describe the particular features of the modeled system faithfully.

Two kinetic ODE models have been suggested for the glyoxalase pathway with no conclusive selection of either, as both seem to contribute to the reaction in a certain unknown extension (Mannervik *et al.* 1972; Bartfai *et al.* 1973; Mannervik *et al.* 1973; Mannervik 1974; Mannervik *et al.* 1974; Aronsson *et al.* 1978; Ponces Freire *et al.* 2003; Gomes *et al.* 2005b; Mannervik 2008). The implications of the sensitivity analysis of each model for the usefulness of glyoxalase I in therapeutic strategies illustrate the importance of kinetic model selection in Biochemistry (as discussed in chapter 2). In the present chapter a methods from Statistics and Information Theory are used to clear out which of the proposed mechanisms describes most adequately the mechanism of the glyoxalase pathway. The results of the analysis are, however, inconclusive, therefore the usability of this methodology is discussed and a possible solution is suggested.

3.3 The Kullback-Leibler divergence and Akaike's information criterion

Statistics and Information Theory provide criteria to choose the best model among a set of candidates. A possible approach is to assume the existence of an unknown model that is able to describe the real biochemical system exactly (Burnham and Anderson 1998). The task of the modeler would be to approximate that 'true' model. Kullback and Leibler proposed a measure of the divergence between two models, i.e. of how good a model *f* approximates a model *g*:

$$I(f,g) = \int f(x \mid \theta_f) \ln \frac{f(x \mid \theta_f)}{g(x \mid \theta_g)} dx \qquad \text{inequality 3.1}$$

where *I* is the Kullback-Leibler distance of *g* from *f*, *x* is the vector of variables and θ_f and θ_g are the vectors of parameters of models *f* and *g* (Kullback and Leibler 1951; Burnham and Anderson 1998). It is noteworthy that the Kullback-Leibler divergence of *f* from *g* will be different from the divergence of *g* from *f*, as approximating *f* with *g* is not the same as approximating *g* with *f*.

The Kullback-Leibler distance cannot however be used directly as criterion to compare an ODE model with the true model if we do not pretend to know (or assume) something about the latter. In addition, even if it was possible to compare an approximating model to the true model, the Kullback-Leibler divergence would not be 0 since estimates of the parameters, and not their real values, would have to be used for the computation (Burnham and Anderson 1998). Akaike solved this problem by deriving from the Kullback-Leibler distance a measure of the distance between a postulated model and the true model without making assumptions about the true model (Akaike 1973; 1974; Burnham and Anderson 1998).

$$\mathrm{AIC} = -2 \ln\left(\mathrm{L}\left(\widehat{\theta} \mid y\right)\right) + 2K \qquad 3.2$$

In equation 3.2, *L* stands for a likelihood function where $\widehat{\theta}$ is the vector of estimates of the parameters and *y* the data from which the estimates were computed. *K* is the number

of parameters in the model. AIC can be determined without any knowledge about the true model. Furthermore, AIC accounts for a trade-off between bias, i.e. the deviation of the model from reality (which experimental data arises from) and the variance of the estimated model parameters (Burnham and Anderson 1998). A larger number of parameters in a model favors a decrease of the bias but increases the variance of the estimates of the parameters. In AIC this effect is accounted for as its value increases and the likelihood decreases with the number of parameters. Therefore, the use of AIC is expected to help avoiding over- and underfitting of the selected model. The search for the best bias-variance trade-off is also in accordance with the Principle of Parsimony, since in case more than one model seems to explain (fit) the observations equally well the simplest one (that is the one with less parameters) is chosen (Burnham and Anderson 1998). For a minimum-least-squares estimation with normally distributed errors, AIC is defined according to equation 3, where $\widehat{\varepsilon}_i$ are the estimated residuals for the model (Burnham and Anderson 1998).

$$\mathrm{AIC} = n \ln\left(\frac{\sum \widehat{\varepsilon}_i^2}{n}\right) + 2K \qquad 3.3$$

3.4 Kinetic assay of the glyoxalase pathway

Enzyme kinetic parameters were estimated from S-D-lactoylglutathione time-course data using the DE algorithm implemented in software S-timator (described in chapter 2). Assays were performed both with commercial yeast glyoxalases I and permeabilized yeast cells of the strain BY4741. The S-D-lactoylglutathione time-course data acquired with permeabilized cells from chapter 2 was used as reference for the analysis in the present chapter. Commercial purified yeast glyoxalase I (Sigma) was also assayed with the same setup, using four different concentrations.

3.5 Model parameter estimation

Two alternative mechanisms were considered for the glyoxalase I-catalyzed reaction – a one-substrate mechanism where a glutathione-methylglyoxal hemithioacetal is formed prior to binding the enzyme (figure 3.1 A), and a two-substrate mechanism where glutathione and methylglyoxal sequentially bind to glyoxalase I directly (figure 3.1 B). A two-substrate sequential model for glyoxalase I is mathematically translated by the expression

$$v = \frac{V_1 \, GSH \, MGO}{K + K_{m12} \, GSH + K_{m11} \, MGO + GSH \, MGO} \qquad 3.4$$

This expression was however simplified to

$$v = \frac{V_1 \, GSH \, MGO}{(K_{m12} + GSH)(K_{m11} + MGO)} \qquad 3.5$$

by assuming

$$K = K_{m11} \, K_{m12} \qquad 3.6$$

since while estimating the parameters in the former expression an identifiability problem was evident – the parameter K could not be uniquely estimated. This simple approximation allowed the suppression of this futile parameter.

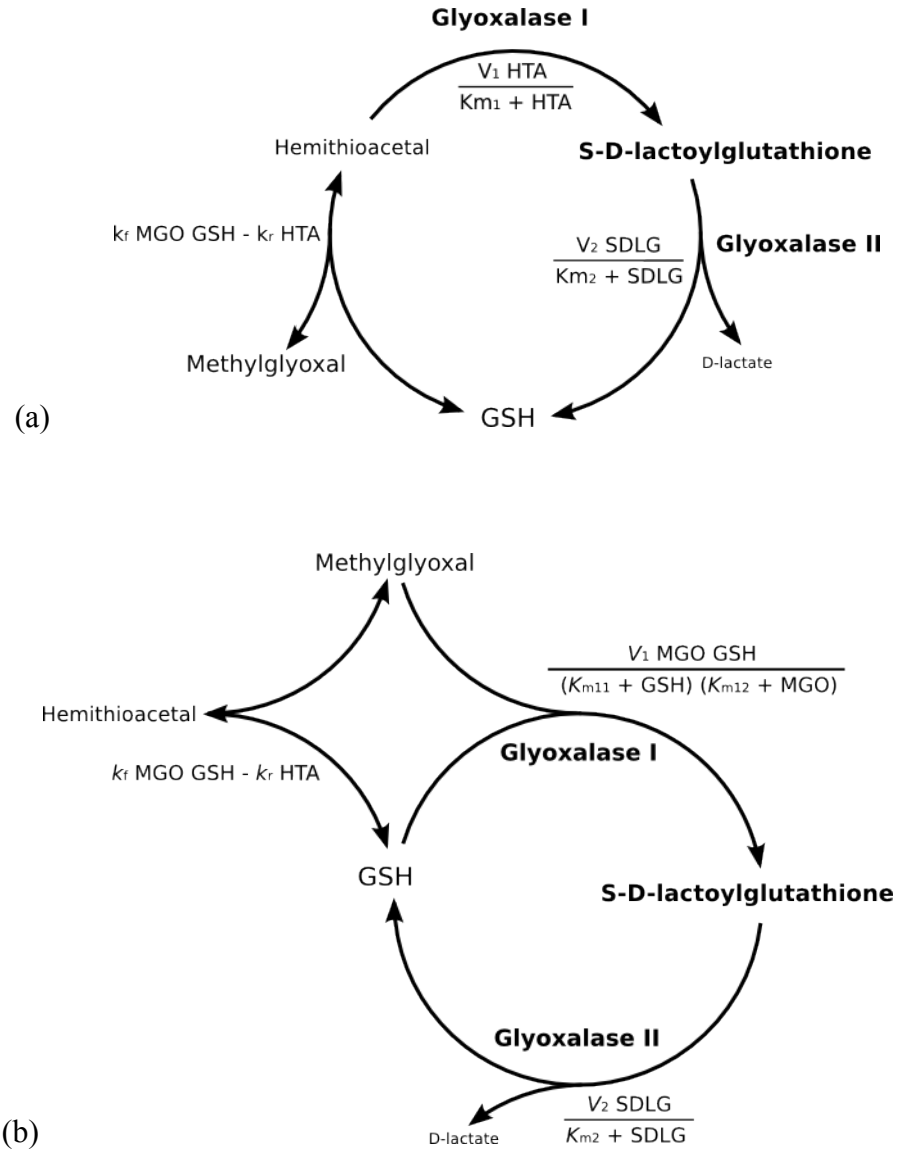


Figure 3.1 Alternative kinetic models for the glyoxalase pathway. In (a) the substrate of glyoxalase I is a glutathione-methylglyoxal hemithioacetal while in (b) glutathione and methylglyoxal are direct substrates of the enzyme.

The parameters of models 1 and 2 were estimated using the two sets of data, acquired with commercial enzyme and permeabilized cells.

3.6 Results and discussion

3.6.1 Parameter estimation from concentration time-course data

Two experimental sets of S-D-lactoylglutathione concentration time-courses were obtained for estimation of the parameters of kinetic models 1 and 2. A first one using permeabilized cells (figure 3.2) and a second one using commercial purified glyoxalase I (figure 3.3). The digitonin-permeabilized cells setup preserves in-cell protein concentrations; therefore this experimental setup is closer to *in vivo* conditions, at least concerning protein-protein interactions. While proteins are kept inside, small molecules diffuse across the permeabilized membrane, so there is full control over substrate concentration. On the other extreme the purified enzyme is free from molecular interactions other than with the substrates and product.

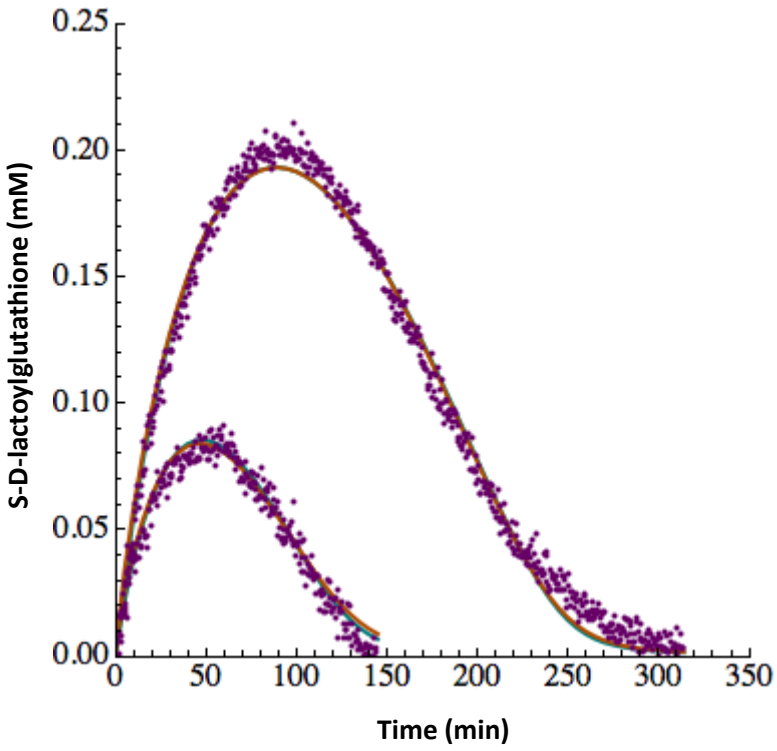


Figure 3.2 Experimental data (magenta) and fitting of models 1 (cyan) and 2 (orange). Glutathione and methylglyoxal initial concentrations were 1.00 and 0.33 mM for the experimental data set 1 and 0.66 mM for both substrates.

Since the BY4741 strain also has active glyoxalase II activity, this enzyme needed to be included in the model to explain the permeabilized cell data. According to previous studies, a single-substrate irreversible Michaelis-Menten equation is adequate to describe the kinetics of the glyoxalase II-catalyzed hydrolysis of S-D-lactoylglutathione (Uotila 1973; Principato *et al.* 1985; Talesa *et al.* 1989; Allen *et al.* 1993; Bito *et al.* 1997; Maiti *et al.* 1997; Martins *et al.* 1999; Martins *et al.* 2001b; Irsch and Krauth-Siegel 2004; Sousa Silva *et al.* 2005a). This equation was therefore adopted and, as expected, the models including this equation consistently fitted the data, giving similar estimates for the activity of glyoxalase II in the mixture and its K_m: 3.31×10^{-3} mM/min and 0.0286 mM for model 1 and 3.45×10^{-3} mM/min and 0.0347 mM for model 2 (activities not normalized to cell volume, unlike in chapter 2).

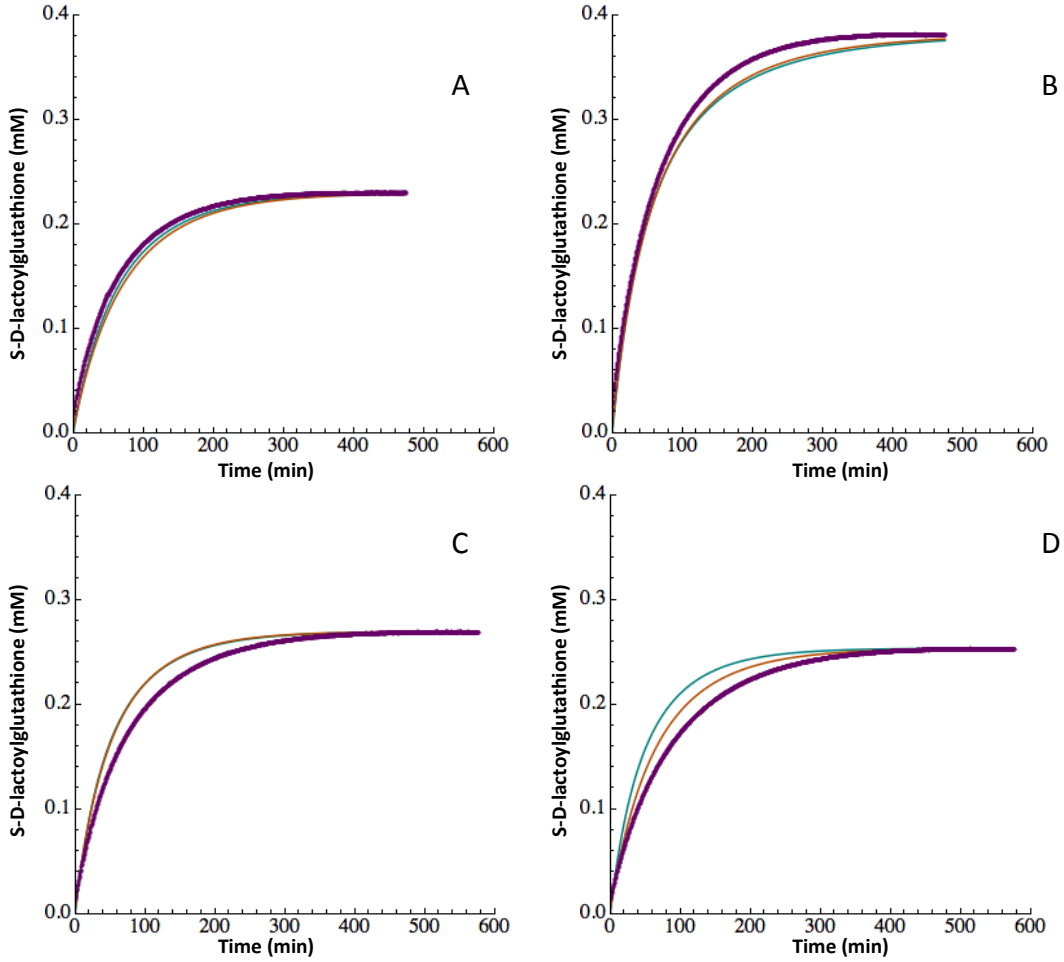


Figure 3.3 S-D-lactoylglutathione concentration time-course data acquired with commercial yeast glyoxalase I (magenta), and corresponding predictions for models 1 (cyan) and 2 (orange). A – 0.229 mM methylglyoxal, 0.571 mM glutathione and 3.23×10^{-6} mM glyoxalase I; B – 0.535 mM methylglyoxal, 0.381 mM glutathione and 4.03×10^{-6} mM glyoxalase I; C – 0.269 mM methylglyoxal, 0.526 mM glutathione and 4.84×10^{-6} mM; D – 0.538 mM methylglyoxal, 0.253 mM glutathione and 3.63×10^{-6} mM glyoxalase I.

The simplification of model 2, where the equality $K = K_{m11}\ K_{m12}$ was assumed, proved to be justified as several equally low local minima were found during the optimization, for both purified enzyme and permeabilized cells data sets. The estimates for models 1 and 2 are outlined in table 3.1 for the assays with permeabilized cells and in table 3.2 for the assays with purified enzyme. The sum of the squared residuals for each set of estimates and the AIC scores for each model and experimental setup are also shown.

Table 3.1 DE estimates for the parameters of models 1 and 2 obtained with permeabilized yeast cells and relevant scores for model discrimination.

Model	**Permeabilized cells**			**RSS** (mM^2)	**AIC**
	Parameter				
		Present work	Literature		
Model 1	V_1 (mM/min)	0.0239	-	0.0386	-9217
	K_{m1} hemithioacetal (mM)	0.214	0.62*		
Model 2	V_1 (mM/min)	0.135	-	0.0407	-9168
	K_{m11} glutathione (mM)	2.569	1.64**		
	K_{m12} methylglyoxal (mM)	1.202	3.56**		

*Martins *et al*. 2001
**Gomes *et al.* 2005

Table 3.2 DE estimates for the parameters of models 1 and 2 obtained with commercial yeast glyoxalase I and relevant scores for model discrimination.

	Commercial glyoxalase I		**RSS** (mM^2)	**AIC**
	Parameter			
	Present work	Literature		
k_{cat1} (min^{-1})	8586	35400*	0.5097	-37971
K_{m1} hemithioacetal (mM)	0.223	0.51*		
k_{cat1} (min^{-1})	17046	-	0.4782	-38233
K_{m11} glutathione (mM)	0.875	-		
K_{m12} methylglyoxal (mM)	1.178	-		

*Marmstal *et al.* 1979

The estimates for the parameters of models 1 and 2 obtained in the present work are within the order of magnitude of the reference values from the literature. The estimates for the K_m's with the purified enzyme and the permeabilized cells are very similar for the hemithioacetal in model 1 and for methylglyoxal in model 2. Though the difference is larger for the K_m's for glutathione in model 2, the estimates are within the same order of magnitude as well (tables 3.1 and 3.2).

3.6.2 Multimodality, epistasis and identifiability

The estimation of the parameters of a model describing a biochemical system is a difficult task, even if the system is a small one such as the glyoxalase pathway. Two major problems in model parameter estimation are objective function multimodality (Polisetty *et al.* 2006; Rodriguez-Fernandez *et al.* 2006) and parameter epistasis (Gutenkunst *et al.* 2007). Multimodality exists when there are several objective function local minima. To deal with multimodal objective functions a stochastic optimization method should be used, as deterministic methods would converge to the minimum nearest an initial guess that might not be the global minimum (Nocedal and Wright 1999; Price 1999). On the other hand, epistasis means that if two or more parameters are perturbed at once the effect of the variation of each perturbed parameter is different from varying that parameter alone (Seo *et al.* 2004). In models of biological systems epistasis seems to be nearly universal as the variation of a set of parameters in some directions of the parameter space generally has a small effect on the predictions of the models (Gutenkunst *et al.* 2007). This is a plausible explanation for the difficulty in estimating the set of parameters collectively from a set of experimental data. Epistasis does not imply however large variances associated to the estimates of the parameters since the variance of any parameter relates to the effect of the variation of that parameter alone (Gutenkunst *et al.* 2007).

The fit of a two-substrate sequential mechanism rate (equation 3.4) equation stalled at practical identifiability problems, meaning that unique optimal parameter estimates for such model were not found for both data sets (Jacquez and Perry 1990). Indeed different values of K would give equally good objective function evaluations in different regions of the parameter space. For this reason the approximated two-substrate equation 3.5 was considered under the assumption of equation 3.6. This proved to be an acceptable approximation since a unique optimal parameter set could be found. The approximated two-substrate equation also has the advantage of having parameters with a straightforward biological interpretation identical to the corresponding parameters of the original equation.

3.6.3 Model selection for glyoxalase I

For permeabilized cells a difference of 49 (0.05%) was observed between the AIC scores, while for the purified enzyme a difference of 262 (0.7%) was observed. The AIC analysis is inconclusive since the difference between the scores is very small and contradictory for the two sets of data – for the permeabilized cells the difference favors the single-substrate model and for the purified enzyme the simplified sequential two-substrate model is favored. A residual analysis was performed to assess the difference between the fittings of the models to the experimental data (figure 3.4).

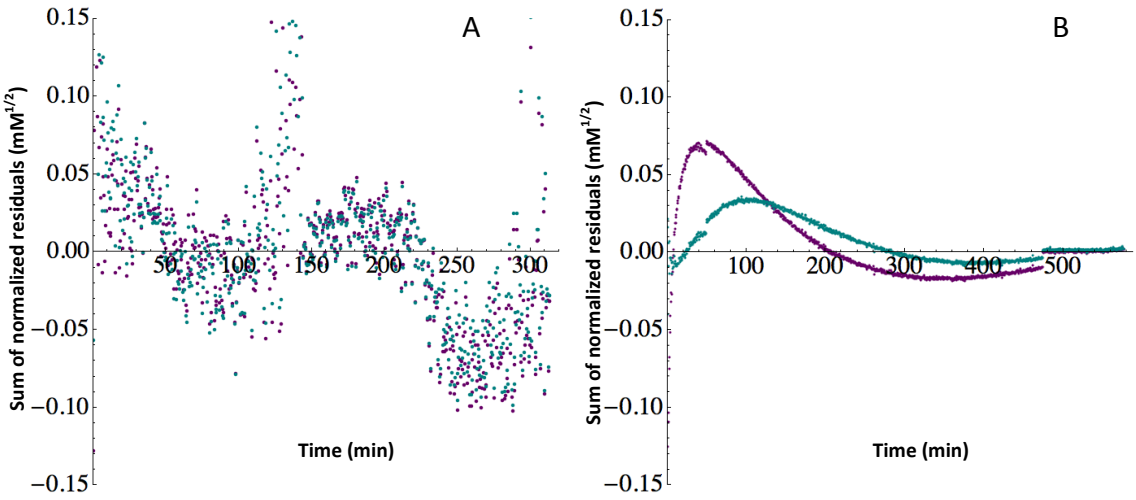


Figure 3.4 Residual plots for model 1 (magenta) and 2 (cyan) for the (A) permeabilized cell and the (B) purified enzyme assays. The residuals are normalized to the squared roots of the values of the experimental time courses and summed up at each instant.

The residual analysis shows a the better performance of model 2 for the data set obtained with purified enzyme – the residuals are smaller for this model and though a systematic bias is observed for both models, it is smaller also for model 2. On the other hand the residual plot for the permeabilized cells data does not show a difference between the two models. Although not so evident, a systematic bias is also observable for the two models with respect to this data.

To conclude, it is not possible to discriminate which of the considered models should be selected to describe the kinetics of glyoxalase I from this experimental data. Despite the

use of a quite complete set of experimental conditions, with time courses covering a vast range of substrate concentrations, and the employment of novel stochastic method for the estimation of the kinetic parameters, no definitive conclusion could be drawn from the determination of the AIC scores or the residual analysis. The glyoxalase I model selection problem could not be tackled by detailed analysis of a large set of kinetic data alone. A way to deal with such a difficult discrimination problem may be to design an experiment optimized to this specific purpose. In chapter 4, a method is described to optimize a time-course experiment to address kinetic model discrimination problems.

4 Optimization of experimental design for kinetic model discrimination

4.1 Abstract

A method is proposed to optimize the design of enzyme time-course assays aiming at selecting a model among a set of candidates. The method searches for the experimental conditions that maximize the Kullback-Leibler distance (a measure of discrepancy) between the models, using the generalized differential evolution algorithm for multiobjective optimization problems. The glyoxalase I enzyme, for which two kinetic models are usually considered, is used to test the method in a concrete problem. The suggested computational procedure proves to give conditions, namely initial substrates concentrations, for which the predictions of the models are dramatically different and therefore is useful to plan experiments for model selection, at least for the case in which two models are under consideration.

4.2 Introduction

In biochemical research multiple mathematical models may be considered to describe a biochemical pathway. In enzyme kinetic studies in particular if the actual enzyme catalysis mechanism is uncertain, several kinetic equations may be considered, each corresponding to a different putative mechanism. The selection of the appropriate kinetic model for a particular reaction is important in at least two dimensions: it emerges from a mechanistic interpretation so its selection reflects the choice of a particular mechanism; and it should be useful to predict the dynamics of the reaction or, at higher levels, the behavior of pathways or larger biochemical networks. In the latter case, the choice of an inappropriate kinetic model might give unrealistic results and render a theoretical analysis of the biochemical system useless. As explained in chapter 2, the glyoxalase pathway is an example of how the choice of different kinetic equations might lead to very different conclusions about the response of even simple pathways to perturbation of their steady state.

In the present chapter a procedure is presented to optimize time-course experiments so that the divergence between the time courses predicted by the models under consideration is maximized. In these conditions the relative competence of the candidate models in describing the experimental data (measured with appropriate statistical criteria) should be clearer. The described procedure is used to find the optimal experimental conditions to select a kinetic model for the glyoxalase I enzyme. As explained in chapters 2 and 3, two different kinetic models are usually employed to describe the kinetics of this enzyme, so this is an appropriate biochemical system to test the procedure proposed herein. The substrate concentrations at time-zero are among the most obvious and easiest experimental variables to be controlled by the user. For this reason, to exemplify the use of the suggested procedure, the question of which substrate and enzyme concentrations should be used in a kinetic assay so that the difference between the outputs of the models were maximum is addressed.

4.3 Optimization criteria

Kullback-Leibler distance

The method presented here was designed under the assumption that at the moment it is used every candidate model is equally regarded as possible. For this reason a strategy was sought to find the experimental conditions for which the divergence between the any pair of candidate models simultaneously, easing the selection of one of the models. The Kullback-Leibler distance is a measure of the divergence between two models, i.e. how good a model *g* is in approximating a model *f* and is defined as

$$I(f,g) = \int f(x \mid \theta_f) \ln \frac{f(x \mid \theta_f)}{g(x \mid \theta_g)} dx \qquad 4.1$$

for probability distribution models, where x is the vector of observable variables, θ_f is the vector of parameters of *f* and θ_g is the vector of parameters of *g*. The Kullback-Leibler divergence was generalized to be used in the space of positive functions (Amisaki and Eguchi 1995) according to the expression

$$I(f,g) = \int f(x \mid \theta_f) \left(\ln \frac{f(x \mid \theta_f)}{g(x \mid \theta_g)} + \frac{g(x \mid \theta_g)}{f(x \mid \theta_f)} - 1 \right) dx \qquad 4.2$$

In the case of ODE kinetic models, x is time and *f* and *g* are the time courses predicted by the models for the experimentally measurable intermediates.

$$I(f,g) = \int f(t \mid \theta_f) \left(\ln \frac{f(t \mid \theta_f)}{g(t \mid \theta_g)} + \frac{g(t \mid \theta_g)}{f(t \mid \theta_f)} - 1 \right) dt \qquad 4.3$$

The Kullback-Leibler distance is used here as the measure of distance between the candidate models – the experimental conditions that maximize the Kullback-Leibler distance between the time courses predicted by every pair of candidate models *f* and *g*, in both directions, are the ones considered to be optimal for an experiment aiming the selection of one of the candidate models.

Pareto optimality

For discrimination between two models, the Kullback-Leibler distance has to be maximized in both directions, maximizing the divergence of model A from model B and vice-versa. This means that there are two objective functions to be maximized in the simplest two-model case. If more than two models are considered, the Kullback-Leibler distance must be maximized between every pair-wise combination of in both directions, as figure 4.1 illustrates.

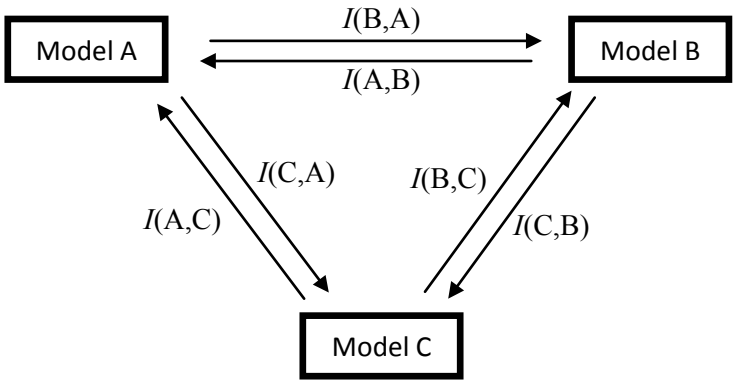


Figure 4.1 Maximization of the Kullback-Leibler distance between pairs of candidate models, for an example set of 3 candidate models. Conditions are sought maximizing the Kullback-Leibler distance (*I*) between any two models, in both directions. At Pareto-optimal solutions one of the *I*(i, j) cannot be increased without decreasing another *I*(x, y).

A possible way to solve to this kind of problems is to assign different weighting factors to each objective according to their relative importance and use a single-objective optimization algorithm. An obvious shortcoming of this approach is that in this case objectives have equal since all candidate models should be tested in conditions which do not favor any model selection from the beginning. In opposition to single-objective optimizations, where a single optimal solution is sought, for multiobjective optimization problems several solutions of equivalent quality can be found, each with different trade-offs regarding the scores for the different objectives. The set of the optimal solutions for a multiobjective optimization problem is named Pareto front, after Vilfredo Pareto, who generalized this concept of optimality (Pareto 1896) originally proposed by Francis Ysidro Edgeworth (Edgeworth 1881; Coello Coello *et al.* 2007). The approximation of

solutions to the Pareto front may be compared by the dominance criterion, defined in definition A.

Definition A

A solution u dominates a solution v, i.e. $v \prec u$ if and only if for the vector of objective functions $f = (f_1, f_2, ..., f_i, ...)$

$$\forall i : f_i(v) \leq f_i(u) \quad \wedge \quad \exists i : f_i(v) < f_i(u)$$

.

The optimal solutions are those for which the score of one objective cannot be improved without decreasing the score of another objective(s). Therefore optimal solutions are non-dominated. In practice optimal solutions are very hard to determine so the optimization process usually gives approximations to the Pareto front (Kukkonen and Lampinen 2005). The user should choose among the solutions given by the algorithm the ones that suit his purpose best.

For the present work Pareto optimality was therefore chosen as the optimization goal and the Kullback-Leibler distance as the optimization criterion.

4.4 Multiobjective optimization

Generalized differential evolution

Differential evolution, the single-objective optimization algorithm described in section 2.3.2, was extended to solve multiobjective problems. Among the different suggested multiobjective versions of the differential evolution algorithm (Abraham *et al.* 2005) the one named Generalized Differential Evolution, despite a remarkable simplicity, seems to perform very well in a variety of problems both regarding computation time and distribution of the final solution set near the Pareto front (Kukkonen and Lampinen 2007). Also it has the advantage of reducing to the single-objective version for one-objective problems. After two revisions, it is presently in its 3rd generation (GDE3) (Kukkonen and Lampinen 2005).

Initialization of the population, mutation and crossover in the GDE3 algorithm occur in the same way as in the single-objective differential evolution. The differences start in the selection process which in GDE3 is based on dominance defined in definition A: if a new solution vector dominates the target vector, the latter is replaced by the former in the new population. In the first version of the multiobjective algorithm, GDE1, this was the solo criterion for selection (Kukkonen and Lampinen 2004b; 2005). In the second generation of the algorithm, when the new and the old candidate solutions were non-dominated with respect to each other, the solution for the next generation would be selected with basis on a measure of crowdedness, i.e. the proximity to other solutions in the space of the objectives. The less crowded of the two would be selected (Kukkonen and Lampinen 2004a; 2005). This addition allowed a better spread of the distribution of the candidates in the proximity of the Pareto front. In GDE3 when the two solutions are non-dominated both are saved so after evaluation of a set of new solutions the dimension of the population usually increases. In this case the solutions of this new population are sorted with basis on dominance. Then the most crowded solutions of the worst non-dominated group of the population are removed to restore the initial size of the population (Kukkonen and Lampinen 2005). As the authors of the algorithm note the procedure for removal of the most crowded vectors should be carefully chosen to ensure a better distribution of the selected solutions in the objective function space.

Pseudo-code for GDE3, as implemented in the present work, may be found in figure 4.2.

Input: D, G_{max}, $NP \geq 4$, $F \in [0, 1+]$, $CR \in [0, 1]$ and $\vec{x}^{(lo)}$, $\vec{x}^{(hi)}$

Initialize P_G: $\left\{\vec{x}_{1,G}, \vec{x}_{2,G}, ..., \vec{x}_{NP,G}\right\}$:

$G = 0$, $m = 0$

$\forall\, i \leq NP \wedge \forall\, j \leq D$: $x_{j,i,G=0} = x_j^{(lo)} + rand_j[0,1] \cdot \left(x_j^{(hi)} - x_j^{(lo)}\right)$

While $G < G_{max}$ or termination criteria not met:

$\forall\, i \leq NP$:

Mutate and recombine (DE/rand/1/Bin):

$r_1, r_2, r_3 \in (1, 2, \ldots, NP)$ randomly chosen with $r_1 \neq r_2 \neq r_3 \neq i$

$j_{rand} \in (1, 2, \ldots, D)$ randomly chosen once for each i

$\forall\, j \leq D$, $u_{j,i,G+1}$:

$u_{j,i,G+1} = x_{j,r_3,G} + F \cdot \left(x_{j,r_1G} - x_{j,r_2G}\right)$ if $rand_j[0,1] < CR \vee j = j_{rand}$
or $x_{j,i,G}$ otherwise

Select $\vec{x}_{i,G+1} = \vec{u}_{i,G+1}$ if $\vec{u}_{i,G+1}$ dominates $\vec{x}_{i,G+1}$ or
$\vec{x}_{i,G+1} = \vec{x}_{i,G}$ otherwise

$m = m + 1$ and $\vec{x}_{NP+m,G+1} = \vec{u}_{i,G+1}$ if $\vec{x}_{i,G}$ does not dominate $\vec{u}_{i,G}$
or vice-versa

While $m > 0$:

Select $\vec{x} \in \left\{\vec{x}_{1,G+1}, \vec{x}_{2,G+1}, ..., \vec{x}_{NP+m,G+1}\right\}$:

$\forall\, i$ $\vec{x}$ is not dominated by $\vec{x}_{i,G+1}$

and

$\forall\, (\vec{x}_{i,G+1} : \vec{x}_{i,G+1}$ is not dominated by $\vec{x})$ $CD(\vec{x}) \leq CD(\vec{x}_{i,G+1})$

Remove $\vec{x}$

$m = m - 1$

$G = G + 1$

Figure 4.2 Pseudo-code for the Generalized Differential Evolution, generation 3 as implemented in the present work. *CD* means crowding distance, *G* means generation and G_{max} is the number of maximum generations, *NP* is the number of solution vectors in the population.

In this specific implementation GDE3, optimizations are always strictly constrained so only solution-vectors within the user-defined boundaries are allowed – if a vector does not lie within the boundaries, other vector is generated.

Non-dominated sorting

As in other multiobjective evolutionary algorithms, the most time-expensive step of GDE3 is the sorting of the non-dominated solutions, eventually surpassed by the evaluation of the objective functions only (Kukkonen and Lampinen 2007; Fang *et al.* 2008). For this reason the development of fast sorting methods for evolutionary algorithms has received considerable attention.

GDE3 was implemented by its authors with use of a non-dominant sorting algorithm based on a measure of *crowding distance* developed for multiobjective evolutionary algorithms in general and especially for the second version of the non-dominated sorting genetic algorithm (NSGA-II) (Deb *et al.* 2002). In NSGA-II the population size doubles before selection of the solutions that integrate the following generation, so in this case sorting of the solutions is especially critical for the total computational cost (Deb *et al.* 2002). The original sorting method implemented in NSGA-II can be summarized as follows (Fang *et al.* 2008):

1. Remove the first solution s_1 from the population P and put it into an empty front F.

2. Compare the second solution s_2 in P with s_1. If s_1 dominates s_2, s_2 remains in P and go to next step; otherwise, remove s_2 from P and put it in F. If s_1 is dominated by s_2, remove s_1 from F and put it back in P; otherwise, s_1 and s_2 are non-dominated.

3. Compare in succession s_i ($i = 3, 4, \ldots, NP$) in P with all solutions in F. If solution s_i is dominated by a solution in F, s_i remains in P. Any solutions in F dominated by s_i are removed from F and put back in P. If s_i is not dominated after all the comparisons, s_i is removed from P and put in F.

After filling this first front F the solutions remaining in P are distributed in subsequent fronts by the same procedure.

This sorting procedure requires many and repeated comparisons between solutions making it quite time-consuming (Fang *et al.* 2008). An algorithm based on an older procedure capable of generating the first non-dominated front (Kung *et al.* 1975) was later suggested to replace the original NSGA-II sorting method, claiming to have a smaller time complexity (Jensen 2003). Recently a new sorting algorithm was proposed which makes use of a *divide-and-conquer* mechanism and a data structure called *dominance tree* which improves the computational and memory efficiency of the sorting process when compared to older algorithms (Fang *et al.* 2008). This algorithm not only avoids repetition but also minimizes the needed number of comparisons among solutions to generate the non-dominated fronts by saving the information about previously done comparisons, which makes it faster. For example, if a solution A was found to dominate a solution B, if a solution C is compared with B and found to be dominated by the latter, then it will be not compared with A since it is surely also dominated by this solution. This is achieved through the use of a tree hierarchical data structure built with the aforementioned divide-and-conquer mechanism. Also, in case two solutions are identical, they are placed in two different non-dominated fronts, which Jensen's algorithm failed to do (Jensen 2003; Fang *et al.* 2008).

The dominance tree data structure's basic unit is a node corresponding to a solution. The nodes are interconnected through dominance or non-dominance relationships. Connected nodes with a non-dominance relationship are called siblings. Nodes with a dominance relationship are named parent and child nodes, being the child node the one containing the dominated solution. A child node cannot dominate its parent's siblings; however the parent's siblings do not have to dominate the child.

The divide-and-conquer method on the basis of the construction of the tree initiates by consecutively *dividing* the entire population in halves, so that each node contains a single solution. Then the individual nodes are successively compared pair-wise and consecutively merged (*conquered*) according to their dominance relationship until the dominance tree is complete. The algorithm is recursive in both phases as the result of dividing or merging nodes is immediately used for another dividing or merging event, respectively. The pseudo-code for the divide-and-conquer non-dominated sorting algorithm is given in figure 4.3.

```
Function GetDominanceTree(NodeList, Size):

    If Size > 1:
        LeftTree = GetDominanceTree(NodeList, Size/2)
        RightTree = GetDominanceTree(NodeList + Size/2, Size - Size/2)

    Else:
        Return NodeList

    Return MergeDominanceTrees(LeftTree, RightTree)


Function MergeDominanceTrees(LeftTree, RightTree):

    LeftNode = LeftTree.Root
    RightNode = RightTree.Root

    While LeftNode ≠ null and RightNode ≠ null:

        If LeftNode dominates RightNode:
            TemporaryNode = RightNode
            RightNode = RightNode.NextSiblingNode
            RightTree.Remove(TemporaryNode)
            MergeDominanceTrees(LeftNode.DominatedChildren, TemporaryNode)

        Elif RightNode dominates LeftNode:
            TemporaryNode = LeftNode
            LeftNode = LeftNode.NextSiblingNode
            LeftTree.Remove(TemporaryNode)
            MergeDominanceTrees(RightNode.DominatedChildren, TemporaryNode)

            If LeftTree = null:
                LeftTree = RightTree
                RightTree = null

        Else:
            RightNode = RightNode.NextSiblingNode

            If RightNode = null:
                RightNode = RightTree.Root
                LeftNode = LeftNode.NextSiblingNode

    LeftNode = LeftTree.Root
    RightNode = RightTree.Root

    While RightTree ≠ null:
        LeftNode.AddSiblingNode(RightNode)
        TemporaryNode ← RightNode
        RightNode ← RightNode.NextSiblingNode
        RightTree.Remove(TemporaryNode)

    Return LeftTree
```

Figure 4.3 Pseudo-code for the divide-and-conquer algorithm by Fang *et. Al* (2008) as implemented in the present work, used to sort solutions according to dominance relationship.

After sorting the non-dominated solutions in non-dominated fronts, the most crowded solutions of the last non-dominated front have to be removed from the population to restore its original size, *NP*. Deb, Jensen and Kukkonen and Lampinen recommend niching techniques based on a measure of the crowding distance and the *k*-nearest neighbor method (Cover and Hart 1967) instead of using a clustering algorithm to find the solutions to be removed, arguing that clustering techniques have a higher time complexity (Deb *et al.* 2002; Jensen 2003; Kukkonen and Lampinen 2005; Kukkonen and Deb 2006).

Removal of most crowded non-dominated solutions

In the present implementation of the GDE3 the *k*-nearest neighbor method (Cover and Hart 1967; Kukkonen and Lampinen 2005) was used to identify the most crowded solutions in the last non-dominated front. Identification and removal of the most crowded solution repeats until the original population size, *NP*, is restored, as explained in detail in figure 4.4.

1. Select solutions giving extreme (maximum or minimum) values for each objective function.
2. If adding these solutions restores *NP*, crowding-distance selection stops here and GDE3 moves to the next generation. If by adding these solutions the number of solutions in the population exceeds *NP*, discard the remaining solutions not corresponding to extreme objective values and move to 3. Otherwise extreme-objective solutions are saved and steps 3 to 6 apply to non-extreme solutions until the population has size *NP*.
3. Compute the Euclidean distance between the objective vectors corresponding to the solutions.
4. For each solution, sort the computed objective-space Euclidean distances from the smallest to the largest.
5. For each solution, sum-up the first *k* smaller Euclidean distances and discard the solution with the smallest sum.
6. While the size of the population is larger than *NP*, repeat steps 3 to 5.

Figure 4.4 Crowding-distance-based selection of solutions from the last non-dominated front using the *k*-nearest neighbor method. In the present work 3 nearest neighbors were considered.

Computational implementation

GDE3 was implemented with the divide-and-conquer sorting method and the *k*-nearest neighbor method for removal of the most crowded non-dominated solutions in a Python (www.python.org) program for which the numpy (numpy.scipy.org) and scipy (www.scipy.org) modules were used for numerical calculus. A scaling factor *F* of 0.5 and a mutation probability *CR* of 0.7 were used. The *odeint* function from the scipy.integrate module, which uses the Adams/Backward Differentiation Formulae method for ODE numerical integration (Hindmarsh 1983; Hindmarsh and Petzold 1983), is used to compute the time courses predicted by the models for each solution.

4.5 The kinetics of the glyoxalase pathway

The kinetics of the glyoxalase pathway was studied in chapters 2 and 3 and the parameter estimates obtained in those chapters were used to compute optimized experimental conditions in the present chapter. The full set of parameters employed here may be found in table 4.1.

Table 4.1 Parameters of the glyoxalase pathway used for optimal experimental design for selection of a model for glyoxalase I.

Enzyme	**Glyoxalase I**		**Glyoxalase II**	**Glyoxalase I**		**Glyoxalase II**	
Organism	*S. cerevisiae* (purified)		*B. taurus* (purified from liver)	*S. cerevisiae* (permeabilized cells)			
	Model 1	Model 2		Model 1	Model 2	Model 1	Model 2
Parameters	k_{cat1} 8586 min^{-1} K_m 0.223 mM	k_{cat1} 17046 min^{-1} K_{m1} 0.875 mM K_{m2} 1.178 mM	k_{cat1} 314.79 min^{-1} K_m 2.862 mM	V_1 0.023917 mM/min K_m 0.214 mM	V_1 0.1347 mM/min K_{m1} 2.569 mM K_{m2} 1.202 mM	V_2 0.003313 mM/min K_{m2} 0.02857 mM	V_2 0.003447 mM/min K_{m2} 0.03472 mM

Commercially available purified glyoxalase 2 from bovine liver (Sigma) was also characterized so that this enzyme could be included in an experimental design strategy aiming at the discrimination of the kinetic model for glyoxalase I. The enzyme was assayed at 30ºC and pH 6.5 (70 mM KH_2PO_4 buffer like in section 2.3.2) using four different enzyme concentrations and four different initial S-D-lactoylglutathione concentrations (figure 4.5). The obtained substrate time courses were used to estimate its k_{cat} and K_m (table 4.2) with the program S-timator, described in section 2.3.2.

4.6 Results and discussion

Two kinds of experimental setup were considered: the first using purified baker's yeast glyoxalase I and bovine liver glyoxalase II and the second using permeabilized yeast cells. For the purified enzyme setup, the glyoxalase II reaction was included in the design, since it was hypothesized that it might contribute to obtain a larger divergence between the outputs of the models. The kinetic parameters of bovine liver glyoxalase II were determined (table 4.2) also from S-D-lactoylglutathione time courses (figure 4.5) to be used with the presented method.

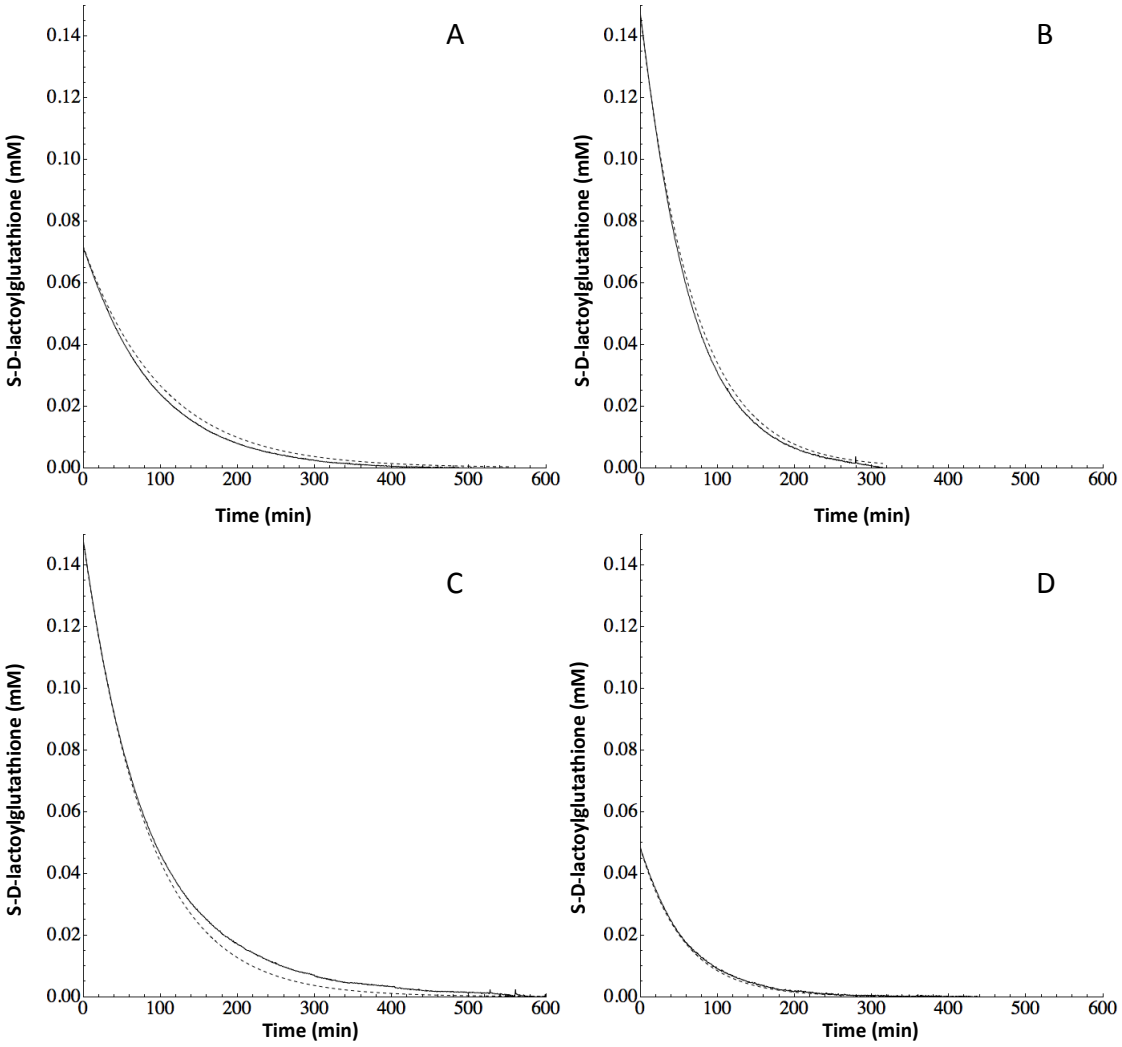


Figure 4.5 S-D-lactoylglutathione time courses used for the estimation of the kinetic parameters of the purified glyoxalase II from bovine liver (solid line) and simulation of the reaction with the estimated parameters (dashed line). Concentrations used in the assays: A – 9.15×10^{-5} mM Glyoxalase II and 0.0719 mM S-D-lactoylglutathione; B – 1.37×10^{-4} mM Glyoxalase II and 0.1479 mM S-D-lactoylglutathione; C – 1.14×10^{-4} mM Glyoxalase II and 0.1484 mM S-D-lactoylglutathione; D – 1.60×10^{-4} mM Glyoxalase II and 0.0484 mM S-D-lactoylglutathione.

Table 4.2 Kinetic parameters for bovine liver glyoxalase II

k_{cat}	314.79 min^{-1}
K_m	2.86 mM

4.6.1 Kullback-Leibler distance as objective function

The surface function representing the Kullback-Leibler distance was drawn for a first analysis of the usability of the algorithm with this particular kind of objective function. In figure 4.6 (C and D) the Kullback-Leibler distance is represented as a function of the concentration of glyoxalase II and the initial concentration of glutathione for an assay with purified enzymes with a fixed concentration of glyoxalase I (0.01 mM) and a fixed initial concentration of methylglyoxal (1.5 mM). The optimization was set to end after a number of consecutive generations without improvement of the solution set (20, as with the single-objective DE in chapter 3). However this proved to be extremely time-consuming and the 5000-maximum-generation criterion was used instead to terminate the optimization, unless stated otherwise. This criterion gives good approximations to the Pareto front, as shown by the convergence of the solutions to well-defined areas in the spaces of solutions (figure 4.6 A) and objectives (figure 4.6 B), so it was used for the remaining optimizations described in this chapter.

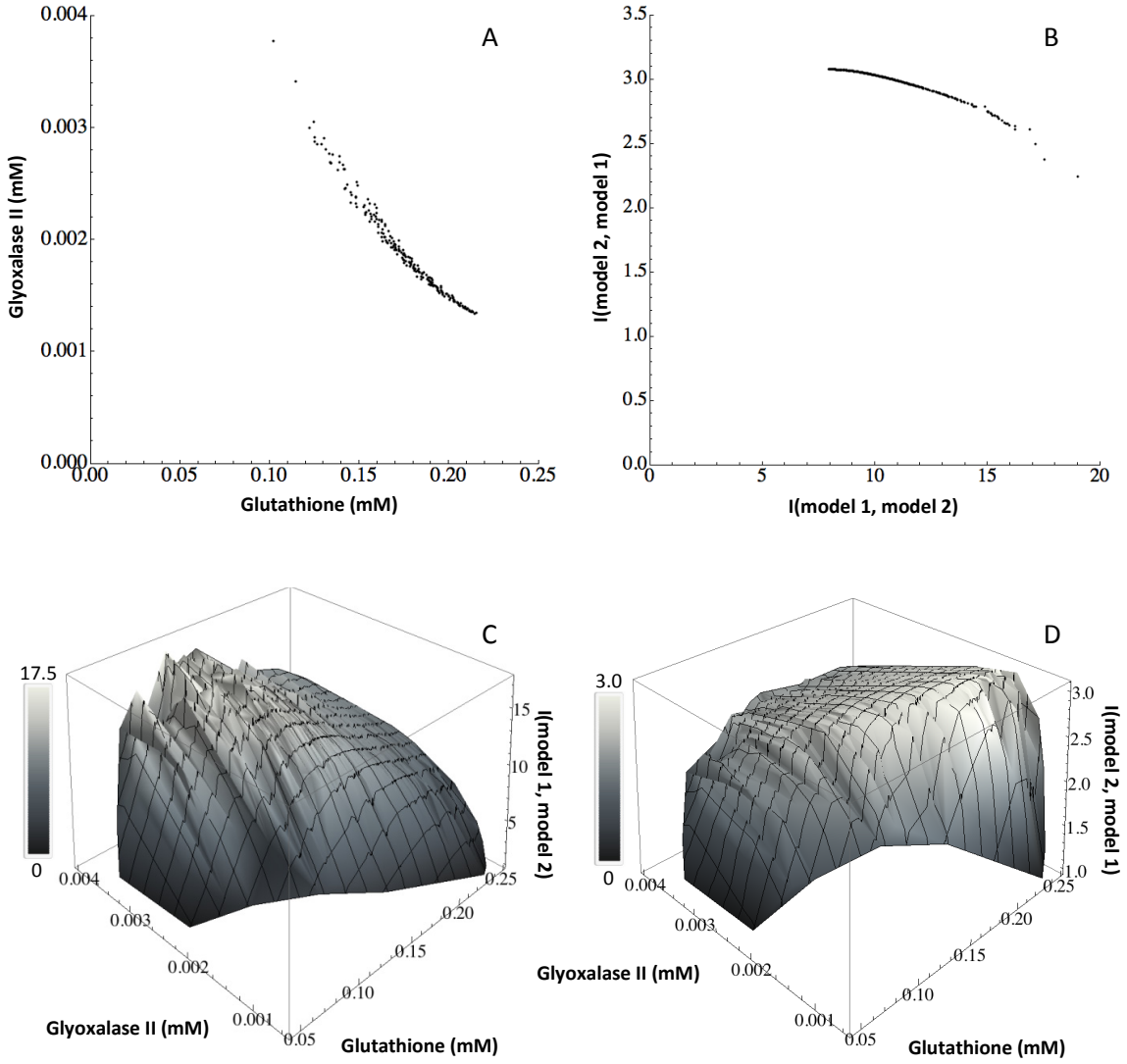


Figure 4.6 A – optimized methylglyoxal and glutathione concentrations (solutions approximating the Pareto front) in the space of solutions; B – Kullback-Leibler distances for the optimized concentrations in the space of objectives; C – Kullback-Leibler distance of model 2 from model 1 within the allowed optimization range; D – Kullback-Leibler distance of model 1 from model 2 within the allowed optimization range.

4.6.2 Discrimination of kinetic models of the glyoxalase system

The global optimal solutions of the optimization problem often cannot be used due to specific experimental limitations: for instance, the corresponding substrate concentrations may lead to intermediate concentrations below the detectable range or above the measurable limit; the necessary amount of reagents may be so high that the

experiment would be extremely expensive. Therefore, the algorithm should be given appropriate ranges for the experimental variables within which it should find the local optimum. In this section the use of the proposed procedure is exemplified and explained using the glyoxalase pathway as a simple application of the procedure to an experimental system.

Assays with purified enzymes

The glyoxalase enzymes are commonly assayed by following the intermediate S-D-glutathione at its characteristic absorption wavelength (240 nm) with an absorption coefficient of 2.86 mM^{-1} cm^{-1}. Since most spectrophotometers show a linear response to absorbance ranging from 0 to 1, approximately, the concentration of this compound should be up to 0.35 mM. So the limiting reagent (either methylglyoxal or glutathione) should not be initially higher than this concentration to avoid that intermediate exceeds the measurable concentration limit. For the case of the enzymes glyoxalase I and II, it is reasonable to use up to approximately 10 fold the quantities used in the assays of chapter 3, i.e. 5.00×10^{-3} mM and 2.00×10^{-3} mM respectively. For the substrates a low-bound concentration of 0.1 mM was set to ensure a measurable variation of S-D-lactoylglutathione. Running the program with such restrictions the optimal solutions resulted in a set of optimized conditions. A sample from this set may be found in table 4.3.

Table 4.3 Some solutions from the first non-dominated front are shown, after 5000 generations, for a population number of 200, a sampling frequency of 0.5 min and a total assay time of 120 min. Bold-highlighted results are used for the simulations in figure 4.7.

	Glutathione (mM)	Methylglyoxal (mM)	Glyoxalase I (mM)	Glyoxalase II (mM)
[Lower bound, Upper bound]	[0.1, 1]	[0.1, 1]	[0, 0.005]	[0, 0.002]
Solutions	**0.132**	**1.000**	**0.004989**	**0.001990**
	0.135	1.000	0.005000	0.001992
	0.192	1.000	0.004999	0.001118
	0.141	1.000	0.004994	0.001838
	0.180	1.000	0.004994	0.001239

The solutions in table 4.3 fulfill the requirements for the experiment to be performed: the concentration of the limiting substrate (glutathione) is below 0.35 mM – higher upper-bound values were nevertheless set for the substrates so that the algorithm would select the liming substrate as well – and the remaining reagents are within the corresponding acceptable concentration range. The optimal concentration value for methylglyoxal is always the upper-bound constraint imposed on the optimization and for glyoxalase I the optimal concentration is also very close to the pre-defined upper limit. This merely indicates that mathematically better solutions lie beyond the ranges set for the optimization. Extending the search range in the solution space should be considered if the computed optimal solutions do not give satisfactory differences between the outputs of the models.

In the particular case of this optimization for the glyoxalase pathway, the two models give significantly different S-D-lactoylglutathione time courses for the initial values in the optimized solutions. As shown on figure 4.7, for the conditions highlighted in bold in table 4.3, for a relatively small time of 120 min (5 fold less than the assays in chapter 3), not only the predicted values for the concentration of S-D-lactoylglutathione for one model are far from the other at all times (except at instant 40 min), but the dynamics of this intermediate is also different. For model 2 there is an immediate consumption of the available glutathione and an approximate steady state takes place during the initial 40

min, followed by a sudden consumption of all S-D-glutathione (and methylglyoxal); for model 1, S-D-glutathione also forms fast in the initial instants but more gradually; then it achieves a peak that is approximately half of the amount predicted by model 2 and decreases slowly and almost linearly until full consumption of S-D-glutathione (and methylglyoxal).

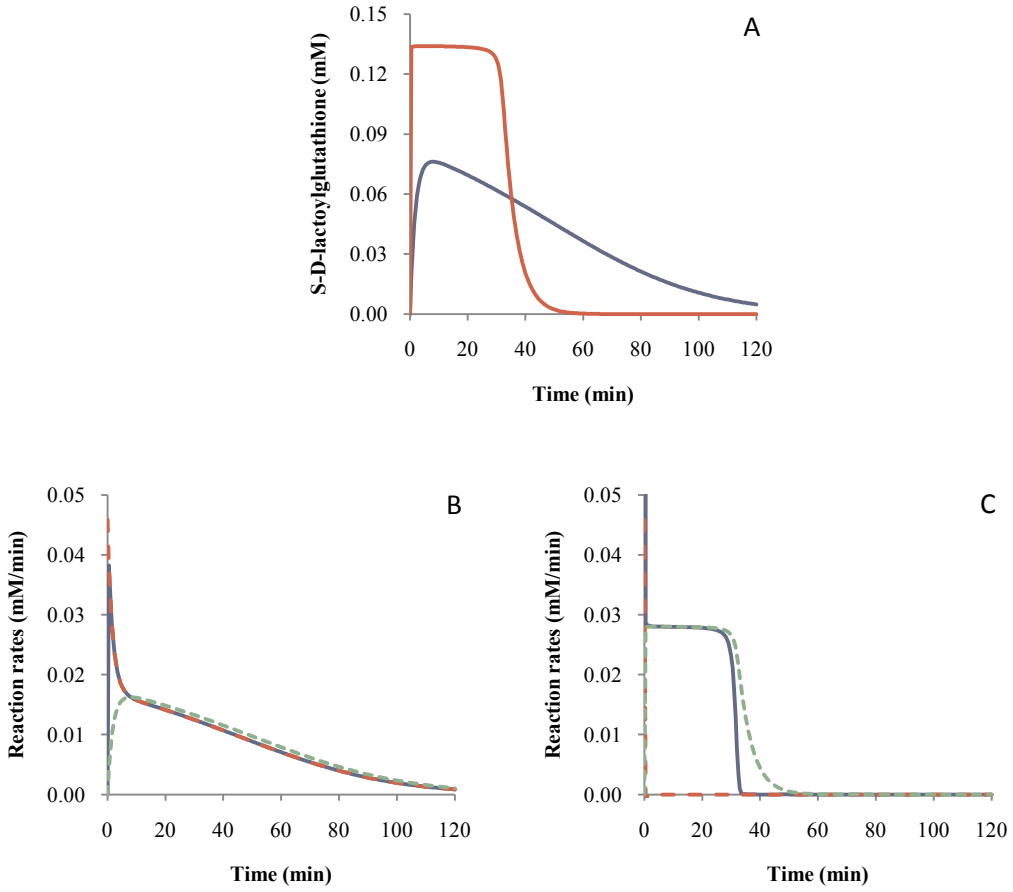


Figure 4.7 (A) Time courses for the S-D-lactoylglutathione concentrations predicted by model 1 (blue) and model 2 (red) for optimized experimental conditions (0.135 mM glutathione, 0.002 mM glyoxalase II, 1 mM methylglyoxal and 0.005 mM glyoxalase I; for extra details see text). (B) Predicted net rate of hemithioacetal formation (red), rate of S-D-lactoylglutathione formation (blue) and rate of S-D-lactoylglutathione hydrolysis (green) for model 1 and (C) model 2.

As shown on figures 4.7-B and 4.7-C, the reason for the optimal solutions to give such a different is that in these conditions, according to model 1, the rates of the enzymatic reactions are limited by the net formation rate of the methylglyoxal-glutathione hemithioacetal while, according to model 2, the enzyme-catalyzed reactions are

extremely fast, limited only by the regeneration of glutathione and the hemithioacetal formation reaction is approximately at equilibrium throughout the assay.

Assays with permeabilized cells

The use of permeabilized cells is an experimental strategy closer to the study of the actual living cell, when compared to studies where purified enzymes are used. In addition to keeping proteins inside the cell in their native concentrations, permeabilized cells pose challenges similar to the ones encountered when intact cells are used: enzyme concentrations inside cells are impossible or at least more difficult to perturb accurately in the desired extent (meaning that there are less variables that can be optimized) and, perhaps more important, data acquired with permeabilized cells is noisier than with a cell-free system.

The program proposed here was used to find the initial concentrations of glutathione and methylglyoxal that give a larger difference between the S-D-lactoylglutathione profiles for the two models of the glyoxalase pathway, for an experiment with permeabilized cells in the conditions described in section 2.3.2. The low-bound concentration for the substrates was chosen so that a measurable variation of S-D-lactoylglutathione might be observed. This is especially critical with permeabilized cells because very for very low substrate concentrations the absorbance of the intermediate is difficult to distinguish from noise. The upper bounds for methylglyoxal and glutathione concentrations were limited to 1 mM since optimal concentrations up to this value were expected to be high enough to allow a significant variation of S-D-lactoylglutathione and not exceed the spectrophotometer's dynamic range. The result of the optimization was the limit 1 mM for methylglyoxal and concentrations for glutathione very close to 0.239 mM.

For the assay with permeabilized cells it was necessary to extend the total time to 500 min so that the optimal concentration for glutathione wouldn't be the pre-defined upper limit as well. For much shorter total times the glyoxalase II-catalyzed regeneration of glutathione was not fast enough to result in a divergence of the time courses predicted by the models based on a feature other than the estimates of glyoxalase I activity for

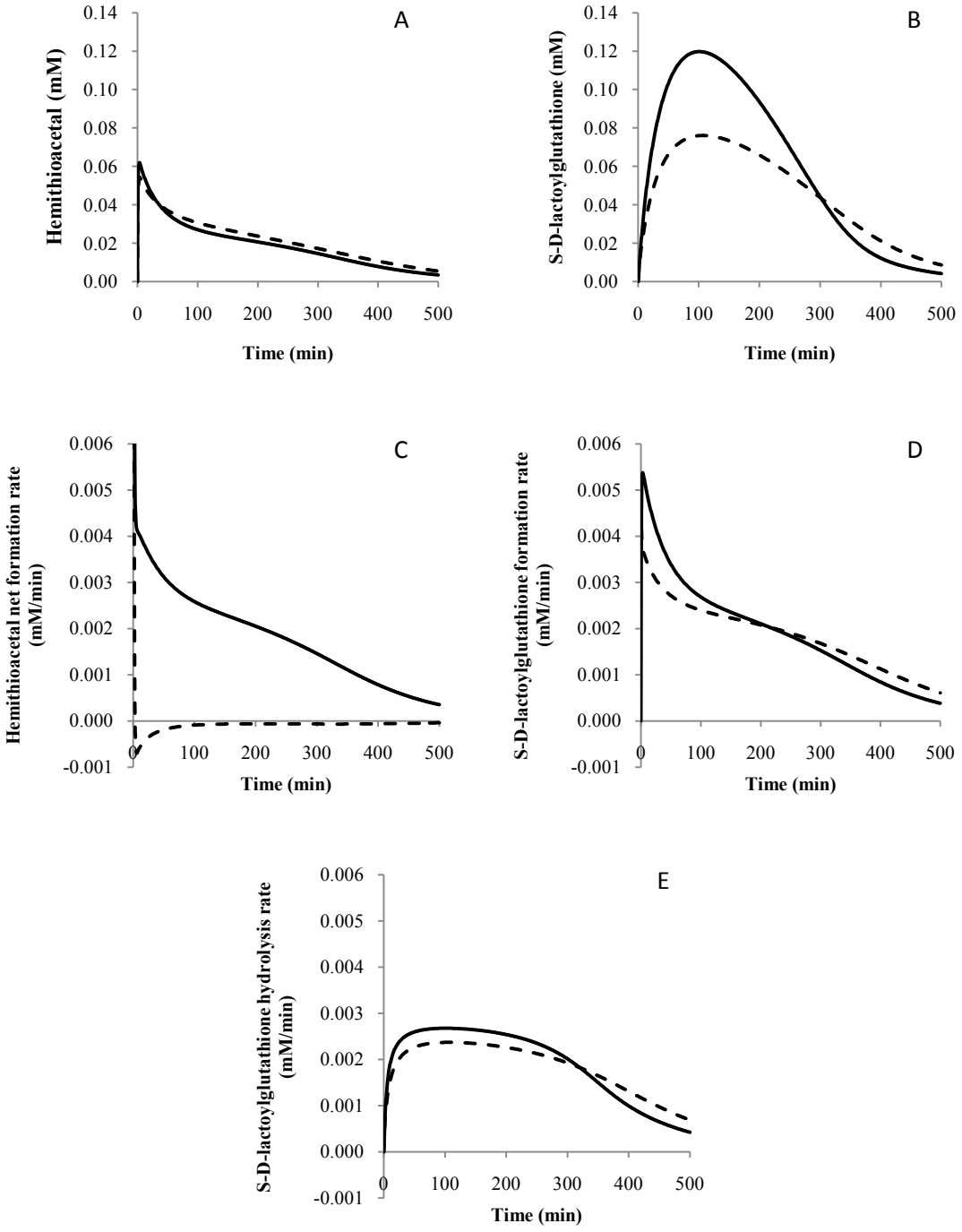


Figure 4.8 Time courses for (A) methylglyoxal-glutathione hemithioacetal and (B) S-D-lactoylglutathione concentrations predicted by model 1 (solid) and model 2 (dashed) for optimized experimental conditions (permeabilized cells grown to early stationary phase with and accessible protein concentration of 16.7 µg/mL and initial concentrations of 1 mM for methylglyoxal and 0.239 mM for glutathione; for extra details see text). In C, D and E the rates of the reactions occurring in the system are shown as functions of time.

each model. The algorithm would give high substrate concentrations as optimal since in that short time the S-D-lactoylglutathione formation rate would be limited by glyoxalase activity only. Using only one feature of the models would be a rather poor discrimination criterion. In addition, S-D-lactoylglutathione cannot exceed a concentration of 0.35 mM during the assay to remain within the spectrophotometer's dynamic limit. Therefore the total time of the assay was extended to 500 min. The S-D-glutathione time courses predicted by models 1 and 2 in these optimized conditions are shown in figure 4.8-B.

A significant difference between the S-D-lactoylglutathione time courses predicted by models 1 and 2 is evident for the assay with permeabilized cells – the maximum concentration predicted for model 1 is 1.6 fold higher than for model 2. Also in this case the observed difference is due to the difference in the net formation rate of hemithioacetal. According to model 1 the hemithioacetal forms at a rate close to the rate of the reaction of glyoxalase I while according to model 2 the net formation rate of the hemithioacetal is approximately 0 throughout the assay.

If the assay with permeabilized cells (figure 4.8-B) and the assay with purified enzymes (figure 4.7-A) are compared, it is evident that the discrimination should be easier with the purified enzymes. With the purified enzymes more experimental variables could be optimized, so a larger difference between the models is expected. Furthermore, permeabilized cells time courses have larger measurement errors and are more difficult to reproduce due to the turbidity of the mixture and degradation of the cells with time, especially if large methylglyoxal concentrations are used. Nevertheless, a significant difference is expected regarding S-D-lactoylglutathione maximum concentration during the assay and this could be enough to discriminate one of the two models.

To conclude, the combination of the GDE3 multiobjective optimization algorithm with the Kullback-Leibler distance as objective function is shown to be successful in providing solutions, within a reasonable computational time, for which the outputs of the candidate models are significantly different, therefore making the selection of a model, easier at least when two candidates are considered.

5 Metabolic impact of redox cofactor perturbations in *Saccharomyces cerevisiae*

Hou J, Lages NF, Oldiges M, Vemuri GN. 2009. Metabolic impact of redox cofactor perturbations in *Saccharomyces cerevisiae*. *Metabolic Engineering*. **11**: 253-261

5.1 Abstract

Redox cofactors play a pivotal role in coupling catabolism with anabolism and energy generation during metabolism. There exists a delicate balance in the intracellular level of these cofactors to ascertain an optimal metabolic output. Therefore, cofactors are emerging to be attractive targets to induce widespread changes in metabolism. We present a detailed analysis of the impact of perturbations in redox cofactors in the cytosol or mitochondria on glucose and energy metabolism in *Saccharomyces cerevisiae* to aid metabolic engineering decisions that involve cofactor engineering. We enhanced NADH oxidation by introducing NADH oxidase or alternative oxidase, its ATP-mediated conversion to NADPH using NADH kinase as well as the interconversion of NADH and NADPH independent of ATP by the soluble, non-proton-translocating bacterial transhydrogenase. Decreasing cytosolic NADH level lowered glycerol production, while decreasing mitochondrial NADH lowered ethanol production. However, when these reactions were coupled with NADPH production, the metabolic changes were more moderated. The direct consequence of these perturbations could be seen in the shift of the intracellular concentrations of the cofactors. The changes in product profile and intracellular metabolite levels were closely linked to the ATP requirement for biomass synthesis and the efficiency of oxidative phosphorylation, as estimated from a simple stoichiometric model. The results presented here will provide valuable insights for a quantitative understanding and prediction of cellular response to redox-based perturbations for metabolic engineering applications.

5.2 Introduction

The bakers yeast, *Saccharomyces cerevisiae*, is widely used as an industrial workhorse and therefore, designing its metabolism that is suited to a particular need is crucial for its use as an efficient cell factory. The availability of its genome information, high throughput tools such as microarrays, well-curated databases that contain cellular information and robust genetic manipulation techniques greatly improved the precision of engineering its metabolism. Despite the progress, efforts to design metabolic networks often meet with unexpected consequences due to the high interconnectivity of cellular components, which tend to counteract the effect of mutations (Vemuri and Aristidou 2005). Cofactors of metabolism such as NADH and ATP are among the most highly connected metabolites in metabolic networks (Nielsen 2003a). Consequently, changing their concentration is bound to have a widespread impact on metabolism.

More recently, engineering metabolism by targeting redox cofactors such as NADH proved to be an effective strategy in *S. cerevisiae* (Bro *et al.* 2004; Geertman *et al.* 2006; Heux *et al.* 2006; Vemuri *et al.* 2007) since close to 100 metabolic reactions involve NADH (Forster *et al.* 2003). NADH is predominantly produced in the catabolism of glucose and redox homeostasis is maintained by the action of NADH dehydrogenases. Since NADH cannot traverse the mitochondrial membrane (von Jagow and Klingenberg 1970), *S. cerevisiae* has two external and one internal NADH dehydrogenases (Bakker *et al.* 2001; Rigoulet *et al.* 2004). Glucose cannot be oxidized as fast as it can be consumed and a portion of it is fermented to ethanol (De Deken 1966), a phenomenon commonly known as the Crabtree effect (Crabtree 1929). This phenomenon is often attributed to the inhibition of respiration (Sonnleitner and Kappeli 1986; Rigoulet *et al.* 2004). Ethanol production is a redox-neutral process and hence, does not contribute to NADH accumulation. The additional NADH generated from concomitant biomass synthesis is oxidized via glycerol production (Vandijken and Scheffers 1986; Overkamp *et al.* 2000). Previously, water-forming NADH oxidase was introduced to alleviate cytosolic NADH accumulation in S. cerevisiae (Heux *et al.* 2006; Vemuri *et al.* 2007), which substantially decreased glycerol synthesis. On the other hand, decreasing the mitochondrial reducing power by introducing an alternative

oxidase from *Histoplasma capsulatum* (Johnson *et al.* 2003) reduced ethanol synthesis by enhancing the mitochondrial capacity to oxidize pyruvate resulting from glucose metabolism (Vemuri *et al.* 2007). Clearly, compartment-specific perturbation of redox cofactors has a great impact on the product profile.

While NADH is the predominant redox product of catabolism, NADP(H) has a greater role in anabolism as many of the reactions involved in the biosynthesis of amino acids, lipids and nucleotides use NADPH as the reducing agent. In *S. cerevisiae*, a majority of the NADPH is generated in the oxidative pathway of the pentose phosphates, which converts glucose-6-phosphate to ribose-5-phosphate in the cytosol. Mitochondrial NADPH is synthesized mainly from the NADP-dependent isocitrate dehydrogenase in the TCA cycle. Besides these, there are other reactions that might produce NADPH, such as the $NADP^+$-dependent acetaldehyde dehydrogenase, which have negligible carbon flux during aerobic growth on glucose (Bruinenberg *et al.* 1983). Baker's yeast also has NADH kinases in the cytosol (Utr1p and Yef1p) and mitochondria (Pos5p) that mediate the ATP-driven conversion of NADH into NADPH without transporting electrons from NADH (Shi *et al.* 2005) of which only Pos5p is well-studied and appears to be dominant (Outten and Culotta 2003; Shianna *et al.* 2006). Unlike many organisms, *S. cerevisiae* does not have transhydrogenases that interconvert NADH and NADPH. Introducing a membrane-bound transhydrogenase from *Escherichia coli* (Anderlund *et al.* 1999) or a soluble one from *Azobacter vinelandii* (Nissen *et al.* 2001) in *S. cerevisiae* catalyzed the conversion of NADPH to NADH. Normally, membrane-bound transhydrogenases convert NADH to NADPH, while the soluble transhydrogenases convert NADPH to NADH, but when introduced in *S. cerevisiae*, both converted NADPH to NADH. These perturbations are bound to have an impact on the yield and rate of ATP synthesis as well as on the efficiency of oxidative phosphorylation (*PO* ratio).

The objective of this study is to evaluate the metabolic impact of decreasing NADH levels in the cytosol or mitochondria of *S. cerevisiae* either with or without linking this to NADPH generation at the expense of ATP. To achieve this goal, we expressed (i) the heterologous *nox* gene (encoding for water-forming NADH oxidase) from *S. pneumonia* (Auzat *et al.* 1999) that converts cytosolic NADH to NAD^+ using molecular oxygen, (ii)

the alternative oxidase, *AOX1*, from *H. capsulatum* (Johnson *et al.* 2003) that enhances the oxidation of mitochondrial NADH, (iii) the native NADH kinase directed to the cytosol (truncated *POS5*) (Strand *et al.* 2003), (iv) the native, mitochondrial NADH kinase (full-length *POS5*) and (v) a non-proton translocating, soluble transhydrogenase (*udhA*) gene from *E. coli* (Boonstra *et al.* 1999) and compared the alteration in metabolic patterns with an isogenic reference strain. Overexpressing NADH oxidase and alternative oxidase directly lowered NADH in the cytosol and mitochondria, respectively, without any direct impact on other cofactors or other aspects of metabolism. NADH kinase decreases NADH level by converting it into NADPH, thereby maintaining a redox balance, but a shift in the level of the redox cofactors. Since this conversion uses ATP, the cellular energy level is also directly affected. Based on the reaction directionality of the other transhydrogenases (Anderlund *et al.* 1999; Nissen *et al.* 2001) in *S. cerevisiae*, the soluble transhydrogenase from *E. coli* is also expected similarly to catalyze the conversion of NADPH to NADH. The impact of these perturbations on the product formation profile, bioenergetic parameters and intracellular metabolites was evaluated. Since these aspects play a pivotal role in determining the efficacy of a targeted, rational perturbation, the results presented here prove valuable to metabolic engineering applications.

5.3 Materials and methods

5.3.1 Yeast strain construction

Genes for NADH oxidase (*nox* from *S. pneumoniae*), alternative oxidase (*AOX1* from *H. capsulatum*), NADH kinase (*POS5* from *S. cerevisiae*) and soluble transhydrogenase (*udhA* from *E. coli*) were amplified using Expand high-fidelity PCR system (Roche Applied Science, Indianapolis, IN). The primers used for the PCR, the source of the corresponding templates for amplification and relevant restriction sites are shown in table 5.1. NADH kinase (encoded by *POS5*) contains a 17 amino acid-long mitochondrial signal peptide that guides the mature protein into the mitochondria. The full-length protein was amplified for perturbing the mitochondrial NADH level. For directing the enzyme into the cytosol, the gene without the first 51 bases was amplified and the ATG translational start codon was artificially inserted through the primer (shown in bold in table 5.1). The gene fragments were amplified and digested with the appropriate restriction enzymes and gel-isolated. A high copy number plasmid derived from pRS426 with native TEF1 promoter and CYC1 terminator (Wattanachaisaereekul *et al.* 2008) was also digested with the corresponding restriction enzymes and was used to ligate the amplified gene fragments. The resulting plasmids were propagated in *E. coli* and sequenced to ensure sequence accuracy. The plasmids without any sequence variations were transformed into *S. cerevisiae* strain CEN.PK113-5D (*MATa SUC2 MAL8C ura3-52*). These strains were designated according to the gene that was overexpressed in them (NOX, AOX, cNDK, mNDK and UDH). A strain containing the plasmid without any insert was used as the isogenic reference (REF). All the strains were maintained on synthetic-complex-medium agar plates lacking uracil. They were grown on synthetic-complex medium lacking uracil and an aliquot of the culture was mixed with glycerol to a final concentration of 20% for long-term storage at −80 °C.

Table 5.1 The forward and reverse primer sequences for the genes from different hosts that were introduced into *S. cerevisiae*.

Gene	Source	Primers (5' -> 3')	Enzymes	Template (Ref)
nox	*S. pneumoniae*	F: CTATACGGATCCATGAGTAAAATCGTTGTAGTCGGTGC R: GTGTCTATCGATTTATTTTTCAGCCGTAAGGGCAG	BamHI ClaI	pNOX2 (Auzat *et al.* 1999)
AOX1	*H. capsulatum*	F: CTATACGGATCCATTAAGGAATTTTTTGCGCCC R: TAGTAGAATTCTCATATCACCTCATCCCGTTCC	BamHI EcoRI	pRS426-AOX1 (Johnson *et al.* 2003)
POS5 (cytosolic)	*S. cerevisiae*	F: TTGCGGATCCATGAGTACGTTGGATTCACATTC R: TGTACGATCGATTTAATCATTATCAGTCTGTCTCTTGGTC	BamHI ClaI	Genomic DNA of *S. cerevisiae* CEN.PK113-7D
POS5 (mitochondrial)	*S. cerevisiae*	F: CGTACAGGATCCATGTTTGTCAGGGTTAAATTGAATAAAC R: TGTACGATCGATTTAATCATTATCAGTCTGTCTCTTGGTC	BamHI ClaI	Genomic DNA of *S. cerevisiae* CEN.PK113-7D
udhA	*E. coli*	F: CTGTCCACTAGTATGCCACATTCCTACGATTACGATGCC R: TATGCCAAGCTTTTCGATAAAGTTTTAAAACAGGCGG	SpeI HindIII	Genomic DNA of *E. coli* MG1655

The sites for restriction enzyme digest used for constructing the expression plasmids are underlined in the primer sequences. Full-length NADH kinase contains an N-terminal mitochondrial signal peptide and naturally localizes in the mitochondria. It was directed into the cytosol by amplifying the gene without its mitochondrial signal peptide (see text for details).

5.3.2 Media and growth conditions

For all cultivations, a previously described mineral salts medium was used (Verduyn *et al.* 1992). Cells were grown at 30 °C and pH 5.0 (controlled with 2 N KOH) in batch cultivations using 20 g/L glucose as the carbon source in 2.2 L Biostat B (B Braun, Germany) glass vessels with a working volume of 2.0 L. Aerobic conditions were maintained by sparging filtered air at 1 vvm and agitating the culture at 600 rpm. Carbon dioxide and oxygen concentrations in the off-gas were monitored using an acoustic gas analyzer (Brüel and Kjær, Nærum, Denmark). The seed culture for the batch cultivations was grown in 500 mL shakeflasks with 100 mL culture at 30 °C and a starting pH of 6.0 while agitating in an orbital shaker at 250 rpm. Seed culture from the shakeflasks was inoculated to a final dry weight of 1 mg/L.

The accelerostats were initiated as batch cultures with 1.0 L working volume. Fresh medium containing 10 g/L glucose was fed after the completion of the batch growth at a flowrate of 0.1 L/h and the culture was allowed to reach a steady state (as indicated by steady value of CO_2 in the off-gas for at least five volume changes). After attaining a steady state, the medium flowrate was accelerated at 0.01 L/h^2. A small aliquot (2 mL) of the culture was withdrawn, filtered through 0.45 μm cellulose acetate filters (Osmonics, Westborough, MA) and stored at -20 °C until subsequent analysis. The fermenters and the gas analyzer were integrated with the multi-fermenter control system (MFCS) for data acquisition and control.

5.3.3 Biomass determination

Cell dry weight was measured by filtering a known volume of the culture through a pre-dried and pre-weighed 0.45 μm pore size nitrocellulose filter (Supor®-450 membrane filters, PALL Life Sciences, Ann Arbor, MI, USA). The filters with the biomass were washed with water, dried for 15 min in a microwave oven at 150 W and weighed again. The optical density was determined at 600 nm using a Hitachi U-1100 spectrophotometer.

5.3.4 Extracellular metabolite analysis

Concentrations of glucose, glycerol, ethanol, acetate and pyruvate were analyzed by an isocratic HPLC with an Aminex HPX-87H ion exchange column (Bio-Rad, Hercules, USA) at 65 °C using 5 mM H2SO4 as mobile phase at a flow rate of 0.6 $mL \cdot min^{-1}$. Glucose, glycerol and ethanol were measured with a refraction index detector (Waters 410 differential refractometer, Millipore, CA, USA) and acetate and pyruvate with

ultraviolet-visible light absorbance detector (Waters 486 tunable absorbance detector set at 210 nm, Millipore, CA, USA).

5.3.5 Quantification of intracellular metabolites

For the analysis of intracellular metabolites duplicate samples were collected in the mid-exponential phase during growth on glucose or on ethanol after the diauxic shift in aerobic batch cultivations. Culture broth (approximately 10 mL) was rapidly extracted into pre-weighed tubes containing 30 ml of pre-chilled methanol (−40 °C). The tubes were weighed again to determine the amount of biomass. The cells were separated by centrifuging at 5000g for 20 min at −20 °C and resuspended in 2 mL cold methanol and 1 mL of chloroform and stored at −80 °C. The metabolites were extracted by adding 2 mL chloroform (−20 °C) and 1 mL of 3 mM Pipes-EDTA (pH 7) followed by vigorous shaking at 500 rpm, 45 min at −20 °C (Smits et al., 1998). The metabolites were separated by centrifugation (3000g, 20 min, −20 °C) and the upper phase was used for quantifying metabolites. Intracellular metabolites of the central carbon metabolites (glycolysis, TCA cycle and the pentose phosphate pathway) and the metabolic cofactors were measured by liquid chromatography coupled with tandem quadrupole mass spectrometry, as described (Luo *et al.* 2007).

5.3.6 Stoichiometric modeling

A simple stoichiometric model based on the complete metabolism of glucose to CO_2 in the EMP-TCA cycle and the pentose phosphate pathway as well as its conversion to biomass, ethanol and glycerol was developed according to the reaction scheme shown in figure 5.1. The condensed catabolic and anabolic reactions have been described

previously (Nielsen 2003b) and are shown in the Appendix B for convenience. We assumed a constant value for biomass composition. We also assumed that NADH is produced only as a result of glucose catabolism and NADPH is produced only in the pentose phosphate pathway. $FADH_2$ was assumed equivalent to NADH in all the calculations. The amount of ATP required for biomass synthesis (Y_{XATP}) was included in the equation for biomass synthesis. An energy dissipation reaction was also included to account for ATP consumption not contributing to biomass (m_{ATP}). The Pirt equation was rewritten to include the bioenergetic parameters (see Appendix B), which were estimated by least squares regression.

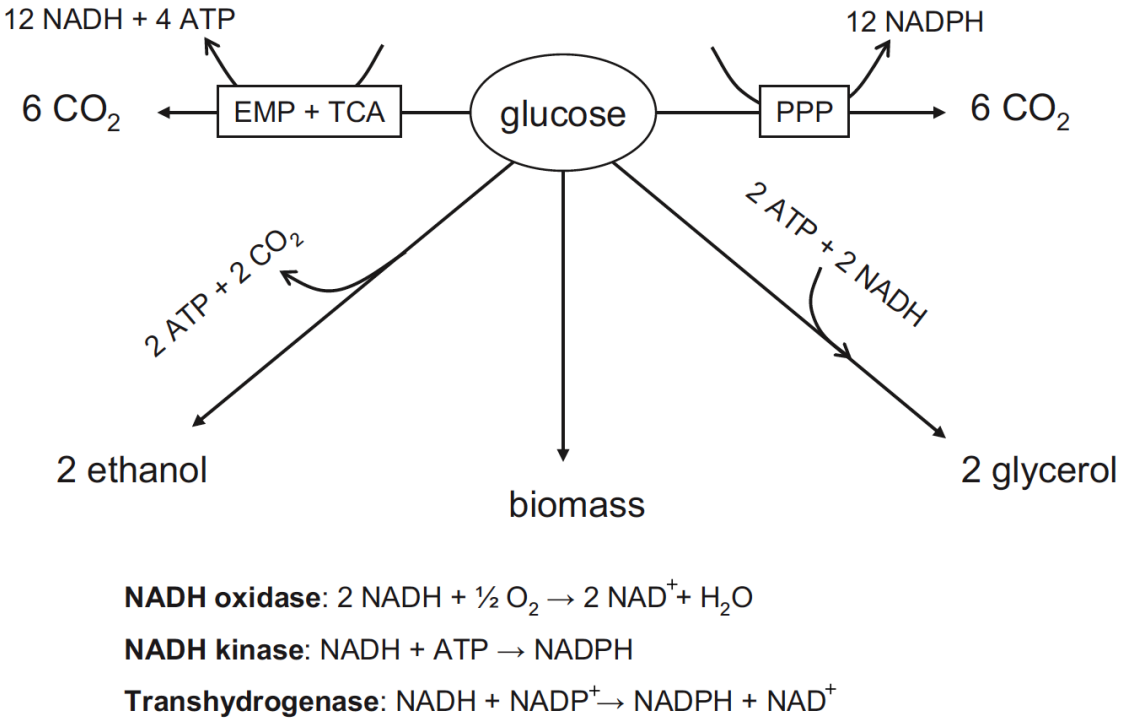


Fig. 5.1 Simplified stoichiometric model illustrating the conversion of glucose to biomass, ethanol and glycerol as well as its complete oxidation to CO_2 in the pentose phosphate pathway (PPP) and in the combined Embden Meyerhof Parnas and the tricarboxylic acid cycle (EMP+TCA) pathways. The associated metabolic cofactors are also shown in the reaction scheme. The reactions catalyzed by the enzymes overexpressed in *S. cerevisiae* are also shown. The reaction catalyzed by the transhydrogenase proceeds in the direction shown in E. coli. The details of the model and other reactions involved are given as a subsection in section 5.2 and further elaborated in the Appendix B at the end of the chapter.

5.4 Results

We introduced enzymes that perturbed either mitochondrial or cytosolic NADH levels in S. cerevisiae either in isolation or concomitantly with other cofactors such as ATP and NADPH. NADH oxidase mediates the direct oxidation of NADH to NAD^+ and NADH kinase drives the ATP-dependent conversion of NADH to NADPH, thereby affecting ATP and NADPH levels as well. The functional expression of the heterologous enzymes was confirmed by enzyme assays.

5.4.1 Growth characteristics in batch cultures

During exponential growth on glucose, the maximum specific growth rate (μ_{max}) of the strains in which only NADH was perturbed (NOX and AOX) was not affected by the perturbations. The μ_{max} decreased when cytosolic NADH kinase or the transhydrogenase were overexpressed by 11% and 17%, respectively (table 5.2). However, overexpressing the mitochondrial NADH kinase increased the μ_{max} by 11% compared with that of the reference strain. The result clearly points to the key role of NADPH in biomass synthesis. The specific glucose uptake rate (r_{glu}) very clearly corresponded with cytosolic NADH consumption. In NOX and cNDK, r_{glu} increased by 7% and 11%, respectively compared with that in REF (table 5.2). In AOX and mNDK, this value decreased by 16% and 6%, respectively. Upon the introduction of the transhydrogenase, it decreased significantly, suggesting the formation of NADH from NADPH. Interestingly, the biomass yield (Y_{SX}) from glucose increased only in the strains where mitochondrial NADH level was decreased (AOX and mNDK) (Table 2). The biomass yields from oxygen (Y_{OX}) were also calculated to indicate the efficiency of oxidative metabolism. The NOX strain had the highest oxygen requirement for biomass synthesis. Since Y_{SX} for NOX remained the same as that for REF (table 5.2), the additional oxygen was used by NADH oxidase. Overexpressing mitochondrial NADH

kinase increased biomass yield on oxygen by 20%, and exhibited a concomitant increase in the biomass yield from glucose.

Table 5.2 Growth parameters of the six strains carrying perturbations in NADH metabolism from batch cultivations on glucose.

Strain	μ_{max} [a]	r_{glu} [b]	r_{eth} [b]	r_{gly} [b]	r_{CO_2} [b]	Y_{sx} [c]	Y_{ox} [c]	C-bal[d]	R-bal[e]
REF	0.34	17.2	24.2	2.69	44.9	0.11	31.39	1.05	−0.25
NOX	0.33	18.3	25.1	0.72	52.1	0.10	21.75	1.00	−0.66
AOX	0.32	14.5	13.6	2.47	45.7	0.13	39.17	0.99	−1.08
cNDK	0.30	19.1	24.7	1.16	46.8	0.09	42.18	0.92	−0.81
mNDK	0.38	16.2	16.5	2.35	42.9	0.13	38.13	0.94	−0.98
UDH	0.28	14.4	21.3	3.19	28.6	0.10	32.13	0.96	−0.17

The average from at least duplicate cultivations are shown and in all cases, the values calculated differed less than 4%.
[a] Maximum specific growth rate (h^{-1}) on glucose.
[b] Specific rates of glucose uptake, ethanol, glycerol and CO_2 production during exponential growth (mmol/g DCW h).
[c] Yield coefficients of biomass from glucose (g DCW/g glu) and oxygen (mg DCW/mmol O_2).
[d] Carbon balance during exponential phase of glucose consumption.
[e] Degree of reduction balance based on the carbon distribution during exponential growth phase on glucose.

The product formation profile differed substantially with the nature of perturbation to NADH level in different compartments. The specific ethanol production rate (r_{eth}) decreased when mitochondrial NADH level was decreased, in AOX and mNDK by 50% and 30%, respectively. However, ethanol production rate was largely unaffected by perturbation to cytosolic NADH, as observed in NOX, cNDK and UDH (table 5.2). The production of glycerol (r_{gly}) depended on the cytosolic NADH/NAD^+ ratio. In NOX and cNDK, the specific production rate of glycerol decreased by 75% and 50%, respectively (table 5.2). The AOX and mNDK strains, in which mitochondrial NADH/NAD^+ ratio was lowered, did not show appreciable decrease in glycerol production. However, the specific glycerol production rate in the UDH strain increased by about 20%, compared with that in REF. This phenotype suggests that transhydrogenation proceeds from NADPH to NADH in *S. cerevisiae*. The specific CO_2 production rate (r_{CO_2}) increased noticeably only when NADH oxidase was introduced, and remained approximately the

same in all the other strains, except in UDH. The presence of the transhydrogenase decreased r_{CO_2} by 36% (table 5.2).

We accounted for more than 90% of all the glucose consumed during exponential growth in all the strains by calculating the amount of carbon distributed to the various products and biomass (table 5.2). The degree of reduction was also satisfactorily balanced for all the strains (considering experimental errors and assumptions for the macromolecular composition).

5.4.2 Distribution of intracellular metabolites

Since the perturbations we made are expected to have a greater impact on the metabolism, we quantified the cofactors and the central carbon metabolism intermediates in all the strains at the mid-exponential growth phase on glucose as well as during growth on ethanol following the diauxic shift.

During growth on glucose, the intracellular ratio of NADH/NAD^+ was the lowest for NOX and AOX, followed by cNDK and mNDK (figure 5.2A). The UDH strain had much higher NADH/NAD^+ ratio compared with the other recombinant strains, but the ratio was lower than that for the REF strain. The NADPH/$NADP^+$ ratio did not change in any of the strains, except cNDK, which had a 4–6 fold increase in this ratio compared with the other strains (figure 5.2B). Quite surprisingly, the NADPH/$NADP^+$ ratio did not differ significantly between mNDK and the REF strains. The ratio of the reduced redox cofactors, NADH/NADPH was about 40% lower in mNDK compared with REF, indicating the effective functioning of NADH kinase in the mitochondria. The NADH/NADPH ratio decreased drastically (by 94%) in cNDK. Although NADH kinase is a native mitochondrial enzyme, it appears to be more active in the cytosol. The intracellular energy ratio (ATP/ADP) is closely linked to the redox ratio (NADH/NAD^+) via the oxidative phosphorylation chain. The lower value of NADH/NAD^+ ratio weakly correlated with a higher ATP/ADP ratio, except for the

UDH strain (figure 5.2C). The UDH strain had the highest ATP/ADP ratio among all the strains. There was not any statistical difference in the ATP/ADP ratio between the other strains. We note that these measurements do not distinguish between cofactor levels in the different compartments, but rather present the concentrations in the entire cell.

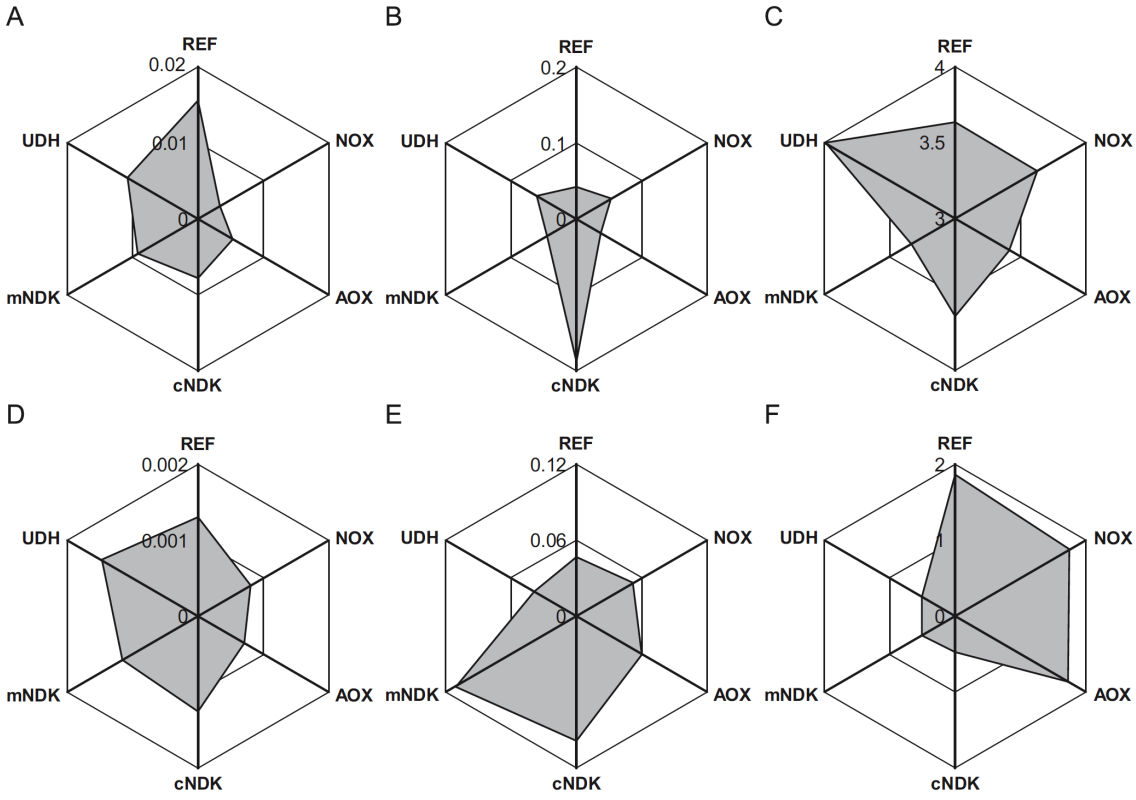


Figure 5.2 Radar plot illustrating the intracellular values of NADH/NAD$^+$ (A and D), NADPH/NADP$^+$ (B and E) and ATP/ADP (C and F) molar ratios in the six strains during growth on glucose (A–C) or on ethanol (D–F) following the diauxic shift. The concentrations of the metabolites were measured and the ratio calculated separately for each cofactor pair. The six vertices indicate the strains. The degree of asymmetry in the plot indicates the extent of differences in the ratios. The values shown are averages of at least duplicates from duplicate biological replicates.

During growth on ethanol (and other non-fermentable carbon sources produced during glucose fermentation, such as glycerol and acetate), the profile of the cofactors significantly differed from that during exponential growth on glucose. The NADH/NAD$^+$ ratio was one order lower in magnitude during the ethanol phase due to the lower concentration of NADH. Although the NADH/NAD$^+$ ratio was still the lowest

for NOX and AOX during ethanol consumption, the difference in the NADH/NAD^+ ratio was not as significant as it was during growth on glucose (figure 5.2D). The UDH strain had the highest NADH/NAD^+ ratio. Similarly, the profile of the NADPH/$NADP^+$ ratio among the strains differed during ethanol consumption. The mitochondrial NADH kinase appears to have a greater impact during growth on ethanol, as indicated by the NADPH/$NADP^+$ ratio. We did not see any change in the ATP/ADP ratio between REF, NOX and AOX. It is to be noted that during growth on glucose, the ATP/ADP ratio was the highest in the UDH strain, while it is the lowest in the UDH strain during growth on ethanol (figure 5.2F). Based on concentration of the cofactors, it is possible to calculate the Gibbs free energy of the transhydrogenase reaction for the UDH strain as

$$\Delta G = \Delta G^{\circ} + R\,T\,\ln\left(\frac{[NADPH][NAD^+]}{[NADP^+][NADH]}\right),$$

assuming that the reaction proceeds as: NADH + $NADP^+ \rightarrow$ NADPH + NAD^+. In this equation, R is the Universal gas constant (8.31 JK^{-1} mol^{-1}) and T is the absolute temperature (300 K). It is reasonable to assume that the energy change involved in the oxidation of NADH is approximately equal to that required for the reduction of $NADP^+$, which leads to $\Delta G° = 0$ for this reaction. The value of the free energy calculated from the concentration of the cofactors was $\Delta G = 4.3$ kJ mol^{-1}, indicating that the reverse reaction (NADPH to NADH) is favored.

Quantification of the central carbon metabolites and the cofactors involved clearly bring out the differences between the growth conditions as well as between the strains. A simple hierarchical clustering separated the metabolic profile during glucose phase from that during the ethanol phase, indicating that the growth conditions had a much stronger impact on the metabolite profiles than the perturbations (figure 5.3). The diauxic shift represents the intracellular modifications the cell undergoes from utilizing glucose as the carbon source to utilizing ethanol as the carbon source. The impact of these modifications is reflected in the metabolite concentrations. The changes in the ratios of the cofactors described in figure 5.2 are also reflected in the clustering scheme. During growth on glucose, the concentrations of many glycolytic metabolites (phosphoglycerate, PEP, acetyl CoA, among others) and the TCA cycle intermediates (citrate, malate, fumarate, etc.) were at a higher concentration. During this phase, the

concentration of the reduced cofactors such as NADH and NADPH was also high. On the other hand, during growth on ethanol, these metabolites were present in lower amounts. Glucose-6-phosphate, pyruvate and ATP were present at higher concentrations during growth on ethanol. Some of the pentose phosphate metabolites were also present at higher concentrations during growth on ethanol (figure 5.3). It is noteworthy to mention that the ΔG for the transhydrogenase reaction in UDH strain is−0.3 kJ mol^{-1} during growth on ethanol, indicating that the formation of NADPH is favored under these conditions.

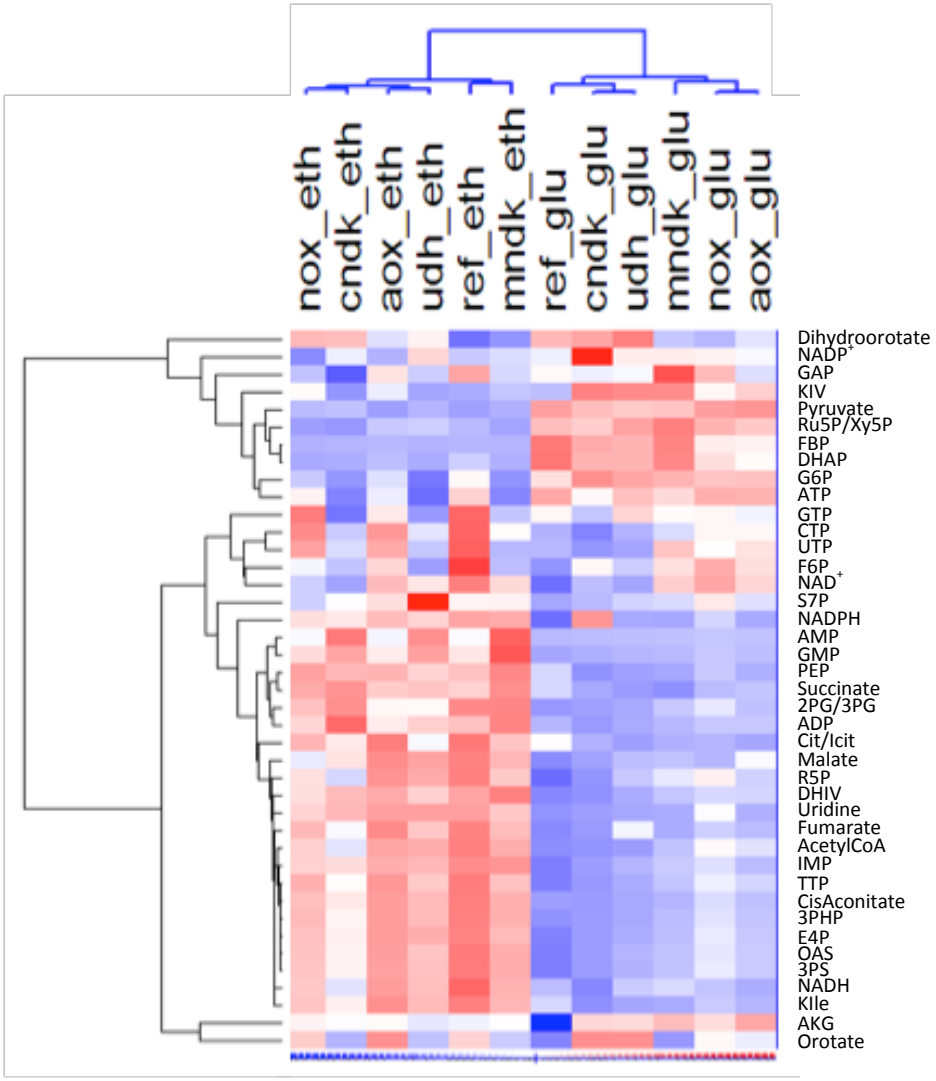


Figure 5.3 Hierarchical clustering of the cofactors and the central carbon metabolites from the six strains during exponential growth on glucose and ethanol shown as a heat map. The values of metabolite concentrations are represented as colors, with blue indicating those that have higher concentrations and red indicating those that have lower concentrations. The metabolites clearly separate the growth phases (columns) and also the pathways they are involved in (rows). The values considered for the clustering are averages of at least duplicates from duplicate biological replicates.

5.4.3 Phenotypic differences in accelerostats

During growth in batch cultures, the cells grow at their maximum specific growth rate and their metabolism is regulated by glucose repression. Accelerostats provide an excellent and rapid means to remove the effect of glucose and study the impact of the perturbations. The accelerostats were initiated as glucose-limited chemostats at a dilution rate of 0.1 h^{-1}. Under these conditions, glucose was completely oxidized to biomass and CO_2. The biomass yield from glucose remained close to 0.5 g/g for all the strains, except for NOX and UDH, which exhibited lower biomass yield (0.45 and 0.4 g/g, respectively). As the dilution rate increased, the rate of metabolism also increased, as indicated by faster r_{glu} and r_{CO_2}. We did not observe any fermentation products until the dilution rate reached about 0.2 h^{-1}, when REF, NOX and cNDK produced ethanol (figure 5.4A). Ethanol production in AOX and mNDK was delayed until the dilution rate reached approximately 0.25 h^{-1}. On the contrary, UDH already commenced ethanol production at 0.15 h^{-1}. During the respiratory-fermentative phase of glucose metabolism, the yield of ethanol on biomass increased for NOX, AOX and mNDK, indicating greater amount of ethanol that needs to be produced to sustain a unit of biomass. The difference in glycerol production was even more prominent between the strains. Glycerol production decreased in all the strains except UDH. In this strain, glycerol production commenced even at very low dilution rates. The yield of glycerol on biomass decreased for all the strains compared to the reference (figure 5.4B, inset).

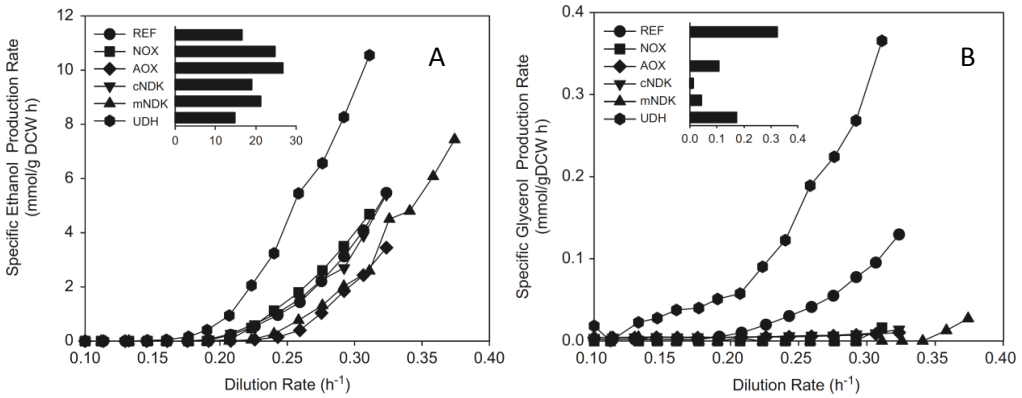


Figure 5.4 Profiles of specific production rates of ethanol (A) and glycerol (B) in the six strains grown in accelerostats as a function of the dilution rate. The transition from completely respiratory conditions to respiratory-fermentative conditions can be clearly seen. The dilution rate at which these strains commence ethanol production is difference from that at which they start producing glycerol. Also shown in the inset is the yield of ethanol (A) or glycerol (B) on biomass (mmol/g DCW). Strains in which cytosolic NADH was dissipated (NOX and cNDK) did not produce any glycerol.

During the acceleration, the rate of change of the growth rate is negligible compared to the turnover rates of the intracellular metabolites, so we assumed a pseudo steady state to calculate the true yield coefficients according to the Pirt equation. The true yield coefficients (Y_{SX}^{true}) decreased in NOX and cNDK, where cytosolic NADH level was decreased, but increased when mitochondrial NADH was lowered (in AOX and mNDK), compared with REF (table 5.3). True biomass yield on oxygen (Y_{OX}^{true}) decreased by 50% in NOX compared to REF. The apparent reduction in the biomass yield is due to increased oxygen consumption by NADH oxidase. However, the biomass yield on oxygen for the other strains increased, most predominantly for AOX and mNDK, where mitochondrial redox level was decreased. Increased biomass yield on glucose as well as on oxygen in the strains with additional NADH oxidation capacity in the mitochondria suggests an inhibitory role for NADH in the electron transport chain.

Table 5.3 Yield coefficients of biomass from glucose and oxygen for the five recombinant strains and the reference strain as determined from the accelerostat experiments using the Pirt equation.

	Y_{SX}^{true}	Y_{OX}^{true}	Y_{XATP}	*PO*
REF	0.52	59.79	85.51	1.23
NOX	0.38	33.75	106.02	0.82
AOX	0.56	118.25	76.95	1.50
cNDK	0.46	76.27	70.09	1.34
mNDK	0.58	111.60	85.53	1.43
UDH	0.48	98.85	80.44	1.18

Estimated ATP requirement for biomass synthesis and the operational PO ratio is also shown.
[a] True biomass yield coefficient from glucose (g DCW/g glucose).
[b] True biomass yield coefficient from oxygen (mg DCW/mmol O_2).
[c] ATP requirement for biomass synthesis (mmol/g DCW).

5.4.4 Bioenergetic parameters

We used a stoichiometric model to estimate the bioenergetic parameters (Y_{XATP} and PO) in the different strains. The ATP requirement for biomass synthesis in AOX and cNDK decreased (about 10–15%). However, the NOX strain appeared to have an increased demand for ATP (table 5.3). The differences in the biomass yield coefficients also manifested in the estimation of the PO ratio. The efficiency of ATP synthesis appeared to have improved in the AOX and mNDK strains (table 5.3), where mitochondrial NADH was decreased. The substantial decrease in the value of PO ratio for the NOX strain is due to the additional oxygen consumption by NADH oxidase. However, our estimations indicated that the PO ratio for the UDH strain slightly decreased. Given the over-simplification in the model scheme and the error involved in the estimations, it is not clear if this decrease has any biological relevance.

5.5 Discussion

The response to cofactor manipulation, in general, is a network property since it is the consequence of the interaction of a large number of components of the metabolic network. The fundamental anabolic and catabolic processes are directly affected in this perturbation. Hence, the results presented here will prove valuable for a quantitative understanding of the relationship between carbon and energy metabolism. We introduced five perturbations in *S. cerevisiae* that directly perturbed the intracellular level of NADH (enhancing cytosolic NADH oxidation, enhancing mitochondrial NADH oxidation, ATP-dependent conversion of NADH to NADPH in the cytosol, ATP-dependent conversion of NADH to NADPH in the mitochondria and the interconversion of NADH and NADPH). The impact of these perturbations on glucose and energy metabolism was evaluated. Since cellular redox and energy levels are intimately connected in the electron transport chain, studying one of these in isolation is impossible.

5.5.1 Bioenergetics and product formation

The metabolic and bioenergetic response of *S. cerevisiae* to these perturbations indicates clear compensatory mechanisms to adjust to new cofactor levels in the cytosol and mitochondria. We used a simple stoichiometric model to estimate the efficiency of oxidative phosphorylation (*PO* ratio) and the energy expense for biomass synthesis (Y_{XATP}). Increased values for Y_{SX}^{true} and *PO* when mitochondrial NADH was lowered (AOX and mNDK) reflect greater efficiency of respiration. Since NADH is the substrate for the respiratory chain, it appears that excess NADH inhibits ATP synthesis. It is evident that there is an optimal level of mitochondrial redox level for efficient operation of oxidative phosphorylation. It was previously reported that decreasing the rate of ATP synthesis can increase the *PO* ratio (Sonnleitner and Kappeli 1986;

Huttemann *et al.* 2008). Increased efficiency of respiration was evident in lower ethanol production and increased growth rate in AOX and mNDK, likely due to higher split ratio of metabolic flux into the TCA cycle. We observed increased μ_{max} when mitochondrial NADPH level was increased. It appears that draining ATP from biomass synthesis to NADPH generation has favorable results for *S. cerevisiae*. A similar result was previously observed in *Aspergillus nidulans* (Panagiotou *et al.* 2009), exposing a gap in our fundamental understanding of the relative importance of NADPH and ATP in biomass synthesis in the cytosol and the mitochondria. In contrast, overexpressing cytosolic NADH kinase decreased μ_{max} and Y_{SX}. We attribute this to excess drain of NADH and ATP to synthesize NADPH, as indicated by a substantially higher $NADPH/NADP^+$ and a lower $NADH/NAD^+$ ratio in cNDK (figure 5.2). Based on these results, we believe that there is an optimum level of cytosolic NADH kinase activity that is conducive to increasing μ_{max}. On the contrary, mere oxidation of cytosolic NADH to NAD^+ lowered glycerol production. When cytosolic NADH was converted to NADPH at the expense of ATP, we additionally observed a decrease in biomass yield from glucose. This is likely due to the significant increased energy dissipation in phosphorylating NADH. NADH kinase not only mediates the non-respiratory dissipation of NADH, it also consumes ATP directly, as indicated by lower ATP/ADP level, and substantially higher $NADPH/NADP^+$ level in cNDK. Although we see clear correlation between the redox state and the *PO* ratio, the precise biochemical mechanisms cannot be conclusively deciphered due to the still unanswered questions regarding the exact stoichiometry of the ATP synthesis and proton transport in the electron transport chain (Hinkle 2005).

5.5.2 Transhydrogenation in *S. cerevisiae*

Previously, membrane-bound, proton translocating transhydrogenase (*pntAB*) from *E. coli* and the soluble transhydrogenase from *A. vinelandii* have been shown to convert NADPH to NADH in *S. cerevisiae* (Anderlund *et al.* 1999; Nissen *et al.* 2001). In this

study, during growth on glucose, we observed a similar result using a soluble transhydrogenase from *E. coli*. We did not see an increase in the NADH/NAD^+ ratio in the presence of the transhydrogenase, as the excess NADH production was likely compensated by excess glycerol production. However, during growth on ethanol, the production of NADPH was favored according to the thermodynamic analysis based on the concentration of the cofactors, as well as in the NADPH/$NADP^+$ ratio. An increased demand for NADPH for biomass synthesis could be the reason, as no glucose is available to feed the pentose phosphate pathway. Increased departure of the cNDK, mNDK and the UDH strains from REF in figure 5.5 reflects the major role played by NADPH during growth on ethanol. On the other hand, growth on glucose appeared to be most sensitive to cytosolic NADH, as indicated by the farthest divergence of the cNDK and NOX strains from REF.

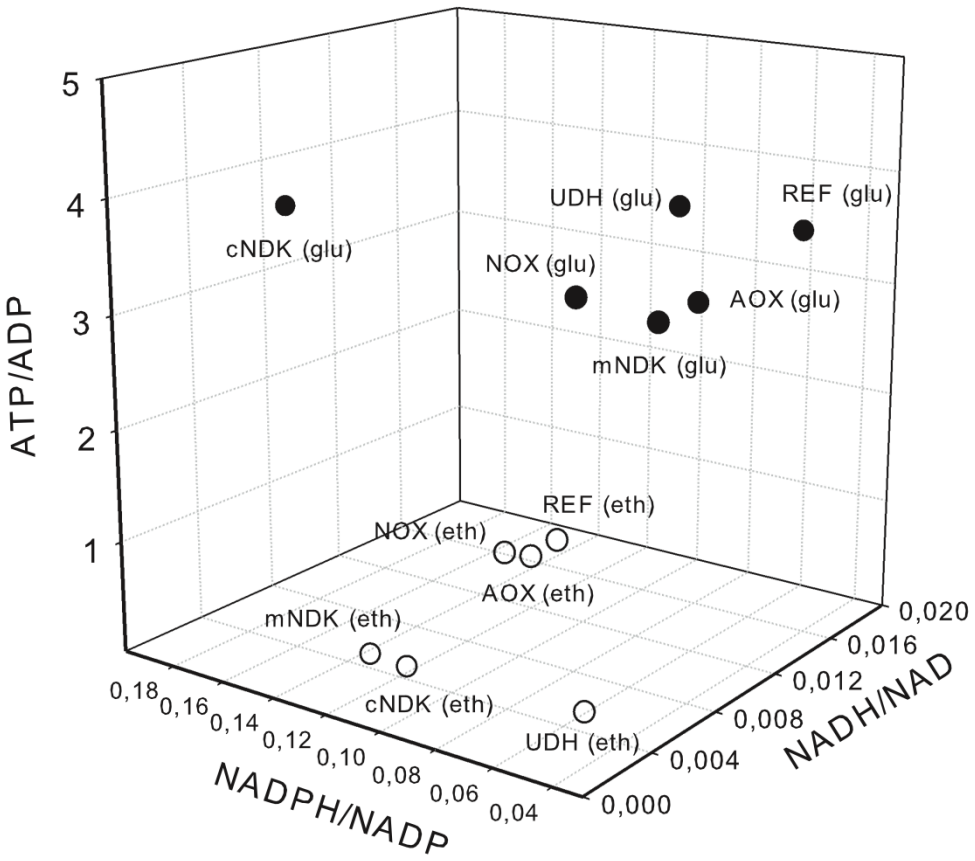


Figure 5.5 Ratios of metabolic cofactor couples in strains with perturbed redox metabolism during exponential growth on glucose (solid circles) and subsequent growth on ethanol (open circles) after all glucose has been depleted.

The results summarize the impact of perturbation in redox and energy on metabolism and product formation in *S. cerevisiae*. Changes in the cofactor pools induce changes at the transcriptional level as well as the enzyme level (Vemuri *et al.* 2007). It is also very likely that many of these changes are brought about by metabolites. One such example is the inhibition of respiration by fructose 1,6-bisphosphate (Diaz-Ruiz *et al.* 2008). The intracellular level of FBP measured during growth on glucose correlated with specific ethanol production rate and inversely correlated with specific oxygen consumption rate in the different strains we studied. Overall, the results presented here serve as the basis for interpreting the impact of perturbations that involve cofactors in *S. cerevisiae* and eukaryotic cells in general.

5.6 Acknowledgements

The authors thank Prof. Jens Nielsen for his comments on manuscript, Bianca Klein for assistance with metabolite samples and Tina Johanssen for technical help with the operation of the fermenters. JH acknowledges fellowship from China Scholarship Council and NFL acknowledges Ph.D. grant from Fundação para a Ciência e Tecnologia – Ministério da Ciência, Tecnologia e Ensino Superior, Portugal (SFRH/BD/21947/2005).

6 Concluding remarks

The glyoxalase pathway has the unusual feature of having a substrate formed in a non-enzyme-catalyzed reaction. Whether it is more accurate to assume, from a kinetic point of view, that this substrate forms prior to binding to the enzyme or that glutathione first binds to the active centre and the hemithioacetal is formed with the subsequent binding of methylglyoxal, is a problem that has been investigated but not conclusively solved in the last thirty years (Bartfai and Mannervik 1972; Bartfai *et al.* 1973; Mannervik *et al.* 1974). As discussed in chapter 2, each one of the two mechanisms may lead to completely different conclusions regarding the response of the pathway to perturbations. This is of especial importance for an enzyme system which inhibition was considered to be a potential therapeutic strategy (Thornalley 1993; 2008).

The differences observed for the behavior of the pathway as predicted by the two models suggested the design of an experimental approach to discriminate which of the two kinetic models (one-substrate or the sequential-two-substrate) better describes the kinetics of the pathway. For this analysis (presented in chapter 3) the chosen strategy comprised the acquisition of a set of S-D-lactoylglutathione time-course data representative of the behavior of the pathway for a wide range of substrate concentrations, both with purified glyoxalase I and the more *in vivo*-resembling digitonin-permeabilized cells setup, and their analysis within a model selection framework from Information Theory. The results of these experiments were still not conclusive regarding the model discrimination question. Nevertheless, valuable observations resulted from this work.

The use of several time courses for the parameter estimation and subsequent computation model discrimination analysis is tempting as that means to use more information from the experimental system. This might be indeed desirable for experiments aiming at obtaining more accurate parameter estimates. Using several time courses, however, made parameter estimation more difficult since different parameter values might fit the experimental data giving similar optimization scores – a poorer fit to a time course might be compensated by a better fitting to another time course.

Therefore, while a more complete experimental data set is expected to improve accuracy of parameter estimates, this may result in an increase of the error associated to the estimates and parameter epistasis. Also, a large amount of data does not necessarily make the selection of a model easier, as results in chapter 3 show. For this specific problem, other strategy than just acquiring a large data set and applying statistical analysis had to be followed. This was the motivation behind the development of the procedure presented in chapter 4.

In chapter 4 a computational procedure was proposed to optimize the design of experiments for model discrimination problems. The method was developed to give a set of optimized experimental conditions for which the time courses predicted by the candidate models for the observable variables (concentrations) have a maximum divergence.

The measure of divergence employed was the Kullback-Leibler distance. It proved to be an adequate measure for this kind of problem as shown by the practical example given for the glyoxalase pathway – it was possible to find experimental conditions for which the models predicted very different intermediate concentrations along time. Even for assays using permeabilized yeast cells, which are more restrictive regarding the number of experimental variables that can be controled, a sufficient difference was obtained for discrimination since for one of the models the maximum concentration measured in the assay was significantly (1.6 fold) higher.

A multiobjective optimization strategy was employed in chapter 4 for two reasons. First, the Kulback-Leibler distance (equation 4.3) is a directed measure of the discrepancy between the models, which means that computing the distance from model A to model B has a different result from computing the distance from model B to model A (Kullback and Leibler 1951; Amisaki and Eguchi 1995; Burnham and Anderson 1998). In the present work the Kullback-Leibler distance was computed in both directions. Secondly, the method was intended to be used not only to discriminate between two models but also for problems with a higher number of candidate models. Pareto optimality is the most appropriate criterion to solve multiobjective optimization problems if all the objective functions are considered equally important – otherwise strategies exist to condense the different objective functions in a single one (Abraham *et*

al. 2005). This is the case of a model discrimination problem: every candidate model is considered equally likely to be the one that better approximates to the real system. State-of-art algorithms (generalized differential evolution (Kukkonen and Lampinen 2005), divide-and-conquer sorting (Fang *et al.* 2008) and *k*-nearest neighbor search (Cover and Hart 1967)) were used to develop a computer program to find the optimized experimental conditions. The method was able to find optimized substrate and enzyme concentrations to discriminate the models. Optimization of these variables alone was sufficient to obtain very different predictions for each model.

For assays with purified enzymes a much clearer difference between the S-D-lactoylglutathione time courses was obtained. While with permeabilized cells only substrate concentrations could be controlled, with purified enzymes the experimenter could define enzyme concentrations as well – if more experimental variables are available for manipulation, a higher divergence between the models can be achieved.

The motivation for the research on the glyoxalase pathway is generally methylglyoxal's toxicity but the results presented here systematically point to the key metabolic role of glutathione. The properties of the glyoxalase pathway identified in chapter 2 seem to arise from the regeneration of glutathione, namely the conditions for achievement of a steady state (inequalities 2.9 and 2.10) that determine for enzyme activities a threshold higher than the flux of the pathway. In addition, the physiological activity of glyoxalase II, much higher than necessary to keep a low methylglyoxal concentration, is indicative of the importance of the regeneration of free glutathione. Glutathione regeneration was also very useful for the discrimination procedure presented in chapter 4. For an experiment excluding glyoxalase II the optimal solutions were substrate concentrations as high as allowed, reflecting the different glyoxalase I activities estimated for the two models. In the presence of glyoxalase II, with regeneration of glutathione, the optimization algorithm also explored other differences between the models, such as the possibility of limiting the flux of the pathway to the formation of the hemithioacetal (which was possible for model 1).

The essential role of glutathione in detoxifying processes, not only as cofactor of the glyoxalase pathway and gluthathionyl-S-transferase systems but also as an antioxidant (Meister and Anderson 1983; Meister 1994), is closely related to the redox status of the

cell. The main determinants of the cell redox status are the pyridine nucleotides, NADH/NAD^+ and NADPH/$NADP^+$ (Nelson and Cox 2005). While NADH is generated in catabolic pathways and feeds the respiratory chain for the synthesis of ATP, NADPH is formed mainly in the pentose-phosphate pathway and plays the role of reducing agent in anabolic pathways (Nelson and Cox 2005). In chapter 5 the effect of perturbations to the levels of pyridine nucleotides in living yeast cells was evaluated.

Some surprising results were observed due to the perturbations described in chapter 5. NADH is the substrate of the respiratory chain; however a decrease of the NADH level in the mitochondrion leads to increased ATP synthesis efficiency and true yield of biomass on glucose. Separate pools of NADH/NAD^+ and NADPH/$NADP^+$ exist in all kinds of cell and both ratios are buffered to secure that energy and synthetic activity demands are met (Nelson and Cox 2005). Some organisms express transhydrogenase that catalyzes the interconversion of the two pyridine nucleotides (Anderlund *et al.* 1999; Nissen *et al.* 2001). This enzyme is thought to allow these organisms to handle extreme NADH or NADPH-demanding situations (Hoek and Rydstrom 1988). The soluble transhydrogenase from *E. coli* was expressed in *S. cerevisiae* and shown to catalyse the conversion of NADPH to NADH during growth on glucose, meaning that a high [NADPH] [NAD^+] / ([NADH] [$NADP^+$]) ratio was disrupted with the introduction of the enzyme. This observation suggests that yeast is adapted not only to consume glucose at a very high rate but also to use it to generate a very high reducing potential for the biosynthetic pathways. Also, the introduction of an enzyme that converted NADH to NADPH at the expense of ATP lead to an increase of specific growth rate. The role of NADPH in detoxifying processes may also contribute to explain these observations. The relative importance of NADPH and ATP in biosynthetic pathways and which regulatory mechanisms are on the basis of these effects are interesting question to explore in the future.

Appendix A

Expanded functions of system parameters for model 1:

$E_{1A} = V_{1A} - v_{in}$

$E_{2A} = V_{2A} - v_{in}$

$A_1 = k_{-1} K_{m1} + E_{1A}$

$B_1 = K_{m2A} E_{1A} + K_{m1} E_{2A}$

$C_1 = k_{-1} K_{m1} + V_{1A}$

$D_1 = k_{-1} K_{m1} + L_1$

$F_1 = SH_{Total} L_1 + M_1$

$G_1 = SH_{Total} k_{-1} + v_{in}$

$I_1 = SH_{Total} + K_{m1} + K_{m2A}$

$J_1 = D_1 F_1 + 3 SH_{Total} V_{1A} V_{2A} - V_{2A} C_1 I_1$

$L_1 = V_{1A} + V_{2A}$

$M_1 = K_{m1} V_{2A} + K_{m2A} V_{1A}$

$N_1 = V_{2A} C_1 F_1 + SH_{Total} D_1 V_{1A} V_{2A}$

$P_1 = D_1 F_1 + V_{2A} C_1 I_1 + SH_{Total} V_{1A} V_{2A}$

$Q_1 = F_1 + D_1 I_1$

$$\alpha_1 = k_{-1}\, K_{m1} - v_{in}$$

$$\beta_1 = \tau_1 + K_{m1}\, E_{2A}$$

$$\gamma_1 = E_{2A} - k_{-1}\, K_{m2A}$$

$$\delta_1 = k_{-1}\, SH_{Total}\, E_{2A} + v_{in}\, \gamma_1$$

$$\varepsilon_1 = \alpha_1 - v_{in}$$

$$\eta_1 = SH_{Total}\, \varepsilon_1\, E_{2A} - v_{in}\, \sigma_1$$

$$\tau_1 = K_{m1}\, k_{-1} + E_{1A}^{\,2}$$

$$\lambda_1 = \psi_1 + \mu_1$$

$$\mu_1 = \psi_1 + K_{m2A}\, E_{1A}$$

$$\xi_1 = SH_{Total}\, (K_{m1}\, k_{-1}\, V_{1A} + E_{1A}^{\,2}) + K_{m1A}\, v_{in}$$

$$\sigma_1 = K_{m1}\, \varphi_1 - 2\, K_{m2A}\, v_{in}$$

$$\tau_1 = SH_{Total}\, E_{2A} - K_{m2A}\, v_{in}$$

$$\varphi_1 = k_{-1}\, K_{m2A} + E_{2A}$$

$$\psi_1 = SH_{Total}\, E_{1A} - K_{m1}\, v_{in}$$

$$\chi_1 = \psi_1 + K_{m2A}\, V_{1A}$$

$$\omega_1 = SH_{Total}\, E_{2A}^{\,2} + K_{m2A}\, v_{in}$$

Expanded functions of system parameters for model 2:

$A_2 = SH_{Total}\, k_{-1}\, E_{2B}\, F_2 - v_{in}\, L_2$

$B_2 = k_{-1}\, V_{1B} + k_1\, K_{m12}\, v_{in}$

$C_2 = -4\, K_{m11}\, k_{-1}\, E_{2B}\, v_{in}\, Q_2\, T_2 + U_2^2$

$D_2 = k_{-1}\, K_{m11}\, K_{m12}\, v_{in}\, E_{2B} + k_{-1}\, P_2$

$F_2 = k_{-1}\, V_{1B}\, E_{1B} + k_1\, K_{m12}\, v_{in}\, M_2$

$G_2 = k_{-1} + k_1\, K_{m12}$

$H_2 = k_{-1} - k_1\, K_{m12}$

$J_2 = k_{-1}\, G_2\, V_{1B} - k_1\, K_{m12}\, H_2\, v_{in}$

$L_2 = K_{m11}\, E_{2B}\, J_2 + K_{m2B}\, k_{-1}\, F_2$

$M_2 = V_{1B} + v_{in}$

$N_2 = 2\, V_{1B} - v_{in}$

$P_2 = -K_{m2B}\, E_{1B}\, v_{in} + SH_{Total}\, E_{1B}\, E_{2B} + K_{m11}\, N_2\, E_{2B}$

$Q_2 = k_{-1}\, E_{1B} + k_1\, K_{m12}\, v_{in}$

$R_2 = SH_{Total} + K_{m2B}$

$S_2 = K_{m11}\, H_2\, E_{2B} - K_{m2B}\, k_{-1}\, E_{1B}$

$T_2 = SH_{Total}\, V_{2B} - R_2\, v_{in}$

$U_2 = SH_{Total}\, k_{-1}\, E_{1B}\, E_{2B} + v_{in}\, S_2$

$W_2 = SH_{Total} + K_{m2B}$

$X_2 = k_1\, K_{m12}\, v_{in} + Q_2$

$Y_2 = SH_{Total}\, V_{2B} - W_2\, v_{in}$

$Z_2 = K_{m11}\, H_2\, E_{2B} - K_{m2B}\, k_{-1}\, E_{1B}$

$\alpha_2 = K_{m11}\, H_2\, E_{2B} + K_{m2B}\, X_2$

$\beta_2 = SH_{Total}\, k_{-1}\, E_{1B}\, E_{2B} + v_{in}\, Z_2$

$\gamma_2 = -\, 4\, K_{m11}\, k_{-1}\, E_{2B}\, v_{in}\, Q_2\, Y_2 + \beta_2^2$

$\delta_2 = SH_{Total}\, E_{2B}\, X_2 - v_{in}\, \alpha_2$

$\varepsilon_2 = K_{m11}\, H_2\, v_{in} + SH_{Total}\, k_{-1}\, E_{1B}$

$\varphi_2 = -\, K_{m2B}\, k_{-1}\, E_{1B}\, v_{in} + E_{2B}\, \varepsilon_2$

Expanded functions of system parameters for model 3:

$A_3 = -\, E_{2B}\, B_3 + K_{m2B}\, v_{in}$

$B_3 = SH_{Total} + K_{m11}$

$C_3 = E_{2B}\, K_{m11} + E_{1B}\, K_{m2B}$

$D_3 = E_{2B}\, SH_{Total} - K_{m2B}\, v_{in}$

$F_3 = SH_{Total} + K_{m11}$

$G_3 = E_{2B}\, F_3 - K_{m2B}\, v_{in}$

$H_3 = SH_{Total}\, k_1\, K_{m12} + k_1\, K_{m11}\, K_{m12} - K_{m11}\, k_{-1}$

$I_3 = k_1\, V_{1B} + k_1\, K_{m12}\, v_{in} - k_{-1}\, v_{in}$

$J_3 = E_{2B}\, H_3 - K_{m2B}\, I_3$

$L_3 = E_{1B}\, E_{2B}\, SH_{Total}\, k_{-1} + v_{in}\, J_3$

$M_3 = E_{1B}\, E_{2B}\, SH_{Total} - E_{2B}\, K_{m11}\, v_{in} - E_{1B}\, K_{m2B}\, v_{in}$

Appendix B

The following reactions are considered in the simple stoichiometric model illustrated in figure 5.1 as described previously (Nielsen *et al.* 2003). For the sake of simplicity, all the reactions were expressed in a C-mole basis (for example, glucose, $C_6H_{12}O_6$ is expressed as CH_2O).

1. Biomass growth (μ): - 1.105 CH_2O + $CH_{1.82}O_{0.58}N_{0.16}$ + 0.105 CO_2 + 0.355 NADH - Y_{XATP} ATP - 0.231 NADPH = 0.

2. Pentose phosphate pathway (V_{pp}): - CH_2O + CO_2 + 2 NADPH = 0.

3. Complete glucose oxidation by catabolism in the EMP pathway and the TCA cycle (V_{cat}): - CH_2O + CO_2 + 2 NADH + 2/3 ATP = 0.

4. Ethanol production from glucose (V_{eth}): - CH_2O + 1/3 ATP + 1/3 CO_2 + 2/3 $CH_3O_{1/2}$ = 0.

5. Glycerol production (V_{gly}): - CH_2O - 1/3 ATP - 1/3 NADH + $CH_{8/3}O$ = 0.

6. Oxidative phosphorylation (V_{op}): *PO* ATP + 1/2 O_2 - NADH = 0.

7. Non-growth associated ATP consumption (*mATP*): - ATP = 0.

In these equations, Y_{XATP} is the amount of ATP required to produce 1g of biomass and *PO* is the efficiency of oxidative phosphorylation (moles of ATP formed per mole of atomic oxygen consumed). The reaction network described by the seven equations can be written in a condensed form by setting up a mass balance around the nine metabolites according to the following equation.

$$\begin{pmatrix} -1.105 & -1 & -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1/2 & 0 \\ 0.105 & 1 & 1 & 0 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2/3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -Y_{XATP} & 0 & 2/3 & -1/3 & 1/3 & PO & -1 \\ 0.355 & 0 & 2 & -1/3 & 0 & -1 & 0 \\ -0.231 & 2 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ V_{pp} \\ V_{cat} \\ V_{gly} \\ V_{eth} \\ V_{op} \\ m_{ATP} \end{pmatrix} = \begin{pmatrix} -r_{glu} \\ -r_{O_2} \\ r_{CO_2} \\ r_{eth} \\ r_{gly} \\ \mu \\ r_{ATP} \\ r_{NADPH} \\ r_{NADPH} \end{pmatrix}$$

The turnover rates for the cofactors ATP, NADH and NADPH is negligible compared to the consumption/production rates of the other metabolites. By setting up balances for the cofactors, it is possible to link the three unknown reaction rates (V_{pp}, V_{op} and V_{cat}) with the bioenergetic parameters (Y_{XATP}, PO and m_{ATP}). By eliminating the unknown reaction rates using algebraic manipulation, an equation for glucose uptake rate can be derived that involves the bioenergetic parameters.

$$r_{glu} = \frac{0.814 - 2.086\,PO + Y_{XATP}}{0.667 + 2\,PO}\mu + \frac{m_{ATP}}{0.667 + 2\,PO}$$

This equation is analogous to the Pirt equation (Nielsen *et al.* 2003), $r_{glu} = (1/Y_{SX}^{true})\mu + m_s$, where Y_{SX}^{true} is the true biomass yield coefficient on glucose during completely respiratory metabolism and m_s is the non-growth associated consumption of glucose (maintenance). By regressing specific glucose uptake rate with the specific growth rate, the values of Y_{SX}^{true} and m_s were calculated (table 5.3), from which the values of Y_{XATP} and m_{ATP} were estimated by least squares regression. For all the calculations, the specific growth rate at any instant was calculated as $\mu = D + (1/X)\,(dX/dt)$, where D is the dilution rate (h^{-1}) and X is the biomass concentration (g/L).

7 References

Abdulnur, S. F. (1976). The interactions of glyoxals with proteins and DNA in relation to cancer. *International Journal of Quantum Chemistry, Quantum Biology Symposium* **3**: 59-64.

Abraham, A.*, et al.* (2005). *Evolutionary Multiobjective Optimization: Theoretical Advances and Applications*. Springer-Verlag, London.

Ahmed, M. U.*, et al.* (1997). N-epsilon-(carboxyethyl)lysine, a product of the chemical modification of proteins by methylglyoxal, increases with age in human lens proteins. *Biochem J* **324**: 565-570.

Ahmed, N.*, et al.* (2002). Assay of advanced glycation endproducts (AGEs): surveying AGEs by chromatographic assay with derivatization by 6-aminoquinolyl-N-hydroxysuccinimidyl-carbamate and application to Nepsilon-carboxymethyl-lysine- and Nepsilon-(1-carboxyethyl)lysine-modified albumin. *Biochem J* **364**(Pt 1): 1-14.

Ahmed, N. and P. J. Thornalley (2002). Chromatographic assay of glycation adducts in human serum albumin glycated in vitro by derivatization with 6-aminoquinolyl-N-hydroxysuccinimidyl-carbamate and intrinsic fluorescence. *Biochem J* **364**(Pt 1): 15-24.

Ahmed, N. and P. J. Thornalley (2007). Advanced glycation endproducts: what is their relevance to diabetic complications? *Diabetes Obes Metab* **9**(3): 233-245.

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*. B. N. Petrov and F. Csaki. Akademiai Kiado, Budapest.

Akaike, H. (1974). New Look at Statistical-Model Identification. *Ieee Transactions on Automatic Control* **Ac19**(6): 716-723.

Aleksandrovskii, Y. A. (1992). Antithrombin III, C1 inhibitor, methylglyoxal, and polymorphonuclear leukocytes in the development of vascular complications in diabetes mellitus. *Thromb Res* **67**(2): 179-189.

Allen, R. E.*, et al.* (1993). Purification and characterisation of glyoxalase II from human red blood cells. *Eur J Biochem* **213**(3): 1261-1267.

Amisaki, T. and S. Eguchi (1995). Pharmacokinetic parameter estimations by minimum relative entropy method. *J Pharmacokinet Biopharm* **23**(5): 479-494.

Anderlund, M.*, et al.* (1999). Expression of the Escherichia coli pntA and pntB genes, encoding nicotinamide nucleotide transhydrogenase, in Saccharomyces cerevisiae and its effect on product formation during anaerobic glucose fermentation. *Appl Environ Microbiol* **65**(6): 2333-2340.

Aronsson, A. C.*, et al.* (1978). Glyoxalase I, a zinc metalloenzyme of mammals and yeast. *Biochem Biophys Res Commun* **81**(4): 1235-1240.

Ashlock, D. (2006). *Evolutionary Computation for Modeling and Optimization*. Springer, New York.

Auzat, I.*, et al.* (1999). The NADH oxidase of Streptococcus pneumoniae: its involvement in competence and virulence. *Mol Microbiol* **34**(5): 1018-1028.

Baggetto, L. G. and A. L. Lehninger (1987). Isolated tumoral pyruvate dehydrogenase can synthesize acetoin which inhibits pyruvate oxidation as well as other aldehydes. *Biochem Biophys Res Commun* **145**(1): 153-159.

Bakhti, M.*, et al.* (2007). Consequential alterations in haemoglobin structure upon glycation with fructose: prevention by acetylsalicylic acid. *J Biochem* **141**(6): 827-833.

Bakker, B. M.*, et al.* (2001). Stoichiometry and compartmentation of NADH metabolism in Saccharomyces cerevisiae. *FEMS Microbiol Rev* **25**(1): 15-37.

Bartfai, T. and B. Mannervik (1972). A procedure based on statistical criteria for discrimination between steady state kinetic models. *FEBS Lett* **26**(1): 252-256.

Bartfai, T*., et al.* (1973). Discrimination between steady-state kinetic models of the Mechanism of action of yeast glyoxalase I. *Biochemistry* **12**(3): 387-391.

Baynes, J. W. and S. R. Thorpe (1999). Role of oxidative stress in diabetic complications: a new perspective on an old paradigm. *Diabetes* **48**(1): 1-9.

Bito, A*., et al.* (1997). Identification and phenotypic analysis of two glyoxalase II encoding genes from Saccharomyces cerevisiae, GLO2 and GLO4, and intracellular localization of the corresponding proteins. *J Biol Chem* **272**(34): 21509-21519.

Bito, A*., et al.* (1999). Heterologous expression, purification, and kinetic comparison of the cytoplasmic and mitochondrial glyoxalase II enzymes, Glo2p and Glo4p, from Saccharomyces cerevisiae. *Protein Expr Purif* **17**(3): 456-464.

Boonstra, B*., et al.* (1999). The udhA gene of Escherichia coli encodes a soluble pyridine nucleotide transhydrogenase. *J Bacteriol* **181**(3): 1030-1034.

Bouma, B*., et al.* (2003). Glycation induces formation of amyloid cross-beta structure in albumin. *J Biol Chem* **278**(43): 41810-41819.

Bro, C*., et al.* (2004). Genome-wide transcriptional response of a Saccharomyces cerevisiae strain with an altered redox metabolism. *Biotechnol Bioeng* **85**(3): 269-276.

Bruggeman, F. J. and H. V. Westerhoff (2007). The nature of systems biology. *Trends Microbiol* **15**(1): 45-50.

Bruinenberg, P. M*., et al.* (1983). A Theoretical-Analysis of NADPH Production and Consumption in Yeasts. *Journal of General Microbiology* **129**(Apr): 953-964.

Bucala, R*., et al.* (1984). Modification of DNA by reducing sugars: a possible mechanism for nucleic acid aging and age-related dysfunction in gene expression. *Proc Natl Acad Sci U S A* **81**(1): 105-109.

Bucala, R. and A. Cerami (1992). Advanced glycosylation: chemistry, biology, and implications for diabetes and aging. *Adv Pharmacol* **23**: 1-34.

Burnham, K. P. and D. R. Anderson (1998). *Model Selection and Inference: A Practical Information-Theoretic Approach*. Springer-Verlag, New York.

Cajelli, E*., et al.* (1987). Methylglyoxal-induced mutation to 6-thioguanine resistance in V79 cells. *Mutat Res* **190**(1): 47-50.

Cameron, A. D*., et al.* (1997). Crystal structure of human glyoxalase I--evidence for gene duplication and 3D domain swapping. *EMBO J* **16**(12): 3386-3395.

Cameron, A. D*., et al.* (1999). Crystal structure of human glyoxalase II and its complex with a glutathione thiolester substrate analogue. *Structure* **7**(9): 1067-1078.

Carlson, M. (1999). Glucose repression in yeast. *Curr Opin Microbiol* **2**(2): 202-207.

Casazza, J. P*., et al.* (1984). The metabolism of acetone in rat. *J Biol Chem* **259**(1): 231-236.

Castellani, R*., et al.* (1996). Glycoxidation and oxidative stress in Parkinson disease and diffuse Lewy body disease. *Brain Res* **737**(1-2): 195-200.

Chen, F*., et al.* (2004). Role for glyoxalase I in Alzheimer's disease. *Proc Natl Acad Sci U S A* **101**(20): 7687-7692.

Clarke, D. M*., et al.* (1986). Nucleotide sequence of the pntA and pntB genes encoding the pyridine nucleotide transhydrogenase of Escherichia coli. *Eur J Biochem* **158**(3): 647-653.

Cliffe, E. E. and S. G. Waley (1961). The mechanism of the glyoxalase I reaction, and the effect of ophthalmic acid as an inhibitor. *Biochem J* **79**: 475-482.

Coello Coello, C. A*., et al.* (2007). *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, New York.

Cooper, R. A. and A. Anderson (1970). The formation and catabolism of methylglyoxal during glycolysis in Escherichia coli. *FEBS Lett* **11**(4): 273-276.

Cordeiro, C. and A. P. Freire (1995). Digitonin permeabilization of Saccharomyces cerevisiae cells for in situ enzyme assay. *Anal Biochem* **229**(1): 145-148.

Cordeiro, C. (1996). *Estudos in situ do metabolismo do metilglioxal na levedura Saccharomyces cerevisiae*. Tese de Doutoramento. Faculdade de Ciências da Universidade de Lisboa, Lisboa.

Cordell, P. A*., et al.* (2004). The Human hydroxyacylglutathione hydrolase (HAGH) gene encodes both cytosolic and mitochondrial forms of glyoxalase II. *J Biol Chem* **279**(27): 28653-28661.

Cornish-Bowden, A. (2005). *Fundamentals of Enzyme Kinetics*. Portland Press, London.

Cover, T. M. and P. E. Hart (1967). Nearest Neighbor Pattern Classification. *Ieee Transactions on Information Theory* **It13**(1): 21-+.

Crabtree, H. G. (1928). The carbohydrate metabolism of certain pathological overgrowths. *Biochem J* **22**(5): 1289-1298.

Crabtree, H. G. (1929). Observations on the carbohydrate metabolism of tumours. *Biochem J* **23**(3): 536-545.

Creighton, D. J*., et al.* (1988). Optimization of efficiency in the glyoxalase pathway. *Biochemistry* **27**(19): 7376-7384.

Creighton, D. J*., et al.* (2003). Glyoxalase I inhibitors in cancer chemotherapy. *Biochem Soc Trans* **31**(Pt 6): 1378-1382.

Dakin, H. D. and H. W. Dudley (1913a). An enzyme concerned with the formation of hydroxy acids from ketonic aldehydes. *J Biol Chem* **14**: 155-157.

Dakin, H. D. and H. W. Dudley (1913b). On glyoxalase. *J Biol Chem* **14**: 423-431.

Dasgupta, D. (1999). Information Processing in the Immune System. *New Ideas in Optimization*. D. Corne, M. Dorigo and F. Glover. McGraw-Hill, London**:** 161-166.

Davis, K. A. and G. R. Williams (1966). Cation activation of glyoxalase I. *Biochim Biophys Acta* **113**: 393-395.

De Deken, R. H. (1966). The Crabtree effect: a regulatory system in yeast. *J Gen Microbiol* **44**(2): 149-156.

Deb, K*., et al.* (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *Ieee Transactions on Evolutionary Computation* **6**(2): 182-197.

Deponte, M*., et al.* (2007). Allosteric coupling of two different functional active sites in monomeric Plasmodium falciparum glyoxalase I. *J Biol Chem* **282**(39): 28419-28430.

Diaz-Ruiz, R*., et al.* (2008). Mitochondrial oxidative phosphorylation is regulated by fructose 1,6-bisphosphate. A possible role in Crabtree effect induction? *J Biol Chem* **283**(40): 26948-26955.

Dorigo, M. and G. Di Caro (1999). The Ant Colony Optimization Meta-Heuristic. *New Ideas in Optimization*. D. Corne, M. Dorigo and F. Glover. McGraw-Hill, London**:** 11-32.

Driessen, H. P*., et al.* (1985). The mechanism of N-terminal acetylation of proteins. *CRC Crit Rev Biochem* **18**(4): 281-325.

Du Yan, S*., et al.* (1997). Amyloid-beta peptide-receptor for advanced glycation endproduct interaction elicits neuronal expression of macrophage-colony stimulating factor: a proinflammatory pathway in Alzheimer disease. *Proc Natl Acad Sci U S A* **94**(10): 5296-5301.

Dukic-Stefanovic, S*., et al.* (2001). AGES in brain ageing: AGE-inhibitors as neuroprotective and anti-dementia drugs? *Biogerontology* **2**(1): 19-34.

Edgeworth, F. Y. (1881). *Mathematical Physics*. P. Keagan, London.

Entian, K.-D. and H.-J. Schüller (1997). Glucose Repression (Carbon Catabolite Repression) in Yeast. *Yeast Sugar Metabolism: Biochemistry, Genetics, Biotechnology and Applications*. F. K. Zimmermann and K.-D. Entian. CRC Press, Boca Raton, FL.

Fang, H. B.*, et al.* (2008). An efficient non-dominated sorting method for evolutionary algorithms. *Evolutionary Computation* **16**(3): 355-384.

Fell, D. (1996). *Understanding the Control of Metabolism*. Portland Press, London.

Ferreira, A. E. N. (2000). in Voit, E.O. *Computational Analysis of Biochemical Systems - A Practical Guide for Biochemists and Molecular Biologists*. Cambridge University Press, Cambridge, http://www.dqb.fc.ul.pt/docentes/aferreira/plas.html.

Forster, J.*, et al.* (2003). Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. *Genome Res* **13**(2): 244-253.

Freire, A. P.*, et al.* (2003). Anti-glycation defences in yeast. *Biochem Soc Trans* **31**: 1409-1412.

Frickel, E. M.*, et al.* (2001). Yeast glyoxalase I is a monomeric enzyme with two active sites. *J Biol Chem* **276**(3): 1845-1849.

Fu, M. X.*, et al.* (1994). Glycation, glycoxidation, and cross-linking of collagen by glucose. Kinetics, mechanisms, and inhibition of late stages of the Maillard reaction. *Diabetes* **43**(5): 676-683.

Geertman, J. M.*, et al.* (2006). Engineering NADH metabolism in Saccharomyces cerevisiae: formate as an electron donor for glycerol production by anaerobic, glucose-limited chemostat cultures. *FEMS Yeast Res* **6**(8): 1193-1203.

Gomes, R.*, et al.* (2005a). Argpyrimidine, a methylglyoxal-derived advanced glycation end-product in familial amyloidotic polyneuropathy. *Biochem J* **385**(Pt 2): 339-345.

Gomes, R. A.*, et al.* (2005b). Protein glycation in Saccharomyces cerevisiae. Argpyrimidine formation and methylglyoxal catabolism. *FEBS J* **272**(17): 4521-4531.

Gomes, R. A.*, et al.* (2008). Protein glycation in vivo: functional and structural effects on yeast enolase. *Biochem J* **416**(3): 317-326.

Gomes, R. J. (2007). *Biochemical effects of protein glycation by methylglyoxal in Saccharomyces cerevisiae*. Tese de Doutoramento. Faculdade de Ciências da Universidade de Lisboa, Lisboa.

Guha, M. K.*, et al.* (1988). Diffusion-dependent rates for the hydrolysis reaction catalyzed by glyoxalase II from rat erythrocytes. *Biochemistry* **27**(24): 8818-8822.

Gutenkunst, R. N.*, et al.* (2007). Universally Sloppy Parameter Sensitivities in Systems Biology. *PLoS Computational Biology* **3**(10).

Haanstra, J. R.*, et al.* (2008). Compartmentation prevents a lethal turbo-explosion of glycolysis in trypanosomes. *Proc Natl Acad Sci U S A* **105**(46): 17718-17723.

Heinrich, R. and T. A. Rapoport (1973). Linear theory of enzymatic chains; its application for the analysis of the crossover theorem and of the glycolysis of human erythrocytes. *Acta Biol Med Ger* **31**(4): 479-494.

Heinrich, R. and T. A. Rapoport (1974). A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. *Eur J Biochem* **42**(1): 89-95.

Heinrich, R. and S. Schuster (1996). *The regulation of cellular systems*. Chapman & Hall, New York.

Heux, S.*, et al.* (2006). Cofactor engineering in Saccharomyces cerevisiae: Expression of a H2O-forming NADH oxidase and impact on redox metabolism. *Metab Eng* **8**(4): 303-314.

Hindmarsh, A. C. (1983). ODEPACK, a Systematized Collection of ODE Solvers. *Scientific Computing*. R. S. *e. a.* Stepleman. North-Holland, Amsterdam**:** 55-64.

Hindmarsh, A. C. and L. R. Petzold (1983). Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations. *SIAM J. Sci. Stat. Comput.* **4**: 136-148.

Hinkle, P. C. (2005). P/O ratios of mitochondrial oxidative phosphorylation. *Biochim Biophys Acta* **1706**(1-2): 1-11.

Hoek, J. B. and J. Rydstrom (1988). Physiological roles of nicotinamide nucleotide transhydrogenase. *Biochem J* **254**(1): 1-10.

Hopkins, A. L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol* **4**(11): 682-690.

Hopkins, F. G. and E. J. Morgan (1945). On the distribution of glyoxalase and glutathione. *Biochem J* **39**(4): 320-324.

Hopper, D. J. and R. A. Cooper (1971). The regulation of Escherichia coli methylglyoxal synthase; a new control site in glycolysis? *FEBS Lett* **13**(4): 213-216.

Hornberg, J. J*., et al.* (2007). Metabolic control analysis to identify optimal drug targets. *Prog Drug Res* **64**: 171, 173-189.

Hovatta, I*., et al.* (2005). Glyoxalase 1 and glutathione reductase 1 regulate anxiety in mice. *Nature* **438**(7068): 662-666.

Huttemann, M*., et al.* (2008). Regulation of oxidative phosphorylation, the mitochondrial membrane potential, and their role in human disease. *J Bioenerg Biomembr* **40**(5): 445-456.

Inoue, Y. and A. Kimura (1996). Identification of the structural gene for glyoxalase I from Saccharomyces cerevisiae. *J Biol Chem* **271**(42): 25958-25965.

Inoue, Y*., et al.* (1998). Expression of the glyoxalase I gene of Saccharomyces cerevisiae is regulated by high osmolarity glycerol mitogen-activated protein kinase pathway in osmotic stress response. *J Biol Chem* **273**(5): 2977-2983.

Irsch, T. and R. L. Krauth-Siegel (2004). Glyoxalase II of African trypanosomes is trypanothione-dependent. *J Biol Chem* **279**(21): 22209-22217.

Iyengar, R. and I. A. Rose (1981). Concentration of activated intermediates of the fructose-1,6-bisphosphate aldolase and triosephosphate isomerase reactions. *Biochemistry* **20**(5): 1223-1229.

Izaguirre, G*., et al.* (1998). Methylglyoxal as substrate and inhibitor of human aldehyde dehydrogenase: comparison of kinetic properties among the three isozymes. *Comp Biochem Physiol B Biochem Mol Biol* **119**(4): 747-754.

Jacquez, J. A. and T. Perry (1990). Parameter estimation:local identifiability of parameters. *American Journal of Physiology - Endocrinology and Metabolism* **258**: E727-E736.

Jencks, W. P. (1987). *Catalysis in chemistry and enzymology.* Dover, New York.

Jensen, M. T. (2003). Reducing the run-time complexity of multiobjective EAs: The NSGA-II and other algorithms. *Ieee Transactions on Evolutionary Computation* **7**(5): 503-515.

Jeong, H*., et al.* (2000). The large-scale organization of metabolic networks. *Nature* **407**(6804): 651-654.

Jerzykowski, T*., et al.* (1978). A re-evaluation of studies on the distribution of glyoxalases in animal and tumour tissues. *Int J Biochem* **9**(11): 853-860.

Johnson, C. H*., et al.* (2003). Characterization of an alternative oxidase activity of Histoplasma capsulatum. *Yeast* **20**(5): 381-388.

Jones, D. P. (2008). Radical-free biology of oxidative stress. *Am J Physiol Cell Physiol* **295**(4): C849-868.

Jones, M. B*., et al.* (2002). Proteomic analysis and identification of new biomarkers and therapeutic targets for invasive ovarian cancer. *Proteomics* **2**(1): 76-84.

Kacser, H. and J. A. Burns (1973). The control of flux. *Symp Soc Exp Biol* **27**: 65-104.

Kang, Y*., et al.* (1996). Effect of methylglyoxal on human leukaemia 60 cell growth: modification of DNA G1 growth arrest and induction of apoptosis. *Leuk Res* **20**(5): 397-405.

Kellum, M. W*., et al.* (1978). A convenient quantitative synthesis of methylglyoxal for glyoxalase I assays. *Anal Biochem* **85**(2): 586-590.

Kermack, W. O. and N. A. Matheson (1957). The effects of some analogues of glutathione on the glyoxalase system. *Biochem J* **65**(1): 48-58.

Kitano, H. (2007). A robustness-based approach to systems-oriented drug design. *Nat Rev Drug Discov* **6**(3): 202-210.

Koivusalo, M. and L. Uotila *Abstr. Commun. 7th Meet. Fed. Eur. Biochem. Soc.* **117**.

Koop, D. R. and J. P. Casazza (1985). Identification of ethanol-inducible P-450 isozyme 3a as the acetone and acetol monooxygenase of rabbit microsomes. *J Biol Chem* **260**(25): 13607-13612.

Krymkiewicz, N. (1973). Reactions of methylglyoxal with nucleic acids. *FEBS Lett* **29**(1): 51-54.

Kukkonen, S. and J. Lampinen (2004a). An extension of generalized differential evolution for multi-objective optimization with constraints. *Parallel Problem Solving from Nature - Ppsn Viii* **3242**: 752-761.

Kukkonen, S. and J. Lampinen (2004b). A differential evolution algorithm for constrained multi-objective optimization: Initial assessment. *Proceedings of the IASTED International Conference on Artificial Intelligence and Applications, Vols 1and 2*: 96-102, 910.

Kukkonen, S. and J. Lampinen (2005). GDE3: The third evolution step of generalized differential evolution. *2005 Ieee Congress on Evolutionary Computation, Proceedings* **1**: 443-450.

Kukkonen, S. and K. Deb (2006). A fast and effective method for pruning of non-dominated solutions in many-objective problems. *Parallel Problem Solving from Nature - Ppsn Ix, Proceedings* **4193**: 553-562.

Kukkonen, S. and J. Lampinen (2007). Performance assessment of generalized differential evolution 3 (GDE3) with a given set of problems. *2007 Ieee Congress on Evolutionary Computation, Vols 1-10, Proceedings*: 3593-3600

4780.

Kullback, S. and R. A. Leibler (1951). On information and sufficiency. *Annals of Mathematical Statistics* **22**: 79-86.

Kume, S*., et al.* (1995). Immunohistochemical and ultrastructural detection of advanced glycation end products in atherosclerotic lesions of human aorta with a novel specific monoclonal antibody. *Am J Pathol* **147**(3): 654-667.

Kung, H. T*., et al.* (1975). Finding Maxima of a Set of Vectors. *Journal of the Acm* **22**(4): 469-476.

Lages, N. F*., et al.* (2008). Metabolismo de aminoácidos. *Organização Molecular da Vida*. A. L. Quintas, A. Ponces Freire and M. J. Halpern. Lidel, Lisboa.

Landgraf, R*., et al.* (2007). Candidate genes of anxiety-related behavior in HAB/LAB rats and mice: focus on vasopressin and glyoxalase-I. *Neurosci Biobehav Rev* **31**(1): 89-102.

Larsson, C*., et al.* (1998). The importance of the glycerol 3-phosphate shuttle during aerobic growth of Saccharomyces cerevisiae. *Yeast* **14**(4): 347-357.

Lederer, M. O. and R. G. Klaiber (1999). Cross-linking of proteins by Maillard processes: characterization and detection of lysine-arginine cross-links derived from glyoxal and methylglyoxal. *Bioorg Med Chem* **7**(11): 2499-2507.

Lee, A. T*., et al.* (1995). A role for DNA mutations in diabetes-associated teratogenesis in transgenic embryos. *Diabetes* **44**(1): 20-24.

Lee, C*., et al.* (1998). Oxidation-reduction properties of methylglyoxal-modified protein in relation to free radical generation. *J Biol Chem* **273**(39): 25272-25278.

Lo, T. W*., et al.* (1994). Binding and modification of proteins by methylglyoxal under physiological conditions. A kinetic and mechanistic study with N alpha-acetylarginine, N alpha-acetylcysteine, and N alpha-acetyllysine, and bovine serum albumin. *J Biol Chem* **269**(51): 32299-32305.

Lohmann, K. (1932). A study of the enzymatic transformation of synthetic methylglyoxal to lactic acid. *Biochem Z* **254**: 332-354.

Luo, B*., et al.* (2007). Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography-mass spectrometry. *J Chromatogr A* **1147**(2): 153-164.

Luthra, M. and D. Balasubramanian (1993). Nonenzymatic glycation alters protein structure and stability. A study of two eye lens crystallins. *J Biol Chem* **268**(24): 18119-18127.

Luttik, M. A*., et al.* (1998). The Saccharomyces cerevisiae NDE1 and NDE2 genes encode separate mitochondrial NADH dehydrogenases catalyzing the oxidation of cytosolic NADH. *J Biol Chem* **273**(38): 24529-24534.

Lyles, G. A. and J. Chalmers (1992). The metabolism of aminoacetone to methylglyoxal by semicarbazide-sensitive amine oxidase in human umbilical artery. *Biochem Pharmacol* **43**(7): 1409-1414.

Maeta, K*., et al.* (2004). Activity of the Yap1 transcription factor in Saccharomyces cerevisiae is modulated by methylglyoxal, a metabolite derived from glycolysis. *Mol Cell Biol* **24**(19): 8753-8764.

Maeta, K*., et al.* (2005). Methylglyoxal, a metabolite derived from glycolysis, functions as a signal initiator of the high osmolarity glycerol-mitogen-activated protein kinase cascade and calcineurin/Crz1-mediated pathway in Saccharomyces cerevisiae. *J Biol Chem* **280**(1): 253-260.

Maiti, M. K*., et al.* (1997). Molecular characterization of glyoxalase II from Arabidopsis thaliana. *Plant Mol Biol* **35**(4): 471-481.

Mannervik, B*., et al.* (1972). Partial purification and characterization of glyoxalase I from porcine erythrocytes. *Eur J Biochem* **29**: 276-281.

Mannervik, B*., et al.* (1973). The steady-state kinetics of glyoxalase I from porcine erythrocytes. Evidence for a random-pathway mechanism involving one- and two-substrate branches. *Eur J Biochem* **37**(2): 270-281.

Mannervik, B. (1974). Glyoxalase I. Kinetic Mechanism and Molecular Properties. *Glutathione. Proceedings of the 16th Conference of the German Society of Biological Chemistry, Tübingen, March 1973*. L. Flohé, H. C. Benöhr, H. Sies, H. D. Waller and A. Wendel. Georg Thieme Publishers, Stuttgart.

Mannervik, B*., et al.* (1974). Random pathway mechanism involving parallel one- and two-substrate branches for glyoxalase I from yeast. *J Biol Chem* **249**(3): 901-903.

Mannervik, B. (2008). Molecular enzymology of the glyoxalase system. *Drug Metabol Drug Interact* **23**(1-2): 13-27.

Marmstal, E*., et al.* (1979). Comparison of glyoxalase I purified from yeast (Saccharomyces cerevisiae) with the enzyme from mammalian sources. *Biochem J* **183**(1): 23-30.

Marres, C. A*., et al.* (1991). Isolation and inactivation of the nuclear gene encoding the rotenone-insensitive internal NADH: ubiquinone oxidoreductase of mitochondria from Saccharomyces cerevisiae. *Eur J Biochem* **195**(3): 857-862.

Martins, A. M*., et al.* (1999). Glyoxalase II in Saccharomyces cerevisiae: in situ kinetics using the 5,5'-dithiobis(2-nitrobenzoic acid) assay. *Arch Biochem Biophys* **366**(1): 15-20.

Martins, A. M*., et al.* (2001a). In situ analysis of methylglyoxal metabolism in Saccharomyces cerevisiae. *FEBS Lett* **499**(1-2): 41-44.

Martins, A. M*., et al.* (2001b). In situ kinetic analysis of glyoxalase I and glyoxalase II in Saccharomyces cerevisiae. *Eur J Biochem* **268**(14): 3930-3936.

McLellan, A. C. and P. J. Thornalley (1992). Synthesis and chromatography of 1,2-diamino-4,5-dimethoxybenzene, . *Eur J Med Chem* **20**: 393-402.

McLellan, A. C*., et al.* (1994). Glyoxalase system in clinical diabetes mellitus and correlation with diabetic complications. *Clin Sci (Lond)* **87**(1): 21-29.

Meister, A. and M. E. Anderson (1983). Glutathione. *Annu Rev Biochem* **52**: 711-760.

Meister, A. (1994). Glutathione-ascorbic acid antioxidant system in animals. *J Biol Chem* **269**(13): 9397-9400.

Meyerhof, O. and K. Lohmann (1934). Über die enzymatische Gleichgewichtsreaktion zwischen Hexosediphosphorsäure und Dioxyacetonphorsäure. *Biochemische Zeitschrift* **271**: 89-110.

Migliore, L.*, et al.* (1990). Genotoxicity of methylglyoxal: cytogenetic damage in human lymphocytes in vitro and in intestinal cells of mice. *Carcinogenesis* **11**(9): 1503-1507.

Miyata, T.*, et al.* (1993). beta 2-Microglobulin modified with advanced glycation end products is a major component of hemodialysis-associated amyloidosis. *J Clin Invest* **92**(3): 1243-1252.

Miyata, T.*, et al.* (1994a). Involvement of beta 2-microglobulin modified with advanced glycation end products in the pathogenesis of hemodialysis-associated amyloidosis. Induction of human monocyte chemotaxis and macrophage secretion of tumor necrosis factor-alpha and interleukin-1. *J Clin Invest* **93**(2): 521-528.

Miyata, T.*, et al.* (1994b). Glycation of human beta 2-microglobulin in patients with hemodialysis-associated amyloidosis: identification of the glycated sites. *Biochemistry* **33**(40): 12215-12221.

Moles, C. G.*, et al.* (2003). Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res* **13**(11): 2467-2474.

Monder, C. (1967). Alpha-keto aldehyde dehydrogenase, an enzyme that catalyzes the enzymic oxidation of methylglyoxal to pyruvate. *J Biol Chem* **242**(20): 4603-4609.

Muller, S.*, et al.* (2003). Thiol-based redox metabolism of protozoan parasites. *Trends Parasitol* **19**(7): 320-328.

Munch, G.*, et al.* (2000). Crosslinking of alpha-synuclein by advanced glycation endproducts--an early pathophysiological step in Lewy body formation? *J Chem Neuroanat* **20**(3-4): 253-257.

Mustacchi, R.*, et al.* (2006). Yeast systems biology to unravel the network of life. *Yeast* **23**(3): 227-238.

Nagaraj, R. H.*, et al.* (1996). Protein cross-linking by the Maillard reaction. Isolation, characterization, and in vivo detection of a lysine-lysine cross-link derived from methylglyoxal. *J Biol Chem* **271**(32): 19338-19345.

Nelder, J. A. and R. Mead (1965). A Simplex-Method for Function Minimization. *Computer Journal* **7**(4): 308-313.

Nelson, D. L. and M. M. Cox (2005). *Lehninger Principles of Biochemistry*. W. H. Freeman and Company, New York.

Neuberg, C. (1913). Über die Zerstörung von Milchsaurealdehyd und Methylglyoxal durch tierische Organe. *Biochem Zeitschr* **49**: 502-506.

Neuberg, C. and M. Kobel (1934). Present Status of the Problem of Sugar Fermentation. *J Bacteriol* **28**(5): 461-471.

Nielsen, J. (2003a). It is all about metabolic fluxes. *J Bacteriol* **185**(24): 7031-7035.

Nielsen, J. (2003b). *Bioreaction Engineering Principles*. Plenum Publishers, New York.

Nielsen, J.*, et al.* (2003). *Bioreaction Engineering Principles*. Kluwer Academic/Plenum Publishers, New York, NY.

Nissen, T. L.*, et al.* (1997). Flux distributions in anaerobic, glucose-limited continuous cultures of Saccharomyces cerevisiae. *Microbiology* **143 (Pt 1)**: 203-218.

Nissen, T. L.*, et al.* (2001). Expression of a cytoplasmic transhydrogenase in Saccharomyces cerevisiae results in formation of 2-oxoglutarate due to depletion of the NADPH pool. *Yeast* **18**(1): 19-32.

Nocedal, J. and S. J. Wright (1999). *Numerical Optimization*. Springer-Verlag, New York.

Nyhlin, N*., et al.* (2000). Advanced glycation end product in familial amyloidotic polyneuropathy (FAP). *J Intern Med* **247**(4): 485-492.

Ohmori, S*., et al.* (1987). Determination of methylglyoxal as 2-methylquinoxaline by high-performance liquid chromatography and its application to biological samples. *J Chromatogr* **414**(1): 149-155.

Outten, C. E. and V. C. Culotta (2003). A novel NADH kinase is the mitochondrial source of NADPH in Saccharomyces cerevisiae. *EMBO J* **22**(9): 2015-2024.

Overkamp, K. M*., et al.* (2000). In vivo analysis of the mechanisms for oxidation of cytosolic NADH by Saccharomyces cerevisiae mitochondria. *J Bacteriol* **182**(10): 2823-2830.

Oya, T*., et al.* (1999). Methylglyoxal modification of protein. Chemical and immunochemical characterization of methylglyoxal-arginine adducts. *J Biol Chem* **274**(26): 18492-18502.

Padayatti, P. S*., et al.* (2001). Argpyrimidine, a blue fluorophore in human lens proteins: high levels in brunescent cataractous lenses. *Invest Ophthalmol Vis Sci* **42**(6): 1299-1304.

Panagiotou, G*., et al.* (2009). Overexpression of a novel endogenous NADH kinase in Aspergillus nidulans enhances growth. *Metab Eng* **11**(1): 31-39.

Pareto, V. (1896). *Cours d'Economie Politique*. F. Rouge, Lausanne.

Pischetsrieder, M*., et al.* (1999). N(2)-(1-Carboxyethyl)deoxyguanosine, a nonenzymatic glycation adduct of DNA, induces single-strand breaks and increases mutation frequencies. *Biochem Biophys Res Commun* **264**(2): 544-549.

Polisetty, P. K*., et al.* (2006). Identification of metabolic system parameters using global optimization methods. *Theor Biol Med Model* **3**: 4.

Pollak, N*., et al.* (2007). The power to reduce: pyridine nucleotides--small molecules with a multitude of functions. *Biochem J* **402**(2): 205-218.

Ponces Freire, A*., et al.* (2003). Anti-glycation defences in yeast. *Biochem Soc Trans* **31**(Pt 6): 1409-1412.

Price, K. V. (1999). An introduction to differential evolution. *New Ideas in Optimization*. D. Corne, M. Dorigo and F. Glover. McGraw-Hill, London**:** 79-108.

Principato, G. B*., et al.* (1982). Glyoxalases and glutathione reductase activity changes in chicken liver during embryo development and after hatching. *Acta Embryol Morphol Exp* **3**(3): 173-179.

Principato, G. B*., et al.* (1985). Purification and characterization of S-2-hydroxyacylglutathione hydrolase (glyoxalase II) from human brain. *IRCS Med Sci* **13**: 952-953.

Raabe, H. M*., et al.* (1996). Biochemical alterations in collagen IV induced by in vitro glycation. *Biochem J* **319 (Pt 3)**: 699-704.

Racker, E. (1951). The mechanism of action of glyoxalase. *J Biol Chem* **190**(2): 685-696.

Rahman, A*., et al.* (1990). Formation of strand breaks and interstrand cross-links in DNA by methylglyoxal. *J Biochem Toxicol* **5**(3): 161-166.

Ray, M. and S. Ray (1984). Purification and partial characterization of a methylglyoxal reductase from goat liver. *Biochim Biophys Acta* **802**(1): 119-127.

Reder, C. (1988). Metabolic control theory: a structural approach. *J Theor Biol* **135**(2): 175-201.

Richard, J. P. (1985). Reaction of triosephosphate isomerase with L-glyceraldehyde 3-phosphate and triose 1,2-enediol 3-phosphate. *Biochemistry* **24**(4): 949-953.

Richard, J. P. (1991). Kinetic parameters for the elimination reaction catalyzed by triosephosphate isomerase and an estimation of the reaction's physiological significance. *Biochemistry* **30**(18): 4581-4585.

Richard, J. P. (1993a). Mechanism for the formation of methylglyoxal from triosephosphates. *Biochem Soc Trans* **21**(2): 549-553.

Richard, J. P. (1993b). Mechanism for the formation of methylglyoxal from triosephosphates. *Biochem Soc Trans* **21**(2): 549-553.

Ridderstrom, M*., et al.* (1996). Molecular cloning, heterologous expression, and characterization of human glyoxalase II. *J Biol Chem* **271**(1): 319-323.

Ridderstrom, M*., et al.* (1998). Involvement of an active-site Zn2+ ligand in the catalytic mechanism of human glyoxalase I. *J Biol Chem* **273**(34): 21623-21628.

Rigoulet, M*., et al.* (2004). Organization and regulation of the cytosolic NADH metabolism in the yeast Saccharomyces cerevisiae. *Mol Cell Biochem* **256-257**(1-2): 73-81.

Rodriguez-Fernandez, M*., et al.* (2006). Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. *BMC Bioinformatics* **7**: 483.

Sakamoto, H*., et al.* (2000). Glyoxalase I is involved in resistance of human leukemia cells to antitumor agent-induced apoptosis. *Blood* **95**(10): 3214-3218.

Sams-Dodd, F. (2005). Target-based drug discovery: is something wrong? *Drug Discov Today* **10**(2): 139-147.

Sauer, U*., et al.* (2004). The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of Escherichia coli. *J Biol Chem* **279**(8): 6613-6619.

Savageau, M. A. (1969a). Biochemical systems analysis. II. The steady-state solutions for an n-pool system using a power-law approximation. *J Theor Biol* **25**(3): 370-379.

Savageau, M. A. (1969b). Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions. *J Theor Biol* **25**(3): 365-369.

Savageau, M. A. (1970). Biochemical systems analysis. 3. Dynamic solutions using a power-law approximation. *J Theor Biol* **26**(2): 215-226.

Savageau, M. A. (1971). Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature* **229**(5286): 542-544.

savageau, M. A. (1972). The behaviour of intact biochemical control systems. *Current Topics in Cell Regulation* **6**: 63-130.

Savageau, M. A. (1974). Optimal design of feedback control by inhibition. *J Mol Evol* **4**(2): 139-156.

Savageau, M. A. (1976). *Biochemical Systems Analysis: a Study of Function and Design in Molecular Biology*. Addison-Weasley, Reading, Mass.

Schneider, M*., et al.* (2006). Detection of DNA-bound advanced glycation end-products by immunoaffinity chromatography coupled to HPLC-diode array detection. *Mol Nutr Food Res* **50**(4-5): 424-429.

Scire, A*., et al.* (2000). Specific interaction of cytosolic and mitochondrial glyoxalase II with acidic phospholipids in form of liposomes results in the inhibition of the cytosolic enzyme only. *Proteins* **41**(1): 33-39.

Seidel, W. and M. Pischetsrieder (1998). Immunochemical detection of N2-[1-(1-carboxy)ethyl]guanosine, an advanced glycation end product formed by the reaction of DNA and reducing sugars or L-ascorbic acid in vitro. *Biochim Biophys Acta* **1425**(3): 478-484.

Seidler, N. W. and I. Seibel (2000). Glycation of aspartate aminotransferase and conformational flexibility. *Biochem Biophys Res Commun* **277**(1): 47-50.

Sell, D. R. and V. M. Monnier (1989). Structure elucidation of a senescence cross-link from human extracellular matrix. Implication of pentoses in the aging process. *J Biol Chem* **264**(36): 21597-21602.

Sellin, S. and B. Mannervik (1983). Reversal of the reaction catalyzed by glyoxalase I. Calculation of the equilibrium constant for the enzymatic reaction. *J Biol Chem* **258**(14): 8872-8875.

Seo, D.-I*., et al.* (2004). New Epistasis Measures for Detecting Independently Optimizable Partitions of Variables. *Genetic and Evolutionary Computation Conference - GECCO 2004*. Berlin/Heidelberg, Springer. **3103:** 150-161.

Shapiro, R. and J. Hachmann (1966). The reaction of guanine derivatives with 1,2-dicarbonyl compounds. *Biochemistry* **5**(9): 2799-2807.

Shapiro, R.*, et al.* (1969). On the reaction of guanine with glyoxal, pyruvaldehyde, and kethoxal, and the structure of the acylguanines. A new synthesis of N2-alkylguanines. *Biochemistry* **8**(1): 238-245.

Shi, F.*, et al.* (2005). Identification of ATP-NADH kinase isozymes and their contribution to supply of NADP(H) in Saccharomyces cerevisiae. *FEBS Journal* **272**(13): 3337-3349.

Shianna, K. V.*, et al.* (2006). Genomic characterization of POS5, the Saccharomyces cerevisiae mitochondrial NADH kinase. *Mitochondrion* **6**(2): 94-101.

Shinohara, M.*, et al.* (1998). Overexpression of glyoxalase-I in bovine endothelial cells inhibits intracellular advanced glycation endproduct formation and prevents hyperglycemia-induced increases in macromolecular endocytosis. *J Clin Invest* **101**(5): 1142-1147.

Silva, M. S.*, et al.* (2008). Catalysis and structural properties of Leishmania infantum glyoxalase II: trypanothione specificity and phylogeny. *Biochemistry* **47**(1): 195-204.

Small, W. C. and L. McAlister-Henn (1998). Identification of a cytosolically directed NADH dehydrogenase in mitochondria of Saccharomyces cerevisiae. *J Bacteriol* **180**(16): 4051-4055.

Sonnleitner, B. and O. Kappeli (1986). Growth of Saccharomyces cerevisiae is controlled by its limited respiratory capacity: Formulation and verification of a hypothesis. *Biotechnol Bioeng* **28**(6): 927-937.

Sousa, M. M.*, et al.* (2001). Familial amyloid polyneurophaty: receptor for advanced glycation end products-dependent triggering of neuronal inflammatory and apoptotic pathways. *Journal of Neuroscience* **21**(19): 7576-7586.

Sousa Silva, M.*, et al.* (2005a). Quantitative assessment of the glyoxalase pathway in *Leishmania infantum* as a therapeutic target by modelling and computer simulation. *FEBS Journal*: in press.

Sousa Silva, M.*, et al.* (2005b). Quantitative assessment of the glyoxalase pathway in *Leishmania infantum* as a therapeutic target by modelling and computer simulation. *FEBS J* **272**: 2388-2398.

Stephanopoulos, G. N.*, et al.* (1998). *Metabolic Engineering: Principles and Methodologies*. Academic Press, London.

Storn, R. and K. Price (1997). Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces. *JOURNAL OF GLOBAL OPTIMIZATION* **11**(4): 341-359.

Strand, M. K.*, et al.* (2003). POS5 gene of Saccharomyces cerevisiae encodes a mitochondrial NADH kinase required for stability of mitochondrial DNA. *Eukaryot Cell* **2**(4): 809-820.

Szent-Gyorgyi, A. (1965). Cell Division And Cancer. *Science* **149**: 34-37.

Talesa, V.*, et al.* (1988). Demonstration of glyoxalase II in rat liver mitochondria. Partial purification and occurrence in multiple forms. *Biochim Biophys Acta* **955**(1): 103-110.

Talesa, V.*, et al.* (1989). Isolation of glyoxalase II from two different compartments of rat liver mitochondria. Kinetic and immunochemical characterization of the enzymes. *Biochim Biophys Acta* **993**(1): 7-11.

Talesa, V.*, et al.* (1990). Isolation of glyoxalase II from bovine liver mitochondria. *Biochem Int* **20**(1): 53-58.

Tan, S. X.*, et al.* (2009). The thioredoxin-thioredoxin reductase system can function in vivo as an alternative system to reduce oxidized glutathione in Saccharomyces cerevisiae. *J Biol Chem*.

Thornalley, P. J. (1990). The glyoxalase system: new developments towards functional characterization of a metabolic pathway fundamental to biological life. *Biochem J* **269**(1): 1-11.

Thornalley, P. J. (1993). The glyoxalase system in health and disease. *Mol Aspects Med* **14**: 287-371.

Thornalley, P. J. (1996). Pharmacology of methylglyoxal: formation, modification of proteins and nucleic acids, and enzymatic detoxification--a role in pathogenesis and antiproliferative chemotherapy. *Gen Pharmacol* **27**(4): 565-573.

Thornalley, P. J. (2003). Glyoxalase I--structure, function and a critical role in the enzymatic defence against glycation. *Biochem Soc Trans* **31**(Pt 6): 1343-1348.

Thornalley, P. J. (2006). Unease on the role of glyoxalase 1 in high-anxiety-related behaviour. *Trends Mol Med* **12**(5): 195-199.

Thornalley, P. J. (2008). Protein and nucleotide damage by glyoxal and methylglyoxal in physiological systems--role in ageing and disease. *Drug Metabol Drug Interact* **23**(1-2): 125-150.

Tressel, T.*, et al.* (1986). Interaction between L-threonine dehydrogenase and aminoacetone synthetase and mechanism of aminoacetone production. *J Biol Chem* **261**(35): 16428-16437.

Trincao, J.*, et al.* (2006). Purification, crystallization and preliminary X-ray diffraction analysis of the glyoxalase II from Leishmania infantum. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**(Pt 8): 805-807.

Turk, Z.*, et al.* (2006). Elevated level of methylglyoxal during diabetic ketoacidosis and its recovery phase. *Diabetes Metab* **32**(2): 176-180.

Tyson, C. B.*, et al.* (1979). Dependency of size of Saccharomyces cerevisiae cells on growth rate. *J Bacteriol* **138**(1): 92-98.

Uotila, L. (1973). Purification and characterization of S-2-hydroxyacylglutathione hydrolase (glyoxalase II) from human liver. *Biochemistry* **12**(20): 3944-3951.

Van Herreweghe, F.*, et al.* (2002). Tumor necrosis factor-induced modulation of glyoxalase I activities through phosphorylation by PKA results in cell death and is accompanied by the formation of a specific methylglyoxal-derived AGE. *Proc Natl Acad Sci U S A* **99**(2): 949-954.

Vander Jagt, D. L.*, et al.* (1992). Reduction of trioses by NADPH-dependent aldo-keto reductases. Aldose reductase, methylglyoxal, and diabetic complications. *J Biol Chem* **267**(7): 4364-4369.

Vandijken, J. P. and W. A. Scheffers (1986). Redox Balances in the Metabolism of Sugars by Yeasts. *Fems Microbiology Reviews* **32**(3-4): 199-224.

Vemuri, G. N. and A. A. Aristidou (2005). Metabolic engineering in the -omics era: elucidating and modulating regulatory networks. *Microbiol Mol Biol Rev* **69**(2): 197-216.

Vemuri, G. N.*, et al.* (2007). Increasing NADH oxidation reduces overflow metabolism in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **104**(7): 2402-2407.

Ventura, O. N. and M. L. Cubas (1992). A semiempirical study of the reaction of the hemimercaptal of methylglyoxal and glutathione at the active center of glyoxalase I. *International Journal of Quantum Chemistry* **44**(5): 699-722.

Verduyn, C.*, et al.* (1992). Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* **8**(7): 501-517.

Vickers, T. J.*, et al.* (2004). A trypanothione-dependent glyoxalase I with a prokaryotic ancestry in Leishmania major. *Proc Natl Acad Sci U S A* **101**(36): 13186-13191.

Vitek, M. P.*, et al.* (1994). Advanced glycation end products contribute to amyloidosis in Alzheimer disease. *Proc Natl Acad Sci U S A* **91**(11): 4766-4770.

Vlassara, H.*, et al.* (1994). Pathogenic effects of advanced glycosylation: biochemical, biologic, and clinical implications for diabetes and aging. *Lab Invest* **70**(2): 138-151.

Voit, E. O. (2000). *Computational Analysis of Biochemical Systems - A Practical Guide for Biochemists and Molecular Biologists*. Cambridge University Press, Cambridge.

von Jagow, G. and M. Klingenberg (1970). Pathways of hydrogen in mitochondria of Saccharomyces carlsbergensis. *Eur J Biochem* **12**(3): 583-592.

Wattanachaisaereekul, S*., et al.* (2008). Production of the polyketide 6-MSA in yeast engineered for increased malonyl-CoA supply. *Metab Eng* **10**(5): 246-254.

Webb, M. R*., et al.* (1977). Phosphorus-31 nuclear magnetic resonance of dihydroxyacetone phosphate in the presence of triosephosphate isomerase. The question of nonproductive binding of the substrate hydrate. *Biochemistry* **16**(12): 2738-2741.

Wendler, A*., et al.* (2009). Glyoxalase II does not support methylglyoxal detoxification but serves as a general trypanothione thioesterase in African trypanosomes. *Mol Biochem Parasitol* **163**(1): 19-27.

Westwood, M. E. and P. J. Thornalley (1997). *Glycation and Advanced Glycation Endproducts*. Springer-Verlag, Heidelberg.

Witte, A. B*., et al.* (2005). Inhibition of thioredoxin reductase but not of glutathione reductase by the major classes of alkylating and platinum-containing anticancer compounds. *Free Radic Biol Med* **39**(5): 696-703.

Xu, Y. and X. Chen (2006). Glyoxalase II, a detoxifying enzyme of glycolysis byproduct methylglyoxal and a target of p63 and p73, is a pro-survival factor of the p53 family. *J Biol Chem* **281**(36): 26702-26713.

Yan, S. D*., et al.* (1994). Glycated tau protein in Alzheimer disease: a mechanism for induction of oxidant stress. *Proc Natl Acad Sci U S A* **91**(16): 7787-7791.

Yang, K*., et al.* (2008). Finding multiple target optimal intervention in disease-related molecular network. *Mol Syst Biol* **4**: 228.

Yim, H. S*., et al.* (1995). Free radicals generated during the glycation reaction of amino acids by methylglyoxal. A model study of protein-cross-linked free radicals. *J Biol Chem* **270**(47): 28228-28233.

Yu, J*., et al.* (2001). Characterization of the Streptococcus pneumoniae NADH oxidase that is required for infection. *Microbiology* **147**(Pt 2): 431-438.