Journal of Biotechnology 161 (2012) 1–13

Contents lists available at SciVerse ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Genome-scale metabolic flux analysis of *Streptomyces lividans* growing on a complex medium

Pieter-Jan D'Huys[a], Ivan Lule[a], Dominique Vercammen[a], Jozef Anné[b], Jan F. Van Impe[a], Kristel Bernaerts[a,*]

[a] *Chemical and Biochemical Process Technology and Control Section, Department of Chemical Engineering, Katholieke Universiteit Leuven, W. de Croylaan 46, B-3001 Leuven, Belgium*
[b] *Bacteriology Section, Department of Microbiology and Immunology, Katholieke Universiteit Leuven, Minderbroederstraat 10, B-3001 Leuven, Belgium*

ARTICLE INFO

Article history:
Received 16 November 2011
Received in revised form 9 April 2012
Accepted 23 April 2012
Available online 26 May 2012

Keywords:
Streptomyces lividans
Constraint-based metabolic modeling
Genome-scale metabolic model
Flux balance analysis
Flux variability analysis
Complex medium
Heterologous protein

ABSTRACT

Constraint-based metabolic modeling comprises various excellent tools to assess experimentally observed phenotypic behavior of micro-organisms in terms of intracellular metabolic fluxes. In combination with genome-scale metabolic networks, micro-organisms can be investigated in much more detail and under more complex environmental conditions. Although complex media are ubiquitously applied in industrial fermentations and are often a prerequisite for high protein secretion yields, such multi-component conditions are seldom investigated using genome-scale flux analysis.

In this paper, a systematic and integrative approach is presented to determine metabolic fluxes in *Streptomyces lividans* TK24 grown on a nutritious and complex medium. Genome-scale flux balance analysis and randomized sampling of the solution space are combined to extract maximum information from exometabolome profiles.

It is shown that biomass maximization cannot predict the observed metabolite production pattern as such. Although this cellular objective commonly applies to batch fermentation data, both input and output constraints are required to reproduce the measured biomass production rate. Rich media hence not necessarily lead to maximum biomass growth.

To eventually identify a unique intracellular flux vector, a hierarchical optimization of cellular objectives is adopted. Out of various tested secondary objectives, maximization of the ATP yield per flux unit returns the closest agreement with the maximum frequency in flux histograms. This unique flux estimation is hence considered as a reasonable approximation for the biological fluxes.

Flux maps for different growth phases show no active oxidative part of the pentose phosphate pathway, but NADPH generation in the TCA cycle and NADPH transdehydrogenase activity are most important in fulfilling the NADPH balance. Amino acids contribute to biomass growth by augmenting the pool of available amino acids and by boosting the TCA cycle, particularly when using glutamate and aspartate. Depletion of glutamate and aspartate causes a distinct shift in fluxes of the central carbon and nitrogen metabolism.

In the current work, hurdles encountered in flux analysis at a genome-scale level are addressed using hierarchical flux balance analysis and uniform sampling of the constrained solution space. This general framework can now be adopted in further studies of *S. lividans*, e.g., as a host for heterologous protein production.

1. Introduction

Constraint-based modeling is nowadays a popular tool to quantify metabolic fluxes in cells and assess the impact of environmental conditions and genetic alterations on the cell's phenotype. In the early days, most applications involved relatively small metabolic reaction networks, including central metabolic pathways, lumped biosynthetic reactions for macromolecules and biomass formation, of well-known organisms. Progress in genome sequencing has pushed constraint-based metabolic modeling to the level of genome-scale metabolic reaction networks and applications treat various kind of cells ranging from microbial cell factories to human oral pathogens (e.g., Edwards and Palsson, 2000; Borodina et al., 2005a; Selvarasu et al., 2009; Mazumdar et al., 2009).

Constraint-based metabolic modeling refers to the quantitative analysis of metabolic fluxes starting from a metabolic steady-state

* Corresponding author. Tel.: +32 16 32 2678; fax: +32 16 32 29 91.
E-mail addresses: jozef.anne@rega.kuleuven.be (J. Anné), jan.vanimpe@cit.kuleuven.be (J.F. Van Impe), kristel.bernaerts@cit.kuleuven.be (K. Bernaerts).

0168-1656/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbiotec.2012.04.010

solution space constrained by the metabolic reaction stoichiometry and the steady-state assumption for intracellular metabolites (e.g., Price et al., 2004; Llaneras and Picó, 2008). Steady-state mass balances of intracellular metabolites define a solution space which is often further bounded by additional equality and inequality constraints reflecting reaction (ir)reversibility, enzyme capability and experimentally observed exchange fluxes. For small networks, sufficient constraints are usually available to calculate a unique set of fluxes. However, for large-scale metabolic networks, the problem typically remains underdetermined. A steady-state solution space, which encompasses all possible phenotypes of the cell under the governing environmental circumstances, is obtained and one has to seek for a unique solution within this space.

Several methods exist to explore this constrained solution space and to seek for (unique) solutions (e.g., Llaneras and Picó, 2008). Flux spectrum analysis (FSA) merely determines the flux interval for each flux only given the above-defined constraints (Llaneras and Picó, 2007). Flux balance analysis (FBA) envisages finding a unique subset of fluxes satisfying some optimality criterion (e.g., Price et al., 2004). By pursuing this optimal behavior, one hopes to remove the systems underdeterminacy and to find a unique flux map. In case of highly dimensional solution spaces, as for genome-scale networks, often alternate optima are found instead of a unique flux vector. Flux variability analysis (FVA) is in an easy way to uncover the existence of these alternate optima (Mahadevan and Schilling, 2003). On the other hand, Monte Carlo sampling (MCS) can be addressed as an unbiased method to characterize the solution space by means of uniform random sampling (Wiback et al., 2004).

This paper focuses on flux analysis for *Streptomyces lividans* TK24 as an attractive host for heterologous protein secretion. Streptomycetes are well-recognized producers of a plethora of antibiotics and other bioactive secondary metabolites. Its large secretory capacity have stimulated researchers to explore its application as a recombinant protein secretion platform. In this effort, *S. lividans* has a proven record of many large and small proteins with industrial and therapeutic applications (see Vrancken and Anné, 2009, for a review), yet protein yields are often unsatisfactory. As reviewed in Anné et al. (2011), several approaches can be followed to augment protein yields, e.g., proper promotor and signal peptide selection, reduction of protease activity, etc.; however, efforts to look into the metabolism have not been pursued in a systematic way.

In view of optimizing heterologous protein production in *S. lividans*, a better understanding of the effects of C- and N-substrates on the metabolism is necessary. Thus far, constraint-based modeling techniques have mainly been applied to micro-organisms that grow on minimal media containing a single carbon source. However, complex media are used in many industrial bioproduction processes. Multiple inputs for the metabolic network complicate constraint-based modeling approaches considerably. Moreover, one may expect quite different metabolic behavior in such a complex environment compared with that in minimal media (e.g., Teusink et al., 2006; Selvarasu et al., 2009).

In this work, genome-scale constraint-based metabolic modeling is exploited to analyze exometabolome data observed from batch fermentations of *S. lividans* TK24 growing on multiple C- and N-substrates. A systematic approach to obtain a plausible unique flux vector is proposed. Experimental data are reported in D'Huys et al. (2011) whom present an exhaustive phenomenological discussion of extracellular observations. Batch growth of *S. lividans* TK24 on glucose and ammonium supplemented with a mixture of amino acids is investigated. Almost all potential substrates are consumed, with a preference for glucose, ammonium, glutamate and aspartate. Secretion of several organic acids and alanine is observed. Now, these experimental data are confronted with constraint-based metabolic modeling. The plethora of substrates taken up motivates the use of a genome-scale metabolic network model, which is setup. First, the consistency of the experimental data with the stoichiometric network model and the validity of the optimal growth principle (reflected by the maximization of biomass as a cellular objective) is evaluated under several constraints. Next, the remaining metabolic solution space (obeying maximum biomass formation and specified boundaries on exchange rates) is narrowed down to a unique solution by sequentially optimizing cellular objectives. Several published objective functions are tested and solutions are compared with flux distributions established by Monte Carlo sampling of the solution space. The flux distribution that has the closest agreement with the maximum frequency in the Monte Carlo flux histograms is considered as a reasonable approximation for the 'real' biological flux distribution. To conclude, this flux distribution is discussed in further detail.

2. Material and methods

2.1. Bacterial strain, media and experiments

Full details are available in D'Huys et al. (2011), but to facilitate the reading of the current paper essential elements are summarized in this section.

Streptomyces lividans TK24 was obtained from the John Innes Centre (Norwich, UK). The stock culture was maintained as a homogenized mycelium suspension at −80 °C in 20% (v/v) glycerol. Pre-cultures were obtained by two subsequent culturing steps in phage medium (Korn-Wendish et al., 1978) (per L 10 g glucose, 5 g tryptone, 5 g yeast extract, 5 g LabM, 0.74 g $CaCl_2{\cdot}2H_2O$, 0.5 g $MgSO_4{\cdot}7H_2O$, and pH 7.2) at 30 °C and 250 rpm. Cells from 250 mL preculture were harvested by centrifugation (10 min at 3200 × *g*). The pellets were resuspended in a small aliquot of reactor medium, i.e., slightly modified NMMP medium consisting of (per L) 10 g glucose, 3 g $(NH_4)_2SO_4$, 5 g or 15 g Bacto™ Casamino acids, 0.6 g $MgSO_4{\cdot}7H_2O$, 2.7 g NaH_2PO_4, 3.92 g K_2HPO_4 and 1 mL spore elements solution (containing per L 1 g $ZnSO_4{\cdot}7H_2O$, 1 g $FeSO_4{\cdot}7H_2O$, 1 g $MnCl_2{\cdot}4H_2O$ and 1 g anhydrous $CaCl_2$) (Kieser et al., 2000).

Batch growth was monitored in a bench-top bioreactor (BioFlo 3000, New Brunswick Scientific Inc.) filled with 3 L NMMP medium. Temperature was set at 30 °C, the rotation speed at 400 rpm, the aeration flow at 2 L/min and pH was kept constant at 7 (using 1N H_2SO_4 and 1N NaOH). Dissolved oxygen was not controlled but was never depleted during the experiments. Antifoam (antifoam Y-30 emulsion, Sigma) was occasionally added to avoid extensive foam formation. At regular time intervals, cell dry weight (DW) was determined in triplicate and the filtrate was used for analysis of the extracellular metabolites. Glucose and ammonium were enzymatically measured (Gluc-DH, Ecoline® S+from Diagnostic Systems, and ammonia testkit from Boehringer Mannheim/R-biopharm, respectively). After methylation, organic acids (pyruvic acid, lactic acid, α-ketoglutaric acid and succinic acid) were determined by GC-FID (Finnigan Trace GC Ultra) with a Stabilwax®-DA column (30 m × 0.32 mm ID, Restek). Amino acids were quantified with the EZ:faast™ GC–MS amino acid analysis kit (Phenomenex) using GC–MS (PerkinElmer Autosystem XL-Turbomass Gold GC–MS). Except for arginine and cysteine, all amino acids could be measured. Samples were analyzed in triplicate.

2.2. Constraint-based metabolic modeling

Core of constraint-based metabolic modeling is the general equation describing the steady-state mass balances for intracellular metabolites (under the assumption of negligible dilution effects due to biomass growth):

$$\mathbf{S}\cdot\mathbf{v}=\mathbf{0} \quad (1)$$

where the flux vector $\mathbf{v}$ ($n \times 1$) contains intracellular fluxes and exchange fluxes (expressed in mmol $gDW^{-1}\,h^{-1}$). The exchange fluxes (i.e., specific consumption or production rates) for the measured metabolites glucose, ammonium, amino acids and organic acids were determined in the exponential growth phase(s), in which a pseudo steady-state (and thus constant yield) was assumed. Details on the calculations are provided in Supplementary data S1 and D'Huys et al. (2011). $\mathbf{S}$ represents the stoichiometric matrix ($m \times n$) summarizing the reaction stoichiometry for m intracellular metabolites. Each element $\mathbf{S}_{ij}$ corresponds with the stoichiometric coefficient of metabolite i in reaction j. A negative value of $\mathbf{S}_{ij}$ indicates that the metabolite is consumed in reaction j, a positive value indicates that metabolite i is produced. In the current work, $\mathbf{S}$ stands for the stoichiometric matrix of the genome-scale metabolic network (*GSMN*) for *Streptomyces lividans* TK24. During the construction of this model, a number of assumptions and adjustments were made to the *S. coelicolor* A3(2) GSMN from Borodina et al. (2005a). Details of the construction proces can be found in Supplementary data S1. The final metabolic network model contains 702 reactions, 496 intracellular metabolites and 106 extracellular metabolites and a full version is available in Supplementary data S2.

Fluxes cannot vary unboundedly and are constrained by an upper and lower limit:

$$v_j^L \leq v_j \leq v_j^U \tag{2}$$

Since in vivo enzyme capabilities of intracellular reactions are not known, intracellular fluxes were arbitrarily bounded by $v^U = 10^5$ and $v^L = -10^5$ for reversible reactions, virtually putting no limit on these reaction fluxes. Reaction irreversibility was imposed by setting one boundary value equal to zero. Doing so, Eqs. (1) and (2) span a steady-state solution space encompassing possible flux distributions consistent with the network stoichiometry and reversibility constraints. Mathematical approaches used to analyze this steady-state solution space and to identify a unique flux set are summarized in the following paragraphs.

Flux balance analysis (FBA). Dealing with a GSMN, the degrees of freedom of the system is much larger than the number of measured exchange fluxes, such that a unique flux vector cannot be found through directly solving the set of linear equations for the unknown fluxes. FBA optimizes some cellular objective function ($Z(\mathbf{v})$) within the feasible steady-state solution space with the aim of finding a unique solution. Setting maximum biomass growth as a primary cellular objective, the following linear programming (LP) problem is obtained:

$$\begin{aligned} \max \quad & v_{biomass} \\ \text{subject to} \quad & \mathbf{S}\cdot\mathbf{v} = \mathbf{0} \\ & v_j^L \leq v_j \leq v_j^U \end{aligned}$$

To gradually reduce the feasible solution space, constraints derived from the experimental data were systematically introduced. Five scenarios were defined as explained in Table 1.

Measured exchange fluxes ($v_{j,m}$) were embedded as soft or hard constraints, i.e., as an inequality ($0 \leq v_j \leq v_{j,m}$) or equality constraint ($v_j = v_{j,m}$), respectively. Measured substrate uptake fluxes were imposed as upper boundaries. In case of multiple proton-mediated transport reactions that can be used, the total uptake flux (i.e., the sum of all uptake fluxes) was constrained by the measured exchange flux. Depending on the scenario, measured secretion fluxes were either unconstrained or introduced as equality constraints, enforcing secretion of the desired by-products. Secretion fluxes of non-detected metabolites were unconstrained or constrained to zero depending on the selected scenario. The oxygen uptake rate and the carbon dioxide production rate were unconstrained, and also phosphate and sulphate uptake, which are abundantly available in the medium, were not constrained. Because GC-FID did not allow discrimination between D- and L-lactate and since D- and L-lactate are completely interchangeable in the GSMN under the prevailing environmental conditions, D-lactate secretion was assumed to represent the experimentally measured lactate secretion. Alanine accumulation during the exponential phase was experimentally recorded and was imposed as D-alanine secretion because it is the only form in which alanine secretion can occur according to the used GSMN. Since extracellular D-alanine (DALAxt) can result from cell wall biosynthesis as well as proton-mediated transport, the overall secretory flux (being the sum of possible secretion fluxes) was constrained by the measured (global) exchange flux. Besides the above-mentioned constraints, a series of equality constraints have been determined when building the model (see Supplementary data S1). These equality constraints are the same for all scenarios. When taking into account experimental measurement errors, exchanges rates were bounded within the 68% confidence interval of the calculated exchange rates (i.e., the measured value ± the standard deviation).

Resulting LP problems were solved with the GNU Linear Programming Kit (GLPK) package Makhorin (using the *glpkmex* Matlab interface) and the IBM ILOG CPLEX optimizer (IMB Inc.). The glpk solver and the CPLEX solver both use a simplex method to solve LP probems. Both solvers not necessarily return the same flux distribution when the problem has alternate solutions. All computations were performed in Matlab R2009b (Version 7.9., Mathworks, MA, USA).

Flux variability analysis (FVA). Flux ranges for the optimum phenotype were computed using flux variability analysis (Mahadevan and Schilling, 2003). For each flux, a maximum and minimum flux value was determined while satisfying all defined constraints as well as the optimized objective value (Z_{opt}).

$$\begin{aligned} \max/\min \quad & v_i \\ \text{subject to} \quad & Z = Z_{opt} \\ & \mathbf{S}\cdot\mathbf{v} = \mathbf{0} \\ & v_j^L \leq v_j \leq v_j^U \end{aligned}$$

FVA ranges give an indication of the capabilities of the network, or in other words, unveils alternate routes which give the same maximum objective function. Reactions without flux variation or with small flux ranges are essential for achieving the considered phenotypic state, whereas more flexibility in their fluxes implies the presence of alternative pathways.

Hierarchical flux balance analysis. To retrieve a biologically meaningful unique solution out of the multiple solutions after optimization of biomass formation, a hierarchical FBA approach was used. This hierarchical FBA approach is based on the hypothesis that a cell can adhere to several optimality principles. Instead of using a multi-objective optimization approach (Sendín et al., 2009), cellular objectives are optimized sequentially until a unique solution is found. Selection of the objective functions is based on the principle of minimal effort which can be expressed in various ways, e.g., as total flux minimization or as minimization of the number of active reactions (Holzhütter, 2004; Schuetz et al., 2007; Murabito et al., 2009). Four objective functions were considered:

- **min** $\sum_j |v_j|$
  Minimization of the total sum of absolute fluxes reflects the minimization of the flux through each enzymatic reaction (Holzhütter, 2004).
- **min** $\sum_j v_j^2$

Table 1
Scenarios tested during flux balance analysis when maximizing biomass growth. Measured exchange fluxes (v_m) are gradually introduced as inequality or equality constraints (with GLC, NH_4: glucose, ammonium; AAs: amino acids; LAC, PYR, AKG, SUCC: lactacte, pyruvate, α-ketoglutarate, succinate; ALA: D-alanine). The number of fluxes to be computed as well as the sequentially added inequality constraints is indicated. In case of multiple exchange reaction, the total flux is constrained (*). Further details are given in the text.

Scenario	Constraints				# unknown fluxes	# added ineq. constraints
	$v_j \leq v_{j,m}$		$v_j = v_{j,m}$			
	GLC, NH4	AAs	LAC, PYR, AKG, SUCC	ALA		
A. Glucose-ammonium free output	X				614	2
B. Amino acids free output	X	X			614	14*
C. Organic acids constrained output	X	X	X		610	
D. Alanine constrained output	X	X	X	X	609*	
E. Experimentally measured constrained output (undetected secretion fluxes set to zero)	X	X	X	X	584	

Minimization of the quadratic sum of fluxes resembles the flux minimization approach but always identifies a unique global optimum (Schuetz et al., 2007).

- **min** *path*

  Opposed to flux minimization, the number of non-zero reactions is minimized coupling the minimal effort principle to the economical use of enzymes (Murabito et al., 2009).

- **max** $v_{ATP}/\sum_j v_j^2$

  Maximization of ATP yield per flux unit reflects that a cell tries to maximize ATP production while minimizing enzyme usage (Schuetz et al., 2007).

To solve the proposed FBA problems, some problems needed to be restated to render a simpler LP or mixed-integer LP programming problem, whereas some objective functions were quadratic functions and needed to be solved using quadratic programming. Details are given in Supplemental data S1.

Random sampling of the solution space. Uniform random sampling of a solution space is a non-optimization based unbiased technique to study the full range of achievable metabolic network states provided by the steady-state solution space. The sampling points correspond to candidate network states that are in agreement with the imposed constraints. When sampling the solution space sufficiently, histograms can be drawn for each flux. These flux distributions can be understood as a probability distributions for every reaction, from which the most likely flux value through a reaction (given the environmental constraints imposed on the organism) can be deduced (Thiele and Palsson, 2007; Schellenberger and Palsson, 2009).

Sampling of the steady-state solution space was done by means of a random walk algorithm, namely, the so-called artificial centering hit-and-run (ACHR), as described by Kaufman and Smith (1998), Thiele et al. (2005). Computations were performed using the COBRA toolbox (Becker et al., 2007) with GLPK as the LP solver in Matlab. To obtain a proper representation of the flux space, 100,000 randomly distributed points were computed with 1000 iterations between each point. Subsequently, histograms were drawn and the flux with the highest frequency was determined for each reaction. The difference with optimal flux vectors resulting from flux balance analysis was quantified by the L1-norm and the mean absolute difference.

3. Results and discussion

3.1. Exometabolome versus optimal biomass growth principle

Concentration profiles of extracellular metabolites of *S. lividans* TK24 grown in NMMP supplemented with casamino acids are the net result of the metabolic activity of *S. lividans* under the given substrate availability. Measured exchange fluxes (and corresponding standard deviations) are listed in Supplementary data S3. In NMMP supplemented with 5 g L^{-1} CAS, depletion of glutamate and aspartate causes a decrease in the growth rate. Hence, a distinction is made between growth phase 1 and 2 (denoted as GP1 and GP2). In NMMP supplemented with 15 g L^{-1} CAS, a single exponential growth phase (GP) is observed since depletion of glucose, glutamate and aspartate almost coincide in time.

To understand this observed macroscopic metabolic activity of *S. lividans* TK24, to investigate the contribution of the different substrates, and to couple exometabolome time profiles with the intracellular metabolism, techniques of constraint-based metabolic modeling are exploited. In a first stage, measured exchange fluxes are combined with a genome-scale metabolic reaction network model within a flux balance analysis approach. As a cellular objective maximization of biomass formation is adopted, since the criterion has proven to be well-suited for batch cultures (e.g., Edwards et al., 2001; Fong and Palsson, 2004). The complexity of the medium imposes the use of a genome-scale metabolic model which contains catabolic reactions for all available substrates. Systematically introducing measured constraints, five scenarios are distinguished: (A) the glucose-ammonium free output system, (B) the amino acids free output system, (C) the organic acids constrained output system, (D) the alanine constrained output system and (E) the experimentally measured constrained output system (see Table 1). All these systems are composed of the same stoichiometric matrix, but differ in the number of constraints imposed. Gradually building up the number of constraints gives additional insight but also avoids ending up with an unfeasible solution space when imposing all constraints at once.

In the glucose–ammonium free output system, the experimentally measured specific uptake rates are used as an upperbound for the uptake fluxes of glucose and ammonium. In the amino acids free output system, the specific uptake rates of the amino acids available for the cell in a certain growth phase are set as upperbounds on amino acid uptake in addition to the constraints of the first system. The organic acids constrained output system includes the specific production rates of the measured organic acids (pyruvate, α-ketoglutarate, lactate and succinate) as equality constraints while the rest of the possible output (secretion) fluxes are left unconstrained. Next, the alanine constrained output system includes the specific alanine production rate in addition to the organic acids constrained output system. Finally, the fully constrained output system includes all experimentally measured specific production rates (organic acids and alanine secretion) as fixed constraints in addition to the constraints of the amino acids free output system, while the remaining output fluxes are put to zero (except for CO_2). Hence, given a certain amount of nutrients, the free output results represent what the cell is stoichiometrically able to achieve, while the constrained output results represent what the cell is able to achieve if it is forced to secrete an amount of the available carbon and nitrogen as overflow products.

Table 2
Specific biomass growth rate (μ) as observed experimentally or as predicted through flux balance analysis with optimal biomass growth and various constraints on the exchange fluxes. Simulated scenarios are explained in Table 1. Complete solutions are listed in Supplementary data S4.

	5 g L^{-1} CAS		15 g L^{-1} CAS
	GP1	GP2	GP
Observed specific biomass growth rate (μ) [h^{-1}]	1.4010E−01	6.3280E−02	1.1290E−01
$\mu\pm$ StDev (68% confidence interval)	[1.3624E−01, 1.4424E−01]	[5.9169E−02, 6.7391E−02]	[1.1268E−01, 1.1312E−01]
A. Glucose ammonium free output	8.8002E−02	5.7704E−02	7.0846E−02
B. Amino acid free output	1.4847E−01	7.9618E−02	1.3604E−01
C. Organic acids constrained output	1.4613E−01	7.9618E−02	1.3604E−01
D. Alanine constrained output	1.4295E−01	5.8982E−02	1.0425E−01
E. Experimentally measured constrained output	1.4295E−01	5.8982E−02	1.0425E−01
F. Alanine constrained output including measurement uncertainty	1.5662E−01	6.8529E−02	1.4794E−01

Table 2 summarizes the predicted specific biomass growth rate applying the various constrained systems. In each growth phase, the maximum specific growth rate realized on glucose and ammonium uptake is clearly smaller than the measured value (Scenario A). Uptake of amino acids thus contributes to biomass synthesis but the addition of the observed amino acids uptake rates overestimates the biomass growth capacity (Scenario B). All substrates are taken up at their maximum rate, except for the ammonium uptake in GP1 which is smaller (0.7304 mmol gDW^{-1} h^{-1}) than experimentally measured (0.8138 mmol gDW^{-1} h^{-1}) but still in the 68% confidence region of the measured value. Except for CO_2 and a very small amount of D-alanine secretion (associated with cell wall synthesis), no other metabolites are secreted. Calculation of flux variability ranges returns flux ranges for exchange fluxes in GP2 and GP (see Fig. 1), but not in GP1. In the former case, alternate solutions exist which return the same biomass growth rate but different metabolite profiles (including organic acids, alcohols such as, e.g., pyruvate, succinate, ethanol, alpha-ketoglutarate, glyoxylate). Thus GP2 and GP encompass a secretion capability. Comparing measured secretory fluxes with the corresponding flux variability analysis (FVA) ranges shows that the measured organic acids are plausible by-products under the optimal growth hypothesis. Consequently, adopting the measured organic acid secretion rates as fixed constraints (organic acid output system) does not change the growth rate in GP2 nor in GP (Table 2). A FVA-range keeps existing for the other mentioned metabolites. In the case of GP1, no flexibility of the metabolic model remains under the specified constraints and objective function. Hence, maximation of biomass formation is not consistent with the observed exometabolome. Applying the organic acids output system (Scenario C), the predicted growth rate drops but remains still too high as compared with the experimental value. Imposing alanine secretion (Scenario D) further reduces the biomass growth rate in alle cases, which is not unexpected since both nitrogen and carbon are expended. The predicted growth rates now resemble the experimentally observed growth rates reasonably close. Finally, the application of all measured exchange fluxes as constraints and setting all undetected secretion fluxes equal to zero (Scenario E) further narrows down the feasible solution space to a space that now contains all possible flux distributions consistent with the experimentally observed phenotype.

Clearly, *S. lividans* is not growing optimally with respect to biomass formation under the presented medium conditions. Whereas carbon overflow still belongs to the feasible solution space of optimum growth in GP2 and GP, alanine secretion is not consistent with the maximal biomass hypothesis. Carbon overflow is a commonly observed response to high glucose intake and/or respiratory limitations. Since the oxygen uptake rate is not bounded in scenarios A to D, oxygen limitation is likely not explaining organic acid production in the current simulations. Constraining the oxygen uptake rate also not predicts the correct organic acid production pattern or rates (data not shown). In GP2 and GP, organic acid production can be explained by the metabolic reaction stoichiometry, whereas in GP1 intracellular enzyme kinetics (e.g., maximum enzyme capacity, allosteric interaction) are probably causing organic acid secretion. Alanine is not a commonly observed by-product during microbial growth. As mitigated in D'Huys et al. (2011), alanine seems to be secreted as an overflow product dealing with excess nitrogen and carbon in the cell.

An additional experiment with 5 g L^{-1} casmino acids but with a lowered ammonium concentration (i.e., 2 g L^{-1} instead of 3 g L^{-1} $(NH_4)_2SO_4$) shows that alanine and pyruvate secretion are apparently interlinked (see Fig. 2). At the highest ammonium concentration (left plot), alanine peaks to a much higher amount whereas the pyruvate level remains moderate. The opposite is observed in the experiment with less ammonium (right plot). Moreover, the difference in alanine concentration almost matches the difference in pyruvate level. Since alanine dehydrogenase (ADH) is the only way to produce alanine, this coupling is not unreasonable, namely, one mole of pyruvate is aminated to form 1 mole of alanine. ADH is a widespread enzyme among

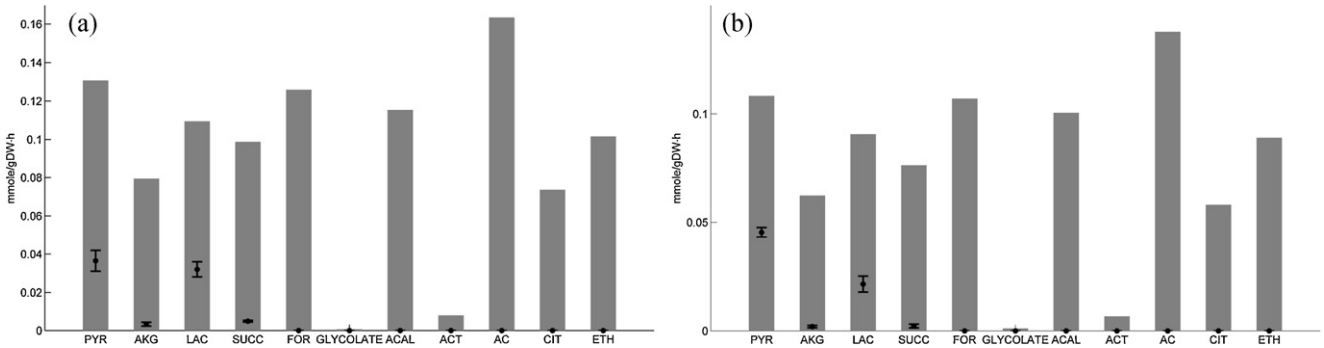

Fig. 1. Flux variability ranges of secreted metabolites versus measured secretion fluxes of organic acids. Scenario B, i.e., the *amino acid free output* system, is depicted. Results for GP2 (a) and GP (b) are shown.

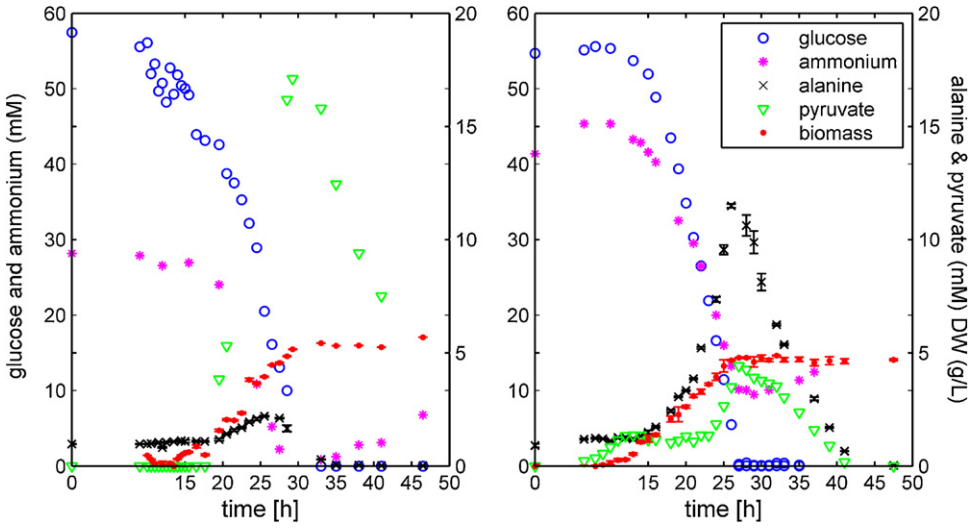

Fig. 2. Experimental data of *S. lividans* TK24 grown in NMMP containing a different initial concentration of ammonium. Left plot: $2\,g\,L^{-1}$ $(NH_4)_2SO_4$. Right plot: $3\,g\,L^{-1}$ $(NH_4)_2SO_4$.

streptomycetes but its role is not yet fully resolved. Brana et al. (1986), Novak et al. (1997) and Fisher (1989), for example, report its induction in the presence of high extracellular ammonium concentrations. Whereas these authors already reported that ADH does not play a direct role as an alternative pathway of ammonium assimilation and wondered whether the induction of the enzyme by high ammonium concentrations has other biological significance, current results point at a role in controlling the intracellular ammonium levels.

Thus far, measured exchanges fluxes are used as *exact* values, but these flux estimates are characterized by experimental error. Introducing bounds on the exchange fluxes according to the experimental error (i.e., $v_{j,m} - StDev_{v_m} \leq v_{j,m} \leq v_{j,m} + StDev_{v_m}$, with $StDev_{v_m}$ the standard deviation on the measured flux) renders a biomass growth rate which is too high as compared to the measured value (see Scenario F in Table 2). After optimization, all substrate uptake rates (except for ammonium) coincide with their upper boundary ($v_{j,m} + StDev_{v_m}$) while all secretion fluxes adopt the corresponding lower boundary value ($v_{j,m} - StDev_{v_m}$). This solution returns a higher biomass growth rate. Although measurements are prone to experimental errors, it is not possible to account for experimental errors within the adopted constraint-based modelling framework. The feasible solution space is extended such that irrealistic predictions on the biomass growth rate are obtained. Since the predicted biomass growth rate using the mean estimated exchange fluxes is very acceptable, Scenario E is retained for further analysis in the following sections.

3.2. Characterization of the optimal growth genome-scale metabolic steady-state solution space

Application of all measured exchange fluxes as constraints narrows down the feasible solution space (Scenario E), but a number of network indeterminacies remain and give rise to alternate solutions for the intracellular flux vector. In other words, the same objective function value – which is here the biomass growth rate – can be realized through different intracellular flux distributions. Multiple solutions may all represent biologically meaningful flux maps but only one flux set coincides with the actual phenotype whereas the other flux distributions may reflect silent phenotypes (Mahadevan and Schilling, 2003; Reed and Palsson, 2004).

To characterize the extent of the remaining metabolic solution space from Scenario E, flux variability analysis and random uniform Monte Carlo sampling are exploited. Although both approaches determine flux ranges, FVA and Monte Carlo sampling are complementary methods. Monte Carlo sampling has the advantage that it characterizes the shape of the solution space and frequency plots can be drawn up for each flux. These histograms represent the flux probability in the solution space. Monte Carlo sampling has the disadvantage that rare flux solutions are not frequently sampled and flux ranges are not necessarily adequate. Extremal values for each flux can be determined through FVA, i.e., each flux is minimized/maximized while satisfying the system constraints as well as the optimum objective function. Flux variability ranges and ranges determined through Monte Carlo sampling are computed for GP1 and GP2 and summarized in Fig. 4. Fig. 5 overlays flux histograms (for the central metabolism) of GP1 and GP2.

The extent of flux ranges illustrate the existence of alternate optima within the central metabolism. A quick comparison of flux ranges between both growth phases immediately reveals that the network is more confined in GP1 as in GP2 (i.e., less flux variability). Plotting flux variability versus Monte Carlo sampling ranges shows a discrepancy in the sense that many Monte Carlo sampled flux ranges do not cover the whole FVA range. The smaller Monte Carlo sampling ranges are most likely due to the fact that the flux space is very narrow at the extremal points (indicated by distributions that tail off) and thus are missed by the Monte Carlo sampling procedure. Extreme fluxes are less likely in the solution space but can *a priori* not be ruled out from being biologically feasible.

The comparison of the Monte Carlo histograms between both growth phases depicts a shift in the metabolism, i.e., distributions shift between growth phases and are even not overlapping in most cases. GP1 and GP2 are thus characterized by a tremendous metabolic shift, which the cells seem to handle without problem because no adaptation phase is observed in the biomass or exometabolome data. Furthermore, Fig. 5 confirms the increased flexibility in GP2. Flux distributions cover broad ranges and many flux combinations will return the same (observed) growth rate in GP2. The uncertainty on the internal flux vector is thus large which is opposite for GP1. In GP1, the feasible set of fluxes is much smaller. Largest difference between GP1 and GP2 is the uptake of three main substrates in GP1, i.e., glucose, asparate and glutamate, whereas glucose is the only main carbon source in GP2. Adding more substrates hence seems to confine the feasible solution space. Extremely wide flux distributions are observed for parallel reactions, as can be seen in the TCA cycle where α-ketoglutarate can be oxidized into succinyl-coA via 2-oxoglutarate synthase (r175 and

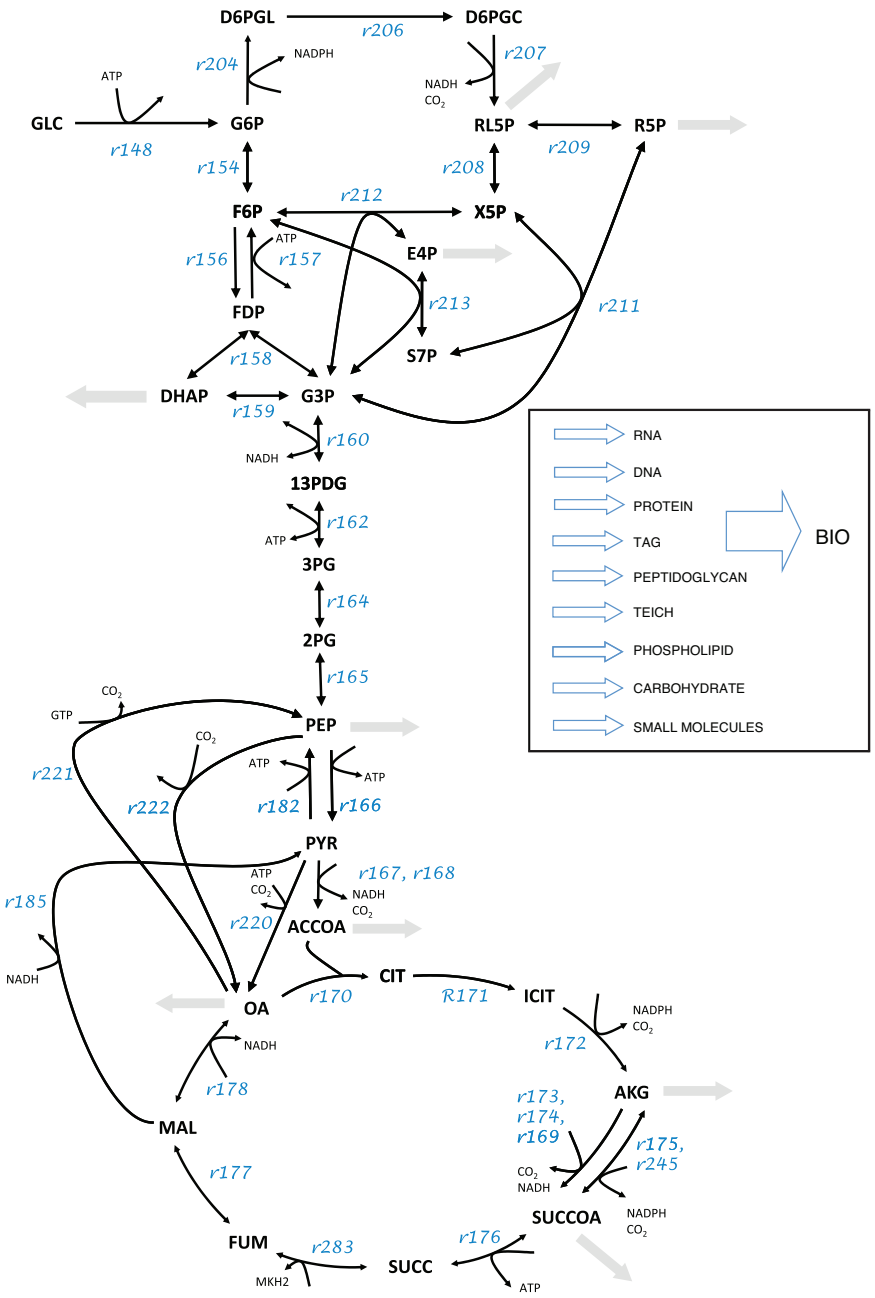

Fig. 3. Graphical representation of the primary metabolism as part of the genome-scale metabolic reaction network model of *S. lividans* TK24. Reaction numbers refer to the numbering applied in Supplemental data S2 and S5.

r245) or oxoglutarate dehydrogenase (r173 and r174). Figs. 4 and 5 also indicate the absence of a flux through the oxidative pentose pathway in GP1.

To conclude, flux ranges that give information about network flexibility and robustness can be determined by means of FVA and Monte Carlo sampling, but the latter additionally gives information about the shape of the remaining optimal growth solution space by means of flux probability histograms.

3.3. Selection of a unique intracellular flux vector

This section searches for a unique flux set by pursuing multiple cellular objectives using hierarchical FBA and compares this unique flux sets with the fluxes that have the highest spacial probability in the optimal growth solution space. The latter are derived from the flux distributions established in the previous section and depicted in Fig. 5. These fluxes correspond to the maximum of the flux probability distributions of the optimal growth solution space. The resulting flux vector, however, does not necessarily fulfils the steady-state assumption.

Alternate solutions can be easily retrieved in large-scale flux balance analysis using FVA. However, many publications exploiting FBA report only a single flux set obtained through linear programming. LP solvers stop after finding a first solution and the physiologically dominant flux map is only predicted by chance. No computational or biological evidence supports this arbitrary choice. Attempts to tackle the problem of alternate optima (i.e., to find a unique flux vector) are proposed in Holzhütter (2004, 2006), Murabito et al. (2009) and Ponce de Leon et al. (2008) whom perform a secondary optimization, and Smallbone and Simenidis (2008) whom propose a geometrical approach. In the present work, a hierarchical FBA approach is followed. Subsequent optimizations

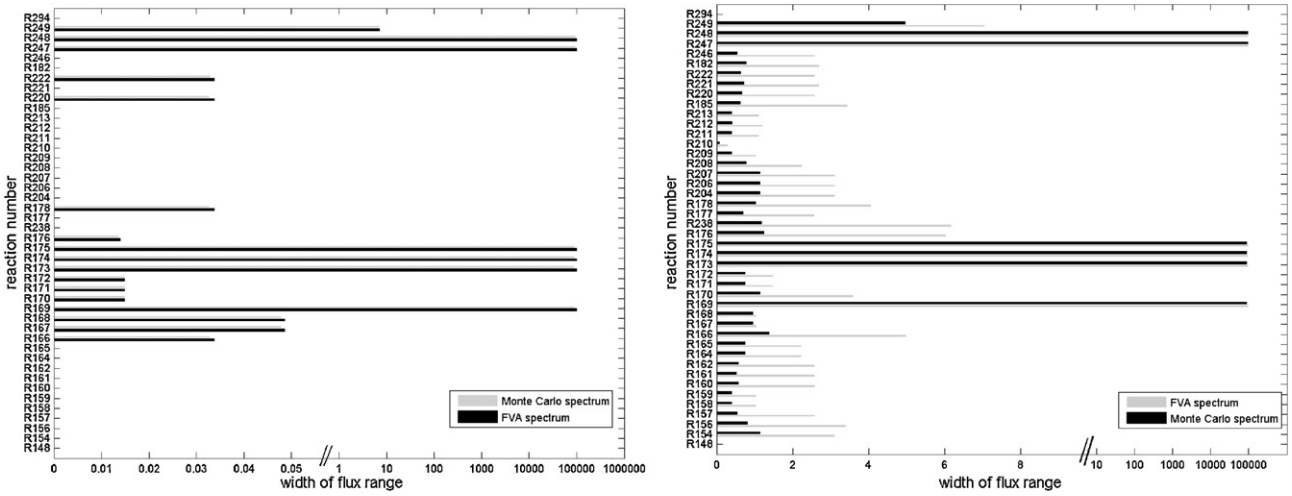

Fig. 4. Width of flux ranges for fluxes of the central metabolism obtained via flux variability analysis and Monte Carlo sampling of the metabolic steady-state solution space of Scenario E (reaction numbers in accordance with Fig. 3 and Supplemental data S2 and S5).

of cellular objectives are executed to retrieve a unique solution out of the optimal growth solution space for GP1 and GP2 (i.e., Scenario E). All selected secondary cellular objectives reflect the principle of *minimum effort*. Rationale behind this principle is that synthesis of enzymes and membrane transporters (i.e., the vital cornerstones of cellular metabolism) requires from the cell a certain *effort* of energy and external resources and it is plausible to assume that during the early phases of natural evolution, the competition for limited external resources represented a permanent pressure on living cells to fulfill their metabolic functions with minimal effort (Holzhütter, 2004). This principle of minimum effort can be mathematically expressed in various ways. Four previously published cellular objectives are applied. Unique flux vectors are obtained by applying two or three cellular objectives. In all cases, Scenario E forms the reference (max $v_{biomass}$) and different secondary cellular targets are optimized. Mathematically, all optimization problems are separable and are successively solved. Results are given in Table 3 and Supplemental data S5.

Minimization of the path (i.e., minimization of the number of active reactions) and minimization of the total flux sum (i.e., minimization of the magnitude of reaction fluxes) are based on the assumption that a cell strives for maximal enzymatic efficiency, either by reducing enzyme usage (Murabito et al., 2009) or by keeping fluxes low (Holzhütter, 2004). Neither criterion guarantees *a priori* a unique solution, such that the combination with a third cellular objectives may be required (see Table 3, top). Both objectives (separate or in combination) reduce the number of active reactions in the network. For example, only 295 reactions out of 702 reactions in the whole network are necessary to describe GP2 when the path is minimized. Since minimization of the total sum of quadratic fluxes puts more weight on minimizing large fluxes, this criterion is also tested. A unique solution is always found because the underlying optimization problem is quadratic and thus convex. For both growth phases, the number of active reactions is slightly higher because the fluxes are spread out over the network particularly within parallel reactions, alternate reaction routes and reaction for (re-)generation of cofactors.

In the fourth case, the predicted intracellular flux vector returns a maximum overall ATP yield per flux unit, which corresponds to the assumption that cells operate to maximize ATP production while minimizing enzyme usage. This non-linear optimization problem is non-convex and one has to avoid only finding a local optimum. This problem is tackled as follows. A first flux estimate is obtained by starting from 100 different initial flux vectors and retaining the flux set with the best objective function value. The original objective function is then reformulated into a convex function which is evaluated in the neighborhood of the first solution. No iteration in this sensitivity analysis results in an improved optimum value which gives strong indication that the global optimum is found in both growth phases. The final solution resulted in 338 active reactions for GP1 and 332 for GP2.

Each hierarchical FBA leads to a different intracellular flux vector (Supplemental data S2). Of course, now the question rises: which candidate flux vector is biologically meaningful and corresponds closest with the actual flux distribution? This question cannot be answered without having intracellular flux measurements and/or information on in vivo enzyme activities. Alternatively, each flux set can be compared with the computed (spacially) *most probable* set of fluxes. This selection criterion is reasonable since Thiele et al. (2005) found that when their human mitochondrial metabolic network was combined with constraints on exchange reactions, their sampling results were in agreement with in vivo intracellular measurements. Computation of the L1-norm provides an overall value for the distance between the flux estimates. The mean absolute difference gives the average difference. Both values are listed in Table 3.

Comparison of the criteria shows that GP1 as well as GP2 is best characterized by max $v_{ATP}/\sum v^2$ as a secondary objective function in combination with optimum biomass formation. The very large magnitude of the criterion values can be explained by the fluxes in parallel reactions, reaction loops and in a few reactions involved in cofactor (re-)generation. This set of reactions represents a minority in the total active network (namely 24 reactions) but flux distributions are very broad (Fig. 5) and the corresponding *most probable* flux values are very large (order of magnitude of 10^3 to 10^4) as compared to the flux estimates derived from the minimum effort principle (order of magnitude between 10^0 and 10^{-4}).

Although our approach to identify the most suitable (secondary) objective function is not underpinned by direct biological evidence, Schuetz et al. (2007) derived the same conclusion for the prediction of intracellular fluxes in *E. coli* during unlimited aerobic batch growth on the basis of experimental data. In vivo flux measurements (based on ^{13}C-labeling experiments) were compared with *in silico* predicted fluxes applying various objective functions which translate the optimality principle. Out of the tested cellular objective functions, maximization of the total ATP yield per flux unit gives the best description of intracellular fluxes during an aerobic batch culture.

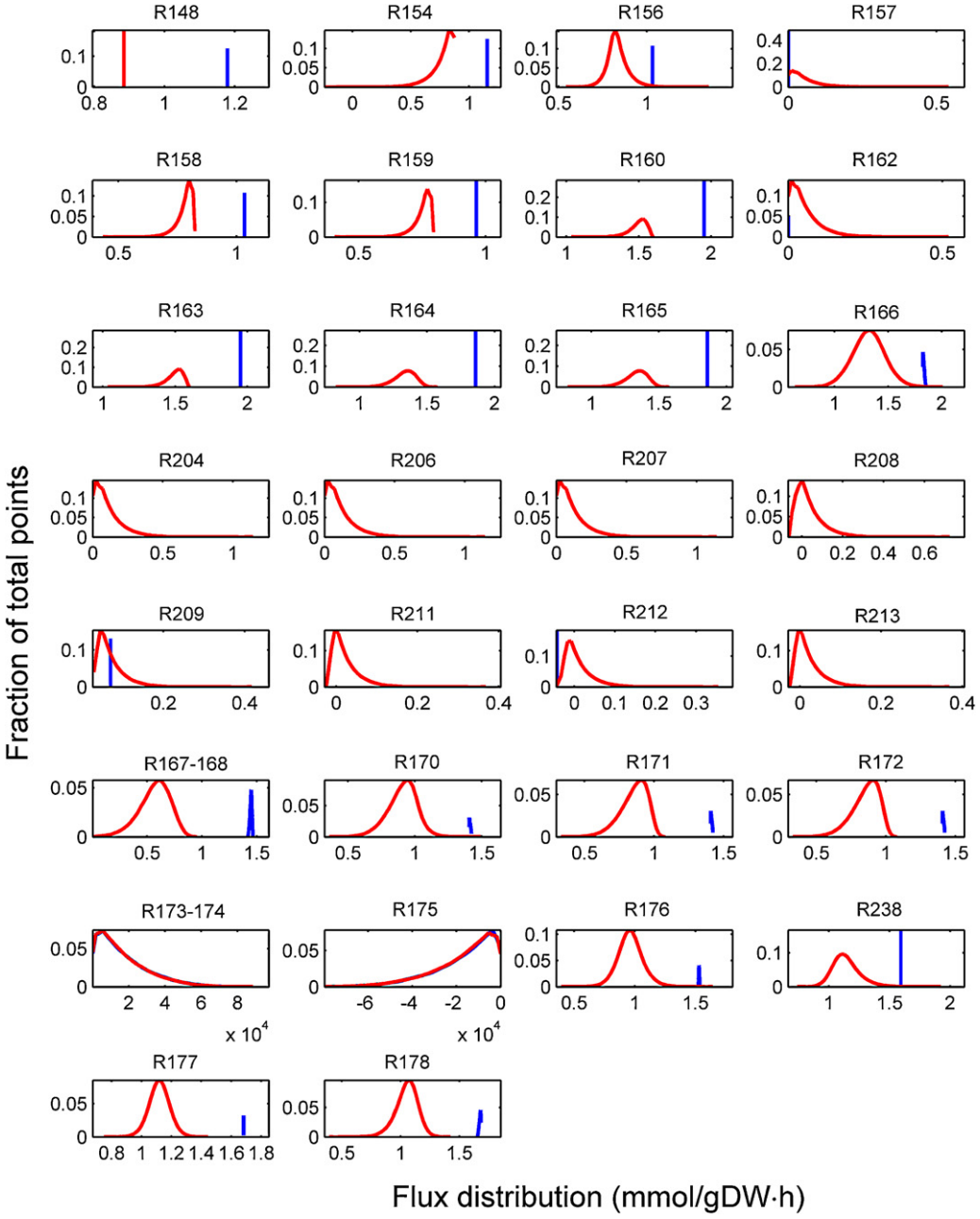

Fig. 5. Flux histograms (blue = GP1, red = GP2) resulting from Monte Carlo sampling of the metabolic steady-state solution space of Scenario E. Fluxes of the central metabolism are depicted (reaction numbers in accordance with Fig. 3 and Supplemental data S2 and S5). Rows 1–3: Glycolysis. Rows 4–5: pentose phosphate pathway. Rows 6–8: TCA cycle. Blue line: GP1. Red line: GP2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

3.4. Study of the intracellular flux distribution

Out of the candidate (unique) flux distributions, the flux solution satisfying maximum ATP yield per unit flux is retained and discussed in further detail. A comparison is made between GP1 and GP2. Special attention is paid to the central carbon and nitrogen metabolism, the fluxes of which are summarized in Table 4. The complete flux vector is available in Supplemental data S4. Not underestimating the fact that there is no hard prove that the *in silico* predicted fluxes agree with in vivo fluxes, our conclusions are underpinned by previous works (Schuetz et al., 2007; Thiele et al., 2005) and indicative insights can be gained in the operation and capabilities of the metabolism of *S. lividans* under the studied experimental conditions.

GP1 is characterized by a higher biomass production rate and glucose uptake rate such that fluxes in general are larger than those during GP2, which is most pronounced in the TCA cycle. Normalized flux values indicate that the glycolytic flux in GP2 is much higher as in GP1, whereas in GP1 the flux through the TCA cycle is significantly larger.

The oxidative part of the pentose phosphate pathway (PPP) is in both cases inactive. In GP1, this zero-flux through glucose-6-phosphate dehydrogenase (G6PDH) and 6-phosphogluconate dehydrogenase (6PGDH) is stringent (see Section 3.3), whereas the zero flux in GP2 only results after imposing the secondary objective function $\max v_{ATP}/\sum v^2$ (see Table 4). The oxidative PPP is generally active in micro-organisms. Both G6PDH and 6PGDH typically generate NADPH required for biosynthesis. A first indication

Table 3
Overview of hierarchical flux balance analyses. Constraints are defined in Scenario E (first level of flux optimization). Flux vectors are available in Supplemental file S5. For each solution, all function values are calculated. The L1-norm and mean absolute difference between the flux vectors obtained from hierchical FBA and Monte Carlo sampling are computed.

Sequential optimization	Aim		*Path*	$\sum \lvert v \rvert$	$\sum v^2$	$\frac{v_{ATP}}{\sum_j v_j^2}$	L1-norm	Mean |diff|
min *path*	Minimization of the overall flux sum (absolute values) followed by a minimization of the active reaction steps	GP1:	309	**86.041**	382.28	4.2730E−02	706059.302	1722,096
s.t. min $\sum \lvert v \rvert$		GP2:	318	**54.389**	142.06	6.5709E−02	884625.965	1996,898
s.t. max μ								
min *path*	Minimization of the number of active reactions	GP1:	**303**	128.20	695.76	2.3478E−02	706061.226	1722,101
s.t. max μ		GP2:	**295**	61.217	233.52	3.8418E−02	884646.414	1996,945
min $\sum v^2$	Minimization of the overall flux (quadratic)	GP1:	318	90.855	**360.54**	4.5307E−02	706054.642	1722,085
s.t. max μ		GP2:	319	58.008	**117.03**	8.1710E−02	884622.272	1996,890
max $\frac{v_{ATP}}{\sum_j v_j^2}$	Maximization of the ATP yield per flux unit	GP1:	338	90.914	360.62	**4.5341E−02**	706054.639	1722,084
s.t. max μ		GP2:	332	66.137	151.42	**8.2973E−02**	884621.220	1996,888

of an alternate route for NADPH production in *Streptomyces lividans* is reported in Butler et al. (2002) and Dekleva and Strohl (1988). Instead of using NADP, 6PGDH is shown to be NAD-dependent. Moreover, blockage of G6PDH still assures normal growth behavior as compared to the parental strain. Butler et al. (2002) hypothesize NADPH is supplied by the TCA cycle and/or by the NAD(P)-transhydrogenase (which can reversibly interchange levels of NADH and NADPH). Current simulations indeed appoint complete NADPH-supply by the TCA cycle, more particularly, by the isocitrate dehydrogenase and ferredoxin-NADP reductase (resulting from the activity of the 2-oxoglutarate synthase) activity, and by transhydrogenase activity in GP2. In GP1, the flux through the NAD(P)-transhydrogenase is estimated such that NADPH is dissipated. Apparently, the high inflow of carbon sources, particularly, glutamate and aspartate, via the TCA cycle generates very high amounts of NADPH which the cell cannot use for biosynthesis. Another potential enzyme supporting NADPH supply is the NADP-dependent glyceraldehyde-3-phosphate dehydrogenase which is present in the genome-scale network. Although this enzyme has a FVA range in GP2 (r162), no flux through this reaction is predicted in neither growth phase.

Genome-scale networks include all possible annotated genes found for an organism, but do not take into account that certain enzymes are only synthesized under particular environmental conditions. The genome-scale network of *S. lividans* contains two enzymatic routes for the formation of α-ketoglutarate into succinyl-coA, i.e., either via the 2-oxoglutarate dehydrogenase complex or via the 2-oxoglutarate synthase (or ferredoxin oxidoreductase) coupled to ferredoxin-NADP reductase (see Fig. 3). In both growth phases, a flux through both enzymatic routes is predicted. The activity of 2-oxoglutarate synthase is although somewhat unexpected. Together with citrate lyase, this enzyme is usually an indication for a possible reductive mode of the TCA cycle. This reductive mode is known to be used by persistent pathogenic bacteria like *Mycobacterium tuberculosis*, *Heliobacter pylori* and *Samonella* Typhimurium during the infection proces when they experience harsh and hostile environments. Probably there is a metabolic switch to anaerobic respiration under these circumstances in which a complete or partial TCA cycle may operate in reductive mode to enable pathogen survival, growth and persistence (Srinivasan and Morowitz, 2006). To know the exact role of 2-oxoglutarate synthase in *S. lividans*, further investigation (e.g., measuring enzyme activity) is necessary. But because this reaction is reversible and thus could possibly also be active in the oxidative mode of the TCA cycle, the reaction has been maintained during the present analysis.

Also remarkable is the absence of a flux through pyruvate dehydrogenase (PDH), i.e., from pyruvate to acetyl-CoA, in GP2. Unlike in GP1 where most of the AcCoA is synthesized directly from pyruvate, acetyl-CoA is synthesized from pyruvate via a *detour* (including reactions of threonine and isoleucine biosynthesis) and the flux through PDH is predicted zero. When compared to the corresponding Monte Carlo histogram, the predicted zero PHD flux is not close to the top (approximately 0.603 mmol gDW^{-1} h^{-1}) which may put some doubt on the flux distribution through PDH. Apart from PDH, fluxes for all other reactions from the central metabolism correspond reasonably well with the top of the histograms.

Another difference between GP1 and GP2 is the usage of the anaplerotic and gluconeogenic reactions. The medium conditions corresponding to GP1, i.e., the availability of glutamate and aspartate, make that the gluconeogenic reactions catalized by malic enzyme and phosphoenolpyruvate carboxykinase are active (although in a very small amount). This corresponds to the results of Nowruzi et al. (2011) who also predicted gluconeogenic activity in the form of malic enzyme activity (which was the only gluconeogenic enzyme in their model) for *S. lividans* during growth on a medium with glutamate, aspartate and phenylalanine. When

Table 4
Flux estimates based on maximization of $v_{ATP}/\sum_j v_j^2$ (as a secondary objective function). Reactions of the central carbon and nitrogen metabolism is listed. Fluxes are normalized with respect to the glucose uptake rate. Original flux values are available in Supplemental data S5.

	Flux		Reaction number and stoichiometry
	GP1	GP2	
	100,000	100,000	r148 : 1 GLC +1 ATP ->1 G6P +1 ADP
	97,975	98,886	r154 : 1 G6P < ->1 F6P
	87,648	93,243	r156 : 1 F6P +1 ATP ->1 FDP +1 ADP
	87,648	93,243	r158 : 1FDP< ->1G3P+1DHAP
	81,700	89,969	r159 : 1 DHAP < ->1 G3P
Glycolysis	165,599	181,269	r160 : 1 G3P +1 PI +1 NAD < ->1 NADH +1 13PDG
	0,000	0,000	r161 : 1 G3P +1 NADP ->1 NADPH +1 3PG
	165,599	181,269	r162 : 1 13PDG +1 ADP < ->1 3PG +1 ATP
	157,739	175,970	r164 : 1 3PG < ->1 2PG
	157,739	175,970	r165 : 1 2PG < ->1 PEP
	154,672	174,685	r166 : 1 PEP +1 ADP ->1 PYR +1 ATP
	0,000	0,000	r204 : 1 G6P +1 NADP ->1 D6PGL +1 NADPH
	0,000	0,000	r206 : 1 D6PGL ->1 D6PGC
	0,000	0,000	r207 : 1 D6PGC +1 NAD ->1 NADH +1 CO2 +1 RL5P
	−6,271	−3,411	r208 : 1 RL5P < ->1 X5P
Pentose phosphate pathway	5,956	3,237	r209 : 1 RL5P < ->1 R5P
	−2,705	−1,569	r211 : 1R5P+1X5P< ->1G3P+1S7P
	−3,566	−1,842	r212 : 1 X5P +1 E4P < ->1 F6P +1 G3P
	−2,705	−1,569	r213 : 1 G3P +1 S7P < ->1 E4P +1 F6P
	120,157	0,000	r167 : 1 PYR +1 LIPO ->1 ADLIPO +1 CO2
	120,157	0,000	r168 : 1 COA+1 ADLIPO ->1 DLIPO +1 ACCOA
	119,360	65,068	r170 : 1 ACCOA +1 OA->1 COA+1 CIT
	119,360	65,068	r171 : 1 CIT < ->1ICIT
	119,360	65,068	r172 : 1 ICIT +1 NADP ->1 NADPH +1 AKG +1 CO2
TCA	39,411	2,900	r173 : 1 AKG +1 LIPO ->1 SDLIPO +1 CO2
	39,411	2,900	r174 : 1 SDLIPO +1 COA->1 DLIPO +1 SUCCOA
	95,677	61,104	r175 : 1 AKG +2 FERI +1 COA< ->1 CO2 +2 FERO +1 SUCCOA
	130,367	49,672	r176 : 1 ADP +1 PI +1 SUCCOA < ->1 ATP +1 SUCC +1 COA
	135,234	50,466	r238 : 1 SUCC +1 MK +2 H < ->1 FUM +1 MKH2
	142,752	51,765	r177 : 1 FUM < ->1 MAL
	142,703	51,765	r178 : 1 MAL +1 NAD < ->1 NADH +1 OA
	0,050	0,000	r185 : 1 MAL +1 NAD ->1 CO2 +1 NADH +1 PYR
Anaplerotic reactions	0,000	29,346	r220 : 1 PYR +1 ATP +1 CO2 ->1 ADP +1 OA +1 PI
	0,048	0,000	r221 : 1 OA+1 GTP->1 PEP+1 CO2 +1 GDP
	0,000	0,000	r222 : 1 PEP+1 CO2->1OA+1PI
	0,000	0,000	r246 : 1 GLU +1 NH3 +1 ATP ->1 GLN +1 ADP +1 PI
	−24,872	−10,848	r247 : 1 AKG +1 GLN +1 NADH < ->1 NAD +2 GLU
Nitrogen metabolism	103,734	98,753	r248 : 1 AKG +1 NH3 +1 NADPH < ->1 GLU +1 NADP
	18,874	28,796	r294 : 1PYR+1NH3 +1NADH< ->1ALA+1 NAD
	79,916	30,445	r275 : 1 OA+1 GLU < ->1 ASP+1 AKG
	51,648	−59,546	r489 : 1 NADPH +1 NAD < ->1 NADP +1 NADH

glutamate and aspartate were no longer available for the cells, this enzyme was found to be inactive.

For our case, both the anaplerotic reaction catalyzed by pyruvate carboxylase and the gluconeogenic reaction catalyzed by the malic enzyme are active in GP2 after depletion of glutamate and aspartate. However, the latter is substantially smaller causing the netto flux to be in the anaplerotic direction. Investigation of the flux-node around oxaloacetate makes clear that the flux by the malate dehydrogenase is too small in this growth phase and extra oxaloacetate should be supplied by the anaplerotic reactions to close the TCA cycle. Unlike in most literature where the phosphoenolpyruvate carboxylase is regarded as the enzyme primary responsible for the TCA cycle metabolite replenishment (Dekleva and Strohl, 1988; Bramwell et al., 1993), pyruvate carboxylase is the only active anaplerotic enzyme during our similations. Pyruvate carboxylation is, however, also reported by Borodina et al. (2005b) as the anaplerotic pathway giving the best fit to ^{13}C-labeling patterns in a glucose fermentation with *S. tenebrarius*.

Furthermore, the ammonium assimilation pathways are the same for GP1 and GP2. During both growth phases, ammonium is assimilated via the glutamate dehydrogenase (GDH)–glutamine-oxoglutarate aminotransferase (GOGAT) couple to make intracellular glutamate and glutamine (r248 and r247 in Table 4). This generation of both glutamate and glutamine by the GDH–GOGAT couple in *Streptomyces* is typical for high ammonium concentrations (Voelker and Altaba, 2001; Hodgson, 2000).

The same approach as explained above for GP1 and GP2 has also been adopted to GP (maximize ATP per flux unit fluxvector shown in Supplemental data S4, rest of the data not shown). GP shows similarity with GP2, which is a bit unexpected since glutmate and aspartate are available for the cell during the whole exponential growth phase. However, the lower influx of glutamate and aspartate, as potentially caused by inhibitory effects of higher amino acid concentrations (D'Huys et al., 2011), might be the explanation for a flux pattern close to the flux pattern of GP2.

3.5. Role of amino acids in the medium

The importance of glutamate and aspartate is immediately obvious from biomass and exometabolome data. Glutamate has a role in ammonium assimilation, but also contributes to many other pathways. To assess the distribution of glutamate and aspartate in the network, flux-sums are computed. This means that the

consumption and production fluxes are summed around the glutamate and aspartate node, and their distribution is analyzed. The total flux through the glutamate pool in GP1 is 1.82 mmol gDW^{-1} h^{-1}. About 16% is delivered by glutamate uptake, and the majority (approximately 67%) comes from the glutamate biosynthetic pathway (via GDH). Because of the absence of glutamine in the casamino acids mixture, circa 32% of the glutamate is used to synthesize glutamine by means of the GOGAT. To meet the biosynthetic demand of glutamate for biomass growth only 1.23% is used, and circa 52% is used to synthesize aspartate via the glutamate-oxaloacetate aminotransferase (GOAT). The flux through the aspartate pool is 1.048 mmol gDW^{-1} h^{-1} in GP1. Uptake supplies 10% whereas 90% is formed via glutamate through the aspartate biosynthetic pathway (GOAT). Synthesis of aspartate is important as a precursor for several other amino acids, i.e., threonine, arginine and asparagine and for nucleotide synthesis. In GP2, without extracellular glutamate and aspartate available, the largest portion (up to 87%) of glutamate comes from GDH activity. Aspartate is synthesized from glutamate by GOAT.

The rapid and preferential uptake of both glutamate and aspartate in GP1 thus makes that the cell has to synthesize less of this important components by itself and makes it possible for the cell to maintain the measured high specific growth rate because both amino acids play an important role in the nitrogen assimilation and as a precursor for a lot of other metabolites.

Finally, we investigate the influence of the availability of amino acids in the medium on the metabolism. Amino acids can act as both carbon and nitrogen sources. To investigate the contribution of the availability of amino acids to biomass growth, a hypothetical case of a minimal medium is simulated for GP1 and GP2 by supplying glucose and ammonium as sole carbon and nitrogen source. During this simulations, glucose and ammonium uptake are left unbounded and the maximum growth rates are fixed to the values predicted by the experimentally constrained output system. The glucose and ammonium uptake rates and the metabolic reaction fluxes of the amino acid biosynthetic reaction of the maximize ATP per flux unit solution for the minimal medium are then compared to the solution for the experimentally constrained output system discussed above. Results are summarized in Supplemental data S3. Besides an increased glucose and ammonium requirement in the minimal medium, there are also higher fluxes through most amino acid biosynthetic pathways in the simulated minimal medium condition. This shows that due to the uptake of amino acids, the biosynthetic demands for protein synthesis could be partially fulfilled.

This *in silico* analysis reveals thus that the cell exploits biosynthetic precursors taken up directly from the medium, through growth-related anabolic pathways. This suggests that the cell could be functioning in an energetically more efficient manner by reducing the energy needed to produce amino acids.

4. Conclusions

Extracellular dynamics of *S. lividans* grown on a nutritious medium have been mapped with a genome-scale metabolic model. Notwithstanding the high number of substrates and inherent experimental uncertainty on measured exchange fluxes, macroscopic observations are consistent with the stoichiometric model, steady-state assumption and applied constraints. Presenting maximum specific substrate uptake rates to the cell and enforcing the correct secretion pattern of organic acids and alanine, the observed specific growth rate is very well predicted by flux balance analysis with the aim for maximum biomass formation. Although organic acid secretion can be consistent with the optimal biomass growth hypothesis, alanine secretion is not. Alanine secretion discards carbon as well as nitrogen from the cell, hereby lowering the ability for biomass formation. Alanine secretion is therefore in contradiction with biomass growth and the reason for alanine secretion probably needs to be sought in the excess ammonium for the cell, as illustrated by the experiments shown.

Imposing known exchange rates as case-specific constraints, a confined steady-state solution space is obtained. Variability on intracellular fluxes, as characterized by flux ranges and flux histograms derived from FVA and MCS, is yet still significant. A unique steady-state flux vector in this solution space is derived through hierarchical flux balance analysis, in which the principle of minimal effort is considered as a second cellular objective. Out of the tested criteria, maximization of the ATP yield per flux unit returns the closest resemblance with the peak values of flux histograms. Strikingly, this agreement is consistent with findings of Thiele et al. (2005) and Schuetz et al. (2007). It gives further confidence in using this criterion as a proper criterion for getting a first good prediction for intracellular fluxes during growth under batch conditions, in cases where in vivo intracellular flux measurements are not available.

Flux maps and histograms demonstrate differences between growth phases. The magnitude of fluxes, the distribution of fluxes as well as the flux variability change when the substrate availability changes. A combination of glucose, ammonium, glutamate and aspartate as main substrates creates a great flux through the TCA cycle, pushes gluconeogenisis and hereby drives biomass growth. Apparently, an excess in NADPH is generated which is dissipated through transhydrogenase activity. After depletion of glutamate and aspartate, the cells shifts to a lower growth rate with higher levels of secreted products. Replenishment of the TCA cycle is required through anaplerotic reactions. The flux ratio between glycolysis and TCA cycle changes involving changes in the energy (ATP) and reductive power (NADPH, NADH) ratio.

Current work forms a basis for further studies of *S. lividans* in complex media. D'Huys et al. (2011) observed a discrepancy between growth rate and mTNF-α production of *S. lividans* in NMMP supplemented with casamino acids. The protein yield increases after exhaustion of glutamate and aspartate and a slow down of biomass growth. It would be interesting to retrieve the link between heterologous protein secretion and intracellular fluxes in order to engineer strains with improved protein production.

Notwithstanding that fast the current flux predictions are reasonable flux estimates, ^{13}C-labeling experiments for in vivo determination of central metabolic fluxes and measurements of enzyme activity are the only way to unambiguously identify fluxes. On the one hand, a genome-scale network model includes all identified genes (enzymatic reactions) but does not take into account the fact that some enzymes will not be synthesized under certain conditions. The model could be adapted to better reflect enzyme usage. On the other hand, ^{13}C-MFA is the only manner to accurately probe fluxes in the central metabolism, but standard approaches relying on feeding of labelled glucose will not be sufficient. Amino acids will need to be considered as, for example, in Metallo et al. (2009).

Acknowledgements

Work supported by the project FWO G.0352.09 of the Research Foundation - Flanders (FWO), in part by projects DBOF/08/033, OT/09/25 and EF/05/006 (OPTEC Optimization in Engineering) of the Katholieke Universiteit Leuven, by the Belgian Program on Interuniversity Poles of Attraction, initiated by the Belgian Federal Science Policy Office, and by FP6-EU Project no. LSHC-CT-2006-037834, and in the frame of ERA-IB project EIB.08.013. J. Van Impe holds the chair Safety Engineering sponsored by the Belgian chemistry and life sciences federation essenscia. The scientific responsibility is assumed by its authors.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbiotec.2012.04.010.

References

Anné, J., Maldonado, B., Van Impe, J., Van Mellaert, L., Bernaerts, K., 2011. Recombinant protein production and streptomycetes. Journal of Biotechnology, http://dx.doi.org/10.1016/j.jbiotec.2011.06.028.

Becker, S., Feist, A., Mo, M., Hannum, G., Palsson, B., Herrgard, M.J., 2007. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox. Nature Protocols 2, 727–738.

Borodina, I., Krabben, P., Nielsen, J., 2005a. Genome-scale analysis of *Streptomyces coelicolor* A3(2) metabolism. Genome Research 15, 820–829.

Borodina, I., Schšller, C., Eliasson, A., Nielsen, J., 2005b. Metabolic network analysis of *Streptomyces tenebrarius*, a Streptomyces species with an active entner-doudoroff pathway. Applied and Environmental Microbiology 71, 2294–2302.

Bramwell, H., Nimmo, H., Hunter, I., Coggins, J., 1993. Phosphoenolpyruvate carboxylase from *Streptomyces coelicolor* A3(2): purification of the enzyme, cloning of the ppc gene and over-expression of the protein in a streptomycete. Biochemical Journal 293, 131–136.

Brana, A., Paiva, N., Demain, A., 1986. Pathways and regulation of ammonium assimilation in *Streptomyces clavuligerus*. Journal of General Microbiology 132, 1305–1317.

Butler, M., Bruheim, P., Jovetic, S., Marinelli, F., Postma, P., Bibb, M., 2002. Engineering of primary carbon metabolism for improved antibiotic production in *Streptomyces lividans*. Applied and Environmental Microbiology 68, 4731–4739.

Dekleva, M.L., Strohl, W.R., 1988. Biosynthesis of e-rhodomycinone from glucose by *Streptomyces* C5 and comparison with intermediary metabolism with other polyketide producing streptomycetes. Canadian Journal of Microbiology 34, 1235–1240.

D'Huys, P., Lule, I., Van Hove, S., Vercammen, D., Wouters, C., Bernaerts, K., Anné, J., Van Impe, J., 2011. Amino acid uptake profiling of wild type and recombinant *Streptomyces lividans* TK24 batch fermentations. Journal of Biotechnology 152, 132–143.

Edwards, J., Palsson, B., 2000. The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. PNAS 97, 5528–5533.

Edwards, J., Ibarra, R., Palsson, B., 2001. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. Nature Biotechnology 19, 125–130.

Fisher, S., 1989. Glutamate synthesis in *Streptomyces coelicolor*. Journal of Bacteriology 171, 2372–2377.

Fong, S., Palsson, B., 2004. Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. Nature Genetics 36, 1056–1058.

Hodgson, D.,A., 2000. Primary metabolism and its control in streptomycetes. Advances in Microbial Physiology 42, 47–238.

Holzhütter, H.G., 2004. The principle of flux-minimization and its application to estimate stationary fluxes in metabolic networks. European Journal of Biochemistry 271, 2905–2922.

Holzhütter, H.G., 2006. The generalized flux-minimization method and its application to metabolic networks affected by enzyme deficiencies. Biosystems 83, 98–107.

Kaufman, D., Smith, R., 1998. Direction choice for accelerated convergence in hit-and-run sampling. Operations Research 46, 84–95.

Kieser, T., Bib, M., Buttner, M., Chater, K., Hopwood, D., 2000. Practical Streptomyces Genetics. The John Innes Foundation.

Korn-Wendish, F., Weingartner, B., Kutzner, H., 1978. A study of twenty actinophages: morphology, serological relationships and host range. In: Freerksen, E., Tarnak, I., Thumin, J. (Eds.), Genetics of the Actinomycetales. Fisher G., Stuttgart & New York, pp. 251–270.

Ponce de Leon, M., Cancela, H., Acerenza, L., 2008. A strategy to calculate the patterns of nutrient consumption by microorganisms applying a two-level optimisation principle to reconstructed metabolic networks. Journal of Biological Physics 34, 73–90.

Llaneras, F., Picó, J., 2007. An interval approach for dealing with flux distributions and elementary modes activity patterns. Journal of Theoretical Biology 15, 290–308.

Llaneras, F., Picó, J., 2008. Stoichiometric modelling of cell metabolism. Journal of Bioscience and Bioengineering 15, 1–11.

Mahadevan, R., Schilling, C., 2003. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metabolic Engineering 5, 264–276.

Makhorin, A. Gnu Linear Programming Kit. Free Software Foundation, Inc., Boston, MA, in press.

Mazumdar, V., Snitkin, E., Amar, S., Segre, D., 2009. Metabolic network model of a human oral pathogen. Journal of Bacteriology 191, 74–90.

Metallo, C., Walther, J., Stephanopoulos, G., 2009. Evaluation of ^{13}C isotopic tracers for metabolic flux analysis in mammalian cells. Journal of Biotechnology 144, 167–174.

Murabito, E., Simeonidis, E., Smallbone, K., Swinton, J., 2009. Capturing the essence of a metabolic network: a flux balance analysis approach. Journal of Theoretical Biology 260, 445–452.

Novak, J., Kopecky, J., Venek, Z., 1997. Nitrogen source regulates expression of alanine dehydrogenase isoenzymes in *Streptomyces avermitilis* in a chemically defined medium. Canadian Journal on Microbiology 43, 189–193.

Nowruzi, K., Elkamel, A., Scharer, J., Cossar, D., Moo-Young, M., 2011. Metabolic flux-based optimisation of recombinant human interleukin-3 expression by *Streptomyces lividans* 66. The Canadian Journal of Chemical Engineering 9999.

Price, N., Reed, J., Palsson, B., 2004. Genome-scale models of microbial cells: evaluating the consequences of constraints. Nature Reviews Microbiology 2, 886–897.

Reed, J.L., Palsson, B., 2004. Genome-scale in silico models of *E. coli* have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states. Genome Research 14, 1797–1805.

Schellenberger, J., Palsson, B., 2009. Use of randomized sampling for analysis of metabolic networks. Journal of Biological Chemistry 284, 5457–5461.

Schuetz, R., Kuepfer, L., Sauer, U., 2007. Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. Molecular Systems Biology 3, 119.

Selvarasu, S., Wong, V., Karimi, I., Lee, D.Y., 2009. Elucidation of metabolism in hybridoma cells grown in fed-batch culture by genome-scale modeling. Biotechnology and Bioengineering 102, 1494–1504.

Sendín, J.O.H., Alonso, A., Banga, J., 2009. Multi-objective optimization of biological networks for prediction of intracellular fluxes. Advances in Soft Computing 49/2009, 197–205.

Smallbone, K., Simenidis, E., 2008. Flux balance analysis: a geometric perspective. Journal of Theoretical Biology 258, 311–315.

Srinivasan, V., Morowitz, H., 2006. Acient genes in contemporary persistent microbial pathogens. Biology Bulletin 210, 1–9.

Teusink, B., Wiersma, A., Molenaar, D., Francke, C., de Vos, W.M., Siezen, R.J., Smid, E.J., 2006. Analysis of growth of *Lactobacillus plantarum* wcfs1 on a complex medium using a genome-scale metabolic model. The Journal of Biological Chemistry 281, 40041–40048.

Thiele, I., Palsson, B.O., 2007. Bringing Genomes to Life: The use of Genome-scale in Silico Models. Humana Press.

Thiele, I., Price, N., Vo, T.D., Palsson, B., 2005. Candidate metabolic network states in human mitochondria. Journal of Biological Chemistry 280, 11683–11695.

Voelker, F., Altaba, S., 2001. Nitrogen source governs the patterns of growth and pristinamycin production in *Streptomyces pristinaespiralis*. Microbiology 147, 2447–2459.

Vrancken, K., Anné, J., 2009. Secretory production of recombinant proteins by *Streptomyces*. Future Microbiology 4, 181–188.

Wiback, S., Famili, I., Greenberg, H., Palsson, B., 2004. Monte Carlo sampling can be used to determine the size and shape of the steady-state flux space. Journal of Theoretical Biology 228, 437–447.