Proceedings of the 11th International Symposium on
Computer Applications in Biotechnology (CAB 2010),
Leuven, Belgium, July 7-9, 2010
Julio R. Banga, Philippe Bogaerts, Jan Van Impe,
Denis Dochain, Ilse Smets (Eds.)

Metabolic Flux Analysis of CHO-320 Cells: Undetermined Network and Effect of Measurement Errors

F. Zamorano * **A. Vande Wouwer** *,1 **G. Bastin** **

* *Service d'Automatique, Université de Mons (UMONS), Faculté Polytechnique, Boulevard Dolez 31, 7000 Mons, Belgium*
** *CESAME, Université Catholique de Louvain, Louvain-La-Neuve, Belgium*

Abstract: The flux distribution in a detailed metabolic network of CHO-320 cells is evaluated using Metabolic Flux Analysis. As in many practical situations, the available information is not sufficient to completely define the metabolic fluxes, and so, the mass balance system of equations is underdetermined. However, the measurements of the time evolution of a number of extracellular components can provide a set of constraints on the metabolic network, so that a range of possible (non-negative) solutions for each metabolic reaction can be computed instead. In this way, metabolic flux intervals can be established for each intra-extracellular flux in the metabolic network. Moreover, the incorporation of simple theoretical assumptions or the addition of further extracellular measurements, result in the determination of certain metabolic fluxes and the delimitation of quite narrow intervals for the others, so providing a good guess of the real flux distribution in CHO-320 cells. A unique flux distribution can also be computed through linear optimization and the definition of some optimality criteria.
In this article, several of the above-mentioned situations are discussed, highlighting the importance of specific measurements and comparing the flux intervals under some particular assumptions such as optimal biomass growth or no Threonine catabolism. In addition, the influence of the measurement errors on the numerical results is explored using Monte Carlo techniques.

Keywords: metabolic flux analysis, mammalian cells, metabolic networks, underdetermined systems, convex analysis, measurement error.

1. INTRODUCTION

Metabolic Flux Analysis is a useful tool to unravel the complexity of cell metabolism through the analysis of an underlying stoichiometric reaction network. It has been the subject of a thorough study by several research groups for almost two decades, giving rise to a large number of publications (Bonarius et al., 1997; Stephanopoulos et al., 1998; Henry et al., 2005; Antoniewicz et al., 2006).

In contrast to other studies, this article is concerned with the metabolic flux analysis (MFA) of a much more detailed metabolic network (100 metabolic reactions) and the investigation of the information that can be extracted considering conventional measurement methods only. This metabolic network is built with the main purpose of embracing all the significant pathways describing the metabolism of CHO-320 cells. In general, when only external measurements are used, this type of complex network leads to an underdetermined system because the number of external measurements is not sufficient to provide all the information needed to obtain a unique solution. In the literature, various suggestions are made to overcome this problem (e.g., (Bonarius et al., 1997) and (Stephanopoulos et al., 1998)): (i) the introduction of additional metabolic theoretical constraints; (ii) the exploitation of linear optimization tools and the definition of suitable objective functions; or (iii) the use of isotopic tracer experiments to determine some intracellular fluxes.

In this study, we focus our attention on the flux analysis of the underdetermined system considering a standard set of measurements (mostly of exo-cellular components) which can be easily achieved in most laboratories. Although isotopic label experiments are very useful for the quantification of intracellular fluxes, it is however a fairly laborious and expensive method. There is therefore a clear incentive to further investigate the feasibility and performance of MFA for underdetermined systems without using isotopic tracer methods.

To tackle our detailed metabolic network and analyze the set of admissible solutions, we use the approach developed in Provost and Bastin (2004), which allows the determination of intervals for the metabolic fluxes. Using this computational approach, we investigate the influence of the set of available extracellular measurements on the size of the metabolic flux intervals.

[1] Author to whom correspondence should be addressed - Email: Alain.VandeWouwer@umons.ac.be, Tel: +32 (0)65374141, Fax: +32 (0)65374136

On the other hand, MFA relies strongly on the experimental data at disposal and in turn on the accuracy of these measurements. It is then of interest to study the effect of measurement errors on the computed flux intervals. In this work, the error distribution is evaluated using Monte Carlo techniques.

This study shows that, although the system is underdetermined, very narrow intervals may be found for most fluxes. In particular, a single well defined value can be obtained for the formation rates of the cellular macromolecules (proteins, DNA, RNA, lipids) that are not measured. We also investigate the potential gain of a simple natural assumption regarding the Threonine (Thr) metabolism and we compare this result with the optimal solution calculated by maximizing the overall biomass production rate.

This paper is organized as follows. Section 2 describes briefly the set of available experimental measurements. Section 3, is devoted to the principles of MFA methodology and to explain how the determination of the flux intervals is done. In Section 4 the numerical results, obtained under different scenarios, are presented. In Section 5, the effect of experimental error over the determination of the flux distribution intervals is assessed. Finally, Section 6 draws the main conclusions of this work.

2. EXPERIMENTAL DATA

The experimental data have been provided by the 'Unité de Biochimie', Université Catholique de Louvain. These experimental data correspond to measurements collected from the exponential growth phase of three different batch cultures of a CHO-320 cell line, carried out in a serum-free medium supplemented with rice protein hydrolysate and glutamine. These three experiments have been designed to have initial concentrations of 16 mM of glucose and 6 mM of glutamine.

This data set contains the time evolution of the extracellular concentrations of the main substrates: Glucose and Glutamine, the main metabolism excretion products: Lactate, Alanine and Ammonia, and the concentration of 14 additional amino acids. It should be noticed that among the 14 additional amino acids measured, Glycine (Gly) and Glutamate (Glu) appear to be produced (See Table 1). The specific uptake and excretion rates are obtained by linear regression of substrates and products during the growth phase.

3. METABOLIC FLUX ANALYSIS

3.1 Metabolic Network of CHO-320 cells

The metabolism of CHO cells considered in this analysis is described by the set of $n = 100$ biochemical reactions listed in Table 1. The metabolic network involves the following pathways:

- Glycolysis;
- Pentose Phosphate Pathway (PPP);
- Tricarboxylic Acid Cycle (TCA);
- the Amino Acid metabolism and Protein Synthesis;
- the Urea Cycle
- the Nucleic Acid Synthesis;

Table 1. Metabolic Reactions for the metabolism of CHO-320 cells

Flux	Reaction
v_1	$\mathbf{Glucose}^a + ATP \rightarrow G6P + ADP$
v_2	$G6P \xleftrightarrow{} {}^bF6P$
v_3	$F6P + ATP \rightarrow DHAP + G3P + ADP$
v_4	$DHAP \xleftrightarrow{} G3P$
v_5	$G3P + NAD^+ + ADP \xleftrightarrow{} 3PG + NADH + ATP$
v_6	$3PG + ADP \rightarrow Pyr + ATP$
v_7	$Pyr + NAD^+ + CoASH \rightarrow AcCoA + CO_2 + NADH$
v_8	$AcCoA + Oxal + H_2O \rightarrow Cit + CoASH$
v_9	$Cit + NAD(P)^+ \rightarrow \alpha KG + CO_2 + NAD(P)H$
v_{10}	$\alpha KG + CoASH + NAD^+ \rightarrow SucCoA + CO_2 + NADH$
v_{11}	$SucCoA + GDP + P_i \xleftrightarrow{} Succ + GTP + CoASH$
v_{12}	$Succ + FAD \xleftrightarrow{} Fum + FADH_2$
v_{13}	$Fum \xleftrightarrow{} Mal$
v_{14}	$Mal + NAD^+ \xleftrightarrow{} Oxal + NADH$
v_{15}	$Pyr + NADH \xleftrightarrow{} \mathbf{Lactate}^a + NAD^+$
v_{16}	$Pyr + Glu \xleftrightarrow{} Ala + \alpha KG$
v_{17}	$G6P + 2NADP^+ + H_2O \rightarrow Rbl5P + 2NADPH + CO_2$
v_{18}	$Rbl5P \xleftrightarrow{} R5P$
v_{19}	$Rbl5P \xleftrightarrow{} X5P$
v_{20}	$X5P + R5P \xleftrightarrow{} F6P + E4P$
v_{21}	$X5P + E4P \xleftrightarrow{} G3P + F6P$
v_{22}	$Mal + NAD(P)^+ \xleftrightarrow{} Pyr + HCO_3^- + NAD(P)H$
v_{23}	$Glu + NAD(P)^+ \xleftrightarrow{} \alpha KG + NH_4^+ + NAD(P)H$
v_{24}	$Oxal + Glu \xleftrightarrow{} Asp + \alpha KG$
v_{25}	$Gln \rightarrow Glu + NH_4^+$
v_{26}	$Thr + NAD^+ + CoASH \rightarrow Gly + NADH + AcCoA$
v_{27}	$3PG + Glu + NAD^+ \rightarrow Ser + \alpha KG + NADH$
v_{28}	$Ser \xleftrightarrow{} Gly$
v_{29}	$Gly + NAD^+ \rightarrow CO_2 + NH_4^+ + NADH$
v_{30}	$Ser \rightarrow Pyr + NH_4^+$
v_{31}	$Thr \rightarrow \alpha Kb + NH_4^+$
v_{32}	$\alpha Kb + CoASH + NAD^+ \rightarrow PropCoA + NADH + CO_2$
v_{33}	$PropCoA + HCO_3^- + ATP \rightarrow SucCoA + ADP + P_i$
v_{34}	$Trp \rightarrow Ala + 2CO_2 + \alpha Ka$
v_{35}	$Lys + 2\alpha KG + 2NADP^+ + NAD^+ + FAD^+ \rightarrow \alpha Ka + 2Glu + 2NADPH + NADH + FADH_2$
v_{36}	$\alpha Ka + CoASH + 2NAD^+ \rightarrow AcetoAcCoA + 2NADH + 2CO_2$
v_{37}	$AcetoAcCoA + CoASH \rightarrow 2AcCoA$
v_{38}	$Val + \alpha KG + CoASH + 3NAD^+ + FAD^+ \rightarrow PropCoA + Glu + 2CO_2 + 3NADH + FADH_2$
v_{39}	$Ile + \alpha KG + 2CoASH + 2NAD^+ + FAD^+ \rightarrow AcCoA + PropCoA + Glu + CO_2 + 2NADH + FADH_2$
v_{40}	$Leu + \alpha KG + CoASH + NAD^+ + HCO_3^- + ATP + FAD^+ \rightarrow AcCoA + AcetoAc + Glu + CO_2 + NADH + ADP + FADH_2$
v_{41}	$AcetoAc + SucCoA \rightarrow AcetoAcCoA + Succ$
v_{42}	$Phe + NADH \rightarrow Tyr + NAD^+$
v_{43}	$Tyr + \alpha KG \rightarrow Fum + AcetoAc + Glu + CO_2$
v_{44}	$Met + Ser + ATP \rightarrow Cys + \alpha Kb + NH_4^+ + AMP$
v_{45}	$Cys \rightarrow Pyr + NH_4^+$
v_{46}	$Asn \rightarrow Asp + NH_4^+$
v_{47}	$Arg \rightarrow Orn + Urea$
v_{48}	$Orn + \alpha KG \xleftrightarrow{} Glu\gamma SA + Glu$

[a] Extracellular measured species

[b] Chosen net direction for reversible reaction

- the Membrane Lipid Synthesis.

It has to be stressed that this metabolic network corresponds to a metabolism of growing cells. Therefore, a single flux direction has been assigned to reversible reactions according to this phase of the cell life.

Flux	Reaction
v_{49}	$Pro \rightarrow Glu\gamma SA$
v_{50}	$Glu\gamma SA + NAD(P)^+ \rightarrow Glu + NAD(P)H$
v_{51}	$His \rightarrow Glu + NH_4^+$
v_{52}	$Orn + CarbP \rightarrow Cln$
v_{53}	$Cln + Asp + ATP \rightarrow ArgSucc + AMP$
v_{54}	$ArgSucc \rightarrow Arg + Fum$
v_{55}	$0.023His + 0.053Ile + 0.091Leu + 0.059Lys + 0.023Met$ $+0.039Phe + 0.059Thr + 0.014Trp + 0.066Val + 0.051Arg$ $+0.072Gly + 0.052Pro + 0.032Tyr + 0.078Ala + 0.043Asn$ $+0.053Asp + 0.019Cys + 0.042Gln + 0.063Glu + 0.068Ser + ATP$ $+3GTP \rightarrow Protein + AMP + Pp_i + 3GDP + 3P_i$
v_{56}	$R5P + ATP \rightarrow PRPP + AMP$
v_{57}	$PRPP + 2Gln + Gly + Asp + 4ATP + CO_2 \rightarrow IMP + 2Glu$ $+Fum + 4ADP + 2H_2O$
v_{58}	$IMP + Asp + 2ATP + GTP \rightarrow ATP_{RN} + Fum + 2ADP + GDP$
v_{59}	$IMP + Gln + 3ATP + NAD^+ + 2H_2O \rightarrow GTP_{RN} + Glu + 2ADP$ $+AMP + NADH$
v_{60}	$HCO_3^- + NH_4^+ + Asp + 2ATP + NAD^+ \rightarrow Orotate + ADP$ $+NADH$
v_{61}	$Orotate + PRPP + ATP \rightarrow UTP_{RN} + CO_2 + 2ADP$
v_{62}	$UTP_{RN} + Gln + ATP \rightarrow CTP_{RN} + Glu + ADP$
v_{63}	$0.285(ATP_{RN} + UTP_{RN}) + 0.215(GTP_{RN} + CTP_{RN}) \rightarrow RNA$
v_{64}	$ATP_{RN} \rightarrow dATP$
v_{65}	$GTP_{RN} \rightarrow dGTP$
v_{66}	$CTP_{RN} \rightarrow dCTP$
v_{67}	$UTP_{RN} \rightarrow dTTP$
v_{68}	$0.285(dATP + dTTP) + 0.215(dGTP + dCTP) \rightarrow DNA$
v_{69}	$DHAP + NADH \rightarrow Glyc3P + NAD^+$
v_{70}	$Choline + 18AcCoA + Glyc3P + 23ATP + 33NADH \rightarrow PC$ $+17ADP + 6AMP + 33NAD^+$
v_{71}	$Ethanolamine + 18AcCoA + Glyc3P + 23ATP + 33NADH \rightarrow PE$ $+17ADP + 6AMP + 33NAD^+$
v_{72}	$PE + Ser \rightarrow PS + Ethanolamine$
v_{73}	$16AcCoA + Ser + Choline + 16ATP + 29NADPH \rightarrow SM + 2CO_2$ $+14ADP + 2AMP + 29NADP^+$
v_{74}	$18AcCoA + 18ATP + 14NADPH \rightarrow Cholesterol + 6CO_2 + 18ADP$ $+14NADP^+$
v_{75}	$0.5PC + 0.2PE + 0.075PS + 0.075SM + 0.15Cholesterol$ $\rightarrow MembraneLipid$
v_{76}	$0.9226Protein + 0.013RNA + 0.0052DNA$ $+0.0297MembraneLipid \rightarrow Biomass$
v_{77}	$\mathbf{Asp_{ext}}^a \rightarrow Asp$
v_{78}	$Cys_{ext} \rightarrow Cys$
v_{79}	$Gly \rightarrow \mathbf{Gly_{ext}}^a$
v_{80}	$\mathbf{Ser_{ext}}^a \rightarrow Ser$
v_{81}	$Glu \rightarrow \mathbf{Glu_{ext}}^a$
v_{82}	$\mathbf{Tyr_{ext}}^a \rightarrow Tyr$
v_{83}	$Ala \rightarrow \mathbf{Ala_{ext}}^a$
v_{84}	$\mathbf{Arg_{ext}}^a \rightarrow Arg$
v_{85}	$\mathbf{Asn_{ext}}^a \rightarrow Asn$
v_{86}	$\mathbf{Gln_{ext}}^a \rightarrow Gln$
v_{87}	$His_{ext} \rightarrow His$
v_{88}	$\mathbf{Ile_{ext}}^a \rightarrow Ile$
v_{89}	$\mathbf{Leu_{ext}}^a \rightarrow Leu$
v_{90}	$\mathbf{Lys_{ext}}^a \rightarrow Lys$
v_{91}	$\mathbf{Met_{ext}}^a \rightarrow Met$
v_{92}	$\mathbf{Phe_{ext}}^a \rightarrow Phe$
v_{93}	$Pro_{ext} \rightarrow Pro$
v_{94}	$\mathbf{Thr_{ext}}^a \rightarrow Thr$
v_{95}	$Trp_{ext} \rightarrow Trp$
v_{96}	$\mathbf{Val_{ext}}^a \rightarrow Val$
v_{97}	$Ethanolamine_{ext} \rightarrow Ethanolamine$
v_{98}	$Choline_{ext} \rightarrow Choline$
v_{99}	$NH_4^+ \rightarrow \mathbf{NH_{4,ext}^+}^a$
v_{100}	$CO_2 \rightarrow CO_{2,ext}$

Continued from previous page

3.2 MFA: principle

MFA is a methodology in metabolic engineering for the quantification of the pathway fluxes from limited experimental data. Here we consider the special case where extracellular measurements in the culture medium are the only available data. On the basis of the metabolic network, the flux distributions can be found by applying steady-state mass balances around the internal balanced metabolites. Each admissible flux distribution is represented by a vector $v = (v_1, v_2, \ldots, v_m)^T$ whose entries are the rates (or fluxes) at which the reactions proceed. The steady-state balance around the internal metabolites is expressed by the algebraic problem

$$\mathbf{N}\mathbf{v} = \mathbf{0} \qquad \mathbf{v} \geqslant \mathbf{0} \tag{1}$$

where the $m \times n$ matrix N is the stoichiometric matrix deduced from the metabolic network (m is the number of internal balanced metabolites and n the number of fluxes). In our case, the stoichiometric matrix $\mathbf{N}$ has dimensions 72×100. It is a rather sparse matrix.

An admissible flux distribution $\mathbf{v}$ must satisfy the steady state balance equation (1) and be compatible with the experimental measurements. The specific uptake and excretion rates of the measured external species are collected in a vector $\mathbf{v}_m$ and are by definition linear combinations of the unknown fluxes v_i. This is expressed as

$$\mathbf{v}_m = \mathbf{N}_m\mathbf{v} \tag{2}$$

where $\mathbf{N}_m$ is a proper given $p \times m$ full-rank matrix with p the number of available measurements.

The aim of MFA is to compute the set of admissible flux distributions $\mathbf{v}$, i.e., the set of non-negative vectors $\mathbf{v}$ that satisfy the system (1)-(2). The problem is said to be *well posed* if the solution set is not empty and if all the solutions are bounded. Otherwise, the system is said to be *ill posed*.

When the problem is well posed, the solution set is a polytope in the positive orthant and each admissible flux distribution $\mathbf{v}$ can be expressed as a convex combination of a set of non-negative basis vectors $\mathbf{f}_i$ which are the vertices of this polytope and form therefore a *unique* convex basis of the solution space. In other words, the solution set of the MFA problem is the set of admissible flux vectors defined as

$$\mathbf{v} = \sum_i \alpha_i \mathbf{f}_i, \quad \alpha_i \geqslant 0, \quad \sum_i \alpha_i = 1. \tag{3}$$

The basis vectors $\mathbf{f}_i$ are obtained by applying the software METATOOL (Pfeiffer et al., 1999) to the matrix

$$\begin{pmatrix} \mathbf{N} & \mathbf{0} \\ \mathbf{N}_m & -\mathbf{v}_m \end{pmatrix}.$$

Once the basis vectors are known, we can compute the limiting values of the flux interval for each metabolic flux:

$$v_i^{\min} \leqslant v_i \leqslant v_i^{\max}$$

$$\text{with} \quad v_i^{\min} \triangleq \min\left\{f_{ki},\ k = 1, \ldots, m\right\},$$
$$v_i^{\max} \triangleq \max\left\{f_{ki},\ k = 1, \ldots, m\right\}$$

where f_{ki} denotes the i-th element of the basis vector $\mathbf{f}_k$.

4. RESULTS

Depending on the number and type of available extracellular measurements, the system can be well- or ill-posed. If the system is well-posed, the number of basis vectors $\mathbf{f}_i$ and the size of the flux intervals will depend on the extracellular measurements that are considered. Some measurement information is critical for the determination of the flux intervals, whereas other measurements are less influential.

For this particular case, the solution set of the system has been analyzed so as to determine whether the system is well-posed or not. In this respect, if matrix $\mathbf{N}_m$ in equation 2 is defined for a set of external measurements as the one given in Section 2, consisting in 19 measurement data, the system is found to be ill-posed. This occurs due to the presence of certain elementary paths linking non-measured inputs to non-measured outputs, which gives rise to an unbounded set of solutions for the reaction rates in these routes. Extracellular species participating in all these unmeasurable paths are CO_2 and urea, meaning that by measuring one of these two species, the solution system will become constrained and thus finite.

Therefore, to be able to work with a well-posed system the set of actual measurements is complemented with an estimative value for the missing data of CO_2. In Lovrecz and Gray (1994) Fig.4(b), a value of $0.68[mmol10^{-9}cell^{-1}h^{-1}]$ is given for the CO_2 excretion rate of CHO cell cultures.

Once the system has been checked for well-posedness, the size of the intracellular flux intervals is analyzed in the light of different assumptions. For this purpose, the (now) 20 experimental measurements are first used to compute the flux intervals. Then, in a second case study, the system is further constrained with the theoretical assumption of no catabolism of Threonine (Thr). Further on, two additional estimated measurements are added, the measurements of Cysteine (Cys) and Proline (Pro). Finally, the assumption of no Thr catabolism is relaxed and the case in which we are able to measure a large number of extracellular species (i.e., the input-output fluxes) defined by the metabolic network is also assessed.

4.1 Intervals obtained with the initial (actual + one estimate) set of experimental measurements (p = 20)

We first perform the MFA with the experimental data described in Section 2 plus the estimate of CO_2. For the sake of place, the distribution intervals obtained for all the unknown metabolic fluxes cannot be presented. Alternatively, the obtained intervals for biomass and its main macromolecules production, the TCA cycle and for the metabolism of amino acids are presented in Fig. 1, Fig. 2 and Fig. 3, respectively. The intervals obtained using this set of experimental data appear delimited by dots in both figures. From these results the two main conclusions are:

1. The MFA problem produces bounded non-negative intervals for all metabolic fluxes. Many of the flux intervals are fairly narrow.
2. Some intervals include zero as a feasible solution. It is the case in particular for reactions v_{55}, v_{63}, v_{68}, v_{75} and v_{76} (see Fig. 1), which correspond to the formation of the cellular macromolecules.

Despite a zero flux is mathematically feasible, it is clear that from a biological viewpoint this is not a valid possibility during the cell growth and it would be desirable to have smaller and more realistic intervals on these species.

In the next subsection, we shall see that this issue is efficiently addressed by introducing a very mild assumption regarding the consumption of Thr.

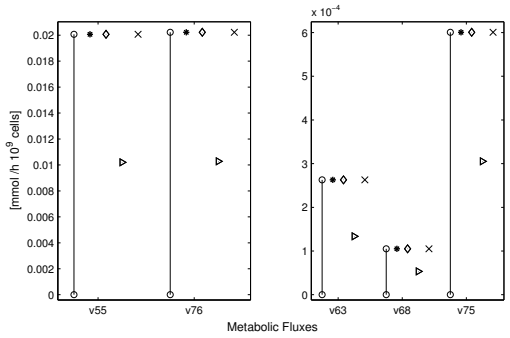

Fig. 1. Flux Distribution Intervals for macromolecules and biomass synthesis. (○): Initial set of measurements; (∗): Initial measurements + no Thr catabolism; (⋄): Additional measurements of Cys and Pro + no Thr catabolism; (▷): Additional measurements of Cys, Pro, His, Trp, Urea and Ethanolamine. (×) depict the optimal solution for the maximization of the biomass for the complete set of measurements.

4.2 Intervals obtained with the initial set of measurements considering no Thr catabolism

As an additional modelling assumption, we now assume that Thr is exclusively used for protein formation in reaction v_{55}. The underlying motivation is that the maximum possible production rate of proteins is given by the essential amino acid with the lower ratio between its uptake rate and its stoichiometric coefficient for protein synthesis. Among all essential amino acids, Thr is the amino acid with this lower ratio. This assumption implies that Thr is not used for catabolic purposes and that fluxes $v26$ and $v31$ are set to zero. The flux intervals obtained for this set of constraints are delimited by asterisks in Figures 1, 2 and 3. The following conclusions can be drawn:

1. The flux intervals of the central metabolism as well those of metabolic reactions directly connected to it, are not significantly modified (see Figs. 2 and 3).
2. Although globally underdetermined, the MFA allows to predict uniquely the specific rates of certain fluxes which are not directly measured. Additionally, the flux intervals for Nucleotide and Lipid pathways are reduced to a single value as well (results not shown).
3. As expected, the protein production rate v_{55} is maximized (This was precisely the goal of choosing the "no Threonine catabolism" constraint). But it is worth to notice that the formation rates of all other macromolecules (v_{63} for RNA, v_{68} for DNA, v_{75} for lipids) as well as the total biomass production rate v_{76} are also maximized (see Fig. 1). Above, our purpose has been to investigate the issue of reducing the range of feasible flux distributions by using experimental data

combined with one additional linear theoretical constraint. A second well-known approach to reduce the range of solutions is to use Linear Programming to compute solutions that optimize some behavioral optimization objective. Typical examples of considered optimization objectives are maximization of cell growth rate, maximization of ATP synthesis or minimization of substrate utilization. Here we see that our results give the flux distribution that maximizes the biomass production rate (or, in other words, that the cell maximizes its resources for growth and duplication) under the constraint of the data set. This is a nice example of a situation were the two approaches for reducing the range of flux distributions are completely equivalent: either to add theoretical linear constraints or to select an optimization criterion. As a matter of comparison we give also the solution obtained with the Matlab optimization function `fmincon`. Obviously, this is only one arbitrary solution in the set of all possible optimal solutions. The optimal fluxes are exactly identical to the fluxes for which there was a single well defined value in the previous solution.

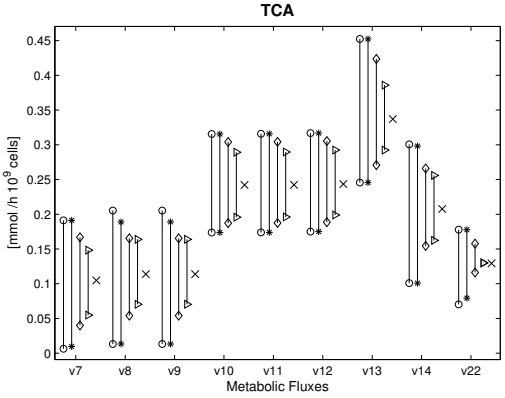

Fig. 2. Flux Distribution Intervals for the Tricarboxylic acid cycle.

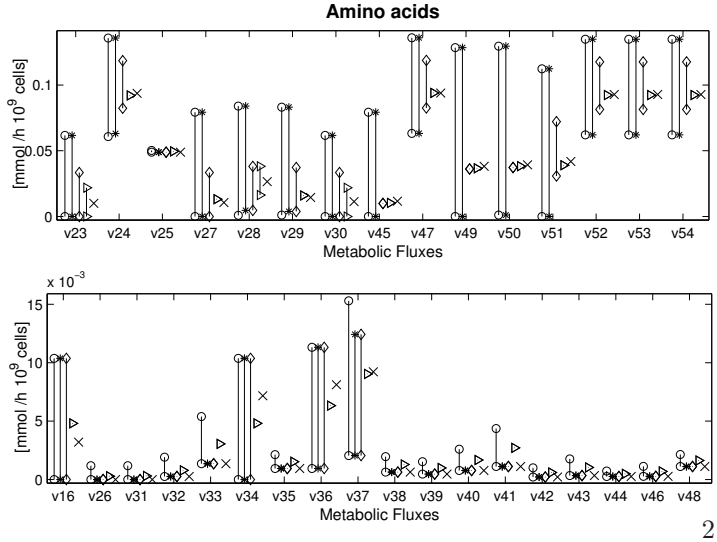

Fig. 3. Flux Distribution Intervals for the amino acid metabolism.

4.3 Intervals obtained considering no Thr catabolism and a larger set of measurements (p = 22)

From the previous results, it is clear that a part of the uncertainty is linked to the fact that measurements of some nutrients in the culture medium are missing. Hence, in order to further constrain the system, in addition to the assumption regarding Thr, we assume that additional measurements of Cys and Pro are available. To estimate the uptake/excretion rate of these two (and a few more significant) extracellular species in the culture medium, we have taken as plausible rates the central (average) values given by the flux intervals computed in Section 4.1.

The intervals for this set of measurements are delimited by diamonds. The MFA results show that:

1. Again, the flux intervals of the central metabolism are not modified. For some of the reactions connected to the central pathways the intervals become narrower than in the two previous cases. The intervals are also narrower for the amino acid metabolism where certain intervals are even reduced to a single value.
2. Zero is excluded from the solution set of a few more reactions, but not completely.

4.4 Intervals obtained with an even larger set of extracellular measurements (p = 27), but no assumption

We now consider that a larger set of measurements is available and we relax the assumption of no Thr catabolism. For this purpose, estimates for the non-measured amino acids and for other extracellular species (consumed or excreted) are estimated as described above (see Subsection 4.3). The estimated measurements are: Histidine, Tryptophan, Urea, choline and ethanolamine. All these species could possibly be measured in the culture medium in more extensive experimental studies.

In this case we have the following comments:

1. The solution gives a biomass formation rate which is smaller than the maximum value of the previous Section 4.2. The reason is that, with the chosen additional pseudo-measurements, the unmeasured amino-acids become slightly limiting. Obviously, it is natural that the solution depends on the selection of the pseudo-measurements but this is not critical with respect to our purpose here. Indeed, the key point is that, with a complete set of measurements, the flux intervals are now almost completely reduced to a single value (53 out of 74). In fact most of the obtained flux intervals (delimited by triangles) are exactly determined while others are very small intervals. Though, the pathways of the central metabolism are not significantly affected remaining as intervals.

   In the amino acid metabolism, only 3 metabolic fluxes remain as intervals (being drastically reduced) while the others are set to a unique value. The same holds for the reactions of the nucleotide and lipid pathways, where the metabolic reactions are totally determined.
2. Regarding the reaction intervals that include zero among the solutions, the number is notably reduced to 6 reactions out of 44 at the beginning. These reactions are: v_7, v_{19}, v_{20}, v_{21}, v_{23} and v_{30}.

This does not mean reversibility or null flux, since the presence of zero is due to the fact that certain reactions are in counterbalance, i.e., some reactions synthesize the same metabolite in parallel and cannot therefore be completely distinguished from the available data.

5. EXPERIMENTAL ERROR ANALYSIS

As an extension, the influence of the measurement accuracy on the numerical results is explored using Monte Carlo techniques. Herein, random errors following a normal distribution with zero mean and 5 or 10% of relative standard deviation, respectively, are considered and the solution of the steady-state mass balance equations ((1)-(2)) under positivity constraints is evaluated for a large number of occurrences of the measurement noise (e.g. ten thousands of runs). For an introduced error of 5% a distribution of errors is then inferred for the upper and lower limits of the flux intervals (an average value and confidence intervals can be evaluated). On the other hand, considering a 10% of error on the experimental measurements can lead to ill-posed systems. For 42 out of ten thousand runs the system do not have a solution, meaning that experimental data with a 10% error will still generate a well-posed system in the 99,58% of the cases.

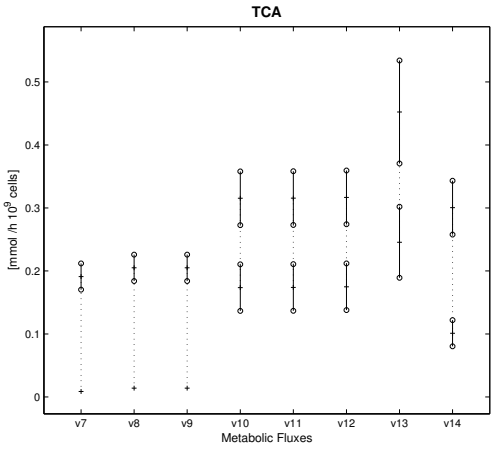

Fig. 4. Flux Distribution Intervals for the TCA cycle considering 5% measurement errors. 95% confidence intervals for the upper and lower bounds.

For a 5% measurement error the Monte Carlo study always results in a well-posed system. As an illustration, the flux intervals for the TCA cycle are depicted in Fig. 4. 95% confidence intervals are built for the upper and lower bounds, assuming a normal error distribution.

The results show that the measurement errors affect the upper limit in a rather linear way, i.e. the distribution of errors on the upper bounds is assimilable to a normal distribution with standard deviation in the same proportion as the introduced error. In contrast, some of the lower bounds are affected in a nonlinear way (due to the positivity constraints) and the distribution of errors is then no longer assimilable to a normal one (See Fig.4).

6. CONCLUSIONS

In this study, the MFA of a detailed metabolic network of $CHO - 320$ has been studied, using the classical quasi-steady state assumption and the measurements of the time evolution of a number of culture components (so as to provide an additional set of constraints on the metabolic network). With commonly available measurements in today's laboratory practice, the mass balance system remains underdetermined due to the insufficient number of measurements. However, the calculation of the solution space allows to define admissible flux ranges for each metabolic reaction.

It is of capital importance to check that the mass balance system is well-posed, as it may occur that the estimated range for the fluxes along certain elementary routes in the network are not bounded due to the absence of measurements at its extreme points (extracellular measurements).

If the system is well-posed, the size of the flux intervals can be significantly reduced either by adding a few more extracellular measurements or by considering simple assumptions, such as the assumption of no Threonine catabolism. In all cases, where the problem is well posed, the computed flux intervals are bounded and a significant number of metabolic fluxes are uniquely determined.

Finally, the effect of measurement errors on the metabolic flux analysis has been assessed using Monte Carlo techniques.

7. ACKNOWLEDGMENT

This article presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors.

REFERENCES

Antoniewicz, M.R., Kelleher, J.K., and Stephanopoulos, G. (2006). Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metabolic Engineering*, 8, 324–337.

Bonarius, H.P.J., Schmid, G., and Tramper, J. (1997). Flux analysis of underdetermined metabolic networks: the quest for the missing constraints. *TIBTECH*, 15, 308–314.

Henry, O., Perrier, M., and Kamen, A. (2005). Metabolic flux analysis of HEK-293 cells in perfusion cultures for the production of adenoviral vectors. *Metabolic Engineering*, 7, 467–476.

Lovrecz, G. and Gray, P. (1994). Use of on-line gas analysis to monitor recombinant mammalian cell cultures.

Pfeiffer, T., Sánchez-Valdenebro, I., Nuño, J., Montero, F., and Schuster, S. (1999). METATOOL: for studying metabolic networks. *Bioinformatics*, 15(3), 251–257.

Provost, A. and Bastin, G. (2004). Dynamic metabolic modelling under the balanced growth condition. *Journal of Process Control*, 14, 717–728.

Stephanopoulos, G.N., Aristidou, A.A., and Nielsen, J. (1998). *Metabolic Engineering: Principles and Methodologies.*