A detailed metabolic flux analysis of an underdetermined network of CHO cells

F. Zamorano [a], A. Vande Wouwer [a,*], G. Bastin [b]

[a] *Department of Automatic Control, University of Mons, Boulevard Dolez 31, 7000 Mons, Belgium*
[b] *Centre for Systems Engineering and Applied Mechanics, Department of Mathematical Engineering, Catholic University of Louvain, 1348 Louvain-La-Neuve, Belgium*

ARTICLE INFO

ABSTRACT

In this article the metabolic flux analysis of growing CHO-320 cells is performed for a detailed metabolic network which involves 100 reactions and embraces all the significant pathways describing the metabolism of CHO cells. The purpose is to investigate the efficiency of the flux analysis when it is based on a relatively small set of extracellular measurements that can be easily achieved in most laboratories. In this case the flux analysis problem leads to a generally underdetermined mass balance system, as data are not sufficient to uniquely define the metabolic fluxes. Our main contribution is to show that, provided the system of mass balance equations is well-posed, although it is underdetermined, very narrow intervals may be found for most fluxes. The importance of checking the well-posedness of the problem is emphasized and the influence of the number of available measurements on the accuracy of the metabolic flux intervals is systematically investigated. In all cases the computed flux intervals are bounded and a single well defined value is obtained for the formation rates of the cellular macromolecules (proteins, DNA, RNA, lipids) that are not measured. The potential gain of a simple theoretical assumption regarding the metabolism of Threonine is also discussed and compared with an optimal solution calculated by maximizing the biomass formation rate. Alternative network structures obtained by inverting the direction of reversible reactions are also considered. Finally, the results of the metabolic flux analysis are exploited to estimate the total energy production resulting from the metabolism of growing CHO-320 cells.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Metabolic flux analysis (MFA) of animal cells has been a subject of scientific interest for almost two decades. It has given rise to a large number of publications. Some significant papers concerned with MFA of animal cells (Hybridoma, CHO, MDCK and HEK) are briefly reviewed in Table 1.

In this article, we study the metabolic flux analysis (MFA) of CHO-320 cells based on a detailed metabolic network comprising 100 reactions. This network is built with the main purpose of embracing all the significant pathways describing the metabolism of CHO cells. In general, when only external measurements are used, this type of complex network leads to an underdetermined system because the external measurements are 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; 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 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 (2003, 2004) which allows the determination of intervals for the metabolic fluxes. In this study, we are interested in determining the non-negative set of solutions, i.e., non-negative flux distributions only. Convex analysis provides a way to calculate and characterize the solution set of a non-negative linear system. The set of basis vectors of the solution set is computed with the toolbox METATOOL (Pfeiffer et al., 1999), a well validated, freely available and compatible with Matlab tool. Using this computational approach, we investigate the influence of the set of available extracellular measurements on the size of the metabolic flux intervals.

* Corresponding author. Tel.: +32 65 37 41 41; fax: +32 65 37 41 36.
E-mail addresses: francisca.zamorano@umons.ac.be
(F. Zamorano), alain.vandewouwer@umons.ac.be (A.Vande Wouwer),
georges.bastin@uclouvain.be (G. Bastin).

Table 1
Overview of MFA publications.

Method	No. of reactions	Main achievements	Reference
MFA + labelling experiments	18–44	A mathematical algorithm for identifying the fluxes which should be selected for experimental measurements, based on the criterion that the calculated fluxes should have low sensitivity to experimental error in the measured fluxes. The algorithm is applied to two biological systems: *Escherichia coli* and **Hybridoma** cells.	Savinell and Palsson (1992a,b)
MFA	–	Classification of metabolic systems and the conversion rates. The set of calculable flux rates is determined based on the singular value decomposition method.	van der Heijden et al. (1994)
MFA + labelling experiments	20	Stoichiometric balance is applied to estimate intracellular fluxes and to study energy metabolism of **Hybridoma** cells. Flux estimates are validated with labelling experiments, with a good agreement.	Zupke and Stephanopoulos (1995)
MFA	34	Material balance is applied to an **Hybridoma** cell stoichiometric reaction network. The role of essential and non-essential amino acids together with the metabolism of glucose and lactate are assessed. The second part of this study deals with the energetic metabolism to determine the stoichiometric ATP production rates.	Xie and Wang (1996a,b)
MFA + theoretical constraints	22	MFA is applied to investigate the metabolism of **Hybridoma** cells under different culture conditions or stress scenarios. The results provide a good insight on the cell metabolism and how they may react in a given stress situation.	Bonarius et al. (1996, 2000)
MFA	34	MFA is used to compare different culture conditions at low glutamine concentrations to improve cellular growth of **HEK-293** cells. A relatively simple metabolic network which, together with the available measurements give place to an overdetermined system.	Nadeau et al. (2000)
MFA	41	The authors present a method based on the null-space of the stoichiometric matrix of the reactions with unknown rates to find out which reaction rates are feasible to be uniquely determined and to calculate them. An example on purple nonsulfur **bacteria** is given.	Klamt and Schuster (2002); Klamt et al. (2002)
MFA	24	Study of a metabolic model of **CHO** cell cultures based on a simple metabolic network. Convex analysis is employed to obtain only the admissible positive set of solutions of an underdetermined system.	Provost (2006); Provost and Bastin (2003, 2004); Provost et al. (2007)
MFA	40	MFA is used to determine the best conditions to perform the infection of adenovirus in **HEK-293** cells by analyzing their metabolic state by means of MFA.	Henry et al. (2005)
MFA + labelling experiments	26	Techniques for determining accurate flux intervals are introduced, based on flux sensitivities with respect to isotope measurements and measurement errors. These tools are applied to glucose metabolism in **human**.	Antoniewicz et al. (2006)
MFA	24	MFA is used to further investigate the metabolism of lactate in **CHO** cells cultured in a medium containing glucose–galactose as carbon source.	Altamirano et al. (2006)
MFA + linear optimization	22	An interval representation of fluxes is introduce to overcome uncertain or partially unknown fluxes. A method is proposed to compute the ranges of possible values for each non-calculable flux, resulting in a flux region called flux-spectrum. This method is applied in a **CHO** cell metabolic network.	Llaneras and Picó (2007)
MFA + theoretical constraints	112	MFA is used to evaluate the metabolism of **MDCK** cells cultivated in glutamine-containing and glutamine-free medium. By applying certain assumptions regarding extracellular measurements and the inactivity of intracellular fluxes, the authors get to resolve the underdetermination of the network.	Sidorenko et al. (2008); Wahl et al. (2008)

Our main contribution is to show that, if the system of mass balance equations is *well-posed*, although it is underdetermined, very narrow intervals may be found for most fluxes using extracellular measurements only. In particular, a single well defined value is obtained for the formation rates of the cellular macromolecules (proteins, DNA, RNA, lipids) that are not measured. It is of upmost importance to check for the well-posedness of the mass balance system, as the absence of measurements at some extreme points of the network may imply unbounded range of fluxes along certain elementary routes.

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. Threonine is the amino acid with the lowest ratio between its uptake rate and its stoichiometric coefficient for protein synthesis. Thus, assuming that Threonine is exclusively used for protein formation, the protein production rate is maximized. The MFA results that we get under this assumption show that we have not only a maximization of the protein production but also a maximization of the overall biomass

production (involving nucleotides and lipids). Thus, in our application to CHO cells, MFA under the constraint of no-Threonine catabolism appears to be equivalent to MFA under the constraint of biomass maximization. As it has been reported in the literature (see, for example, Boghigian et al., 2009), when MFA is applied to exponentially growing cells, it is reasonable to assume the cellular system maximizes its resources to make biomass.

The validity of the chosen network structure is then analyzed by considering alternative network configurations, i.e., the direction of certain fluxes is systematically varied in order to test the feasibility of the chosen structure. Finally, the results obtained from the MFA are exploited to estimate the total energy production resulting from the metabolism of growing CHO-320 cells.

In general, the literature focuses on simplified metabolic networks, where the set of mass balance equations is either determined or overdetermined. Otherwise, metabolic flux determination in complex biological systems is mostly based on the use of isotope labelling and the addition of theoretical constraints. In contrast to the studies listed in Table 1, this article is concerned with the analysis of a more detailed metabolic network (100 metabolic reac-
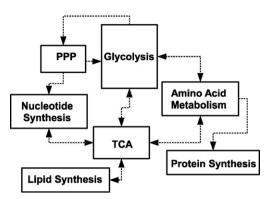
Fig. 1. Schematic representation of the metabolic network for CHO-320 cells.

tions) and the investigation of the information that can be extracted considering conventional measurement methods only (i.e., excluding isotopic labelling techniques). No *a priori* assumptions on the inactivity of intracellular fluxes are considered neither but one regarding the Threonine (Thr) metabolism.

This paper is organized as follows. Section 2 describes how a detailed metabolic network of CHO cells has been built. The set of experimental data is briefly described in Section 3. Section 4, is devoted to the principles of MFA methodology and to explain how the determination of the flux intervals is done. In Section 5 the numerical results, obtained under different scenarios, are presented. In Section 6, alternative network structures obtained by inverting the direction of reversible reactions are considered. The energetic balance of cofactors performed over the metabolic network of CHO-320 cells is explained in Section 7. Finally, Section 8 draws the main conclusions of this work.

2. Metabolic network description

The metabolism of CHO cells considered in this paper is represented by the set of $n = 100$ biochemical reactions listed in Table 2. The metabolic network involves the following pathways (also schematically represented in Fig. 1).

- Glycolysis,
- Pentose Phosphate Pathway (*PPP*),
- Tricarboxylic Acid Cycle (*TCA*),
- the Amino Acid metabolism and Protein Synthesis,
- the Urea Cycle,
- the Nucleic Acid Synthesis,
- the Membrane Lipid Synthesis,
- Biomass Formation.

It has to be stressed that this metabolic network corresponds to a metabolism of growing cells. Therefore, a single flux direction is assigned to reversible reactions according to this phase of the cell life. The choice on the direction of the net flux for some possibly reversible reactions will be assessed further on in Section 6.

The goal of this section is to give the main justifications and motivations that underlie the set-up of this metabolic network.

2.1. Central metabolism

Central metabolism involves Glycolysis, *TCA* and *PPP* pathways. In this study, the considered central metabolism is the usual metabolism of strictly aerobic eukaryotic organisms. With the exception of a few reactions which are specific to each type of mammalian cell, the pathways of the central metabolism herein described are similar to those described by numerous works dealing with MFA of animal cells, like Hybridoma (Bonarius et al., 1996,

2000; Savinell and Palsson, 1992a,b; Xie and Wang, 1996a,b; Zupke and Stephanopoulos, 1995), HEK-293 (Henry et al., 2005; Nadeau et al., 2000), MDCK (Sidorenko et al., 2008; Wahl et al., 2008) and CHO cells (Altamirano et al., 2006; Llaneras and Picó, 2007; Provost, 2006; Provost and Bastin, 2003, 2004; Provost et al., 2007).

Glucose and glutamine are the main carbon and energy sources, the latter also serving as the primary nitrogen source. Major products of glucose and glutamine metabolism are biomass, secreted proteins, energy in the form of *ATP*, reducing power for biosynthesis, carbon dioxide, and the waste products lactate and ammonia.

2.2. Amino acid metabolism and protein synthesis

In mammalian cells, essential amino acids cannot be synthesized and must therefore be provided in the culture medium. Accordingly, only catabolic pathways are considered for essential amino acids. In contrast, for non-essential amino acids, both anabolic and catabolic pathways are taken into account.

Considering that the pathways of amino acid catabolism are quite similar in most organisms, all catabolic and biosynthetic reactions of amino acids have been taken from references KEGG (2008) and Nelson and Cox (2005).

It has been reported in Bonarius et al. (1996) that mammalian tissue is ureotelic, meaning that the excess NH_3 is converted into urea and then excreted. Accordingly, small amounts of urea can be detected during CHO-320 cell cultures. Hence, the Urea Cycle has been included in the network (see Table 2). Additionally, this cell line has the particularity of being auxotrophic with respect to Proline (Pro), and thus cannot synthesize it and relies on its external supply for growth (Kao and Puck, 1967). Therefore, only the catabolic phase of the metabolism of Pro is taken into account.

The reaction v_{55} describing the synthesis of proteins is taken from Provost (2006), where an average composition of proteins for eukaryotic cells is presented and used to simulate the protein synthesis of CHO-320.

The metabolic reaction v_{76} describing the biomass production of CHO-320 cells (see Table 2) is based on the cellular composition of a recombinant t-PA producing cell line (CHO TF 70TR) used in Altamirano et al. (2006).

2.3. Nucleotide metabolism

For nucleotide synthesis, and in turn *DNA* and *RNA* synthesis, we consider only the *de novo* pathways (i.e., they are synthesized from their main precursor *Ribose*5 − *Phosphate*).

In order to simulate nucleic acid synthesis as two simplified reactions (in a similar way as for proteins), average percentages of nucleotide composition have been considered. In Seoka (1988) and Sueoka (1961, 1962) an average composition of nucleic acids is given, at different guanine–cytosine base concentrations for several cell types. On this basis, two overall reactions for both *RNA* and *DNA* synthesis, can be established (see reactions v_{69} and v_{75} in Table 2).

The subindex *RN* used in these sets of reactions stands for Ribonucleotides, and is used to differentiate nucleotides as biomass constituents (making up RNA) from nucleotides as energetic molecules (Section 7).

2.4. Lipid metabolism

Among all different kinds of lipids, we only consider those which play a structural role as components of cellular membranes: Glycerophospholipids as phosphatidylcholine (PC), phosphatidylethanolamine (PE) and phosphatidylserine (PS); Sphingolipids as shpingomyelin (SM); and Sterols as cholesterol.

Table 2
Metabolic reactions for the metabolism of CHO-320 cells.

Flux	Reaction
	Glycolysis
v_1	**Glucose**[a] $+ ATP \rightarrow G6P + ADP$
v_2	$G6P \overset{\leftrightarrow}{\rightarrow}$ [b]$F6P$
v_3	$F6P + ATP \rightarrow DHAP + G3P + ADP$
v_4	$DHAP \overset{\leftrightarrow}{\rightarrow} G3P$
v_5	$G3P + NAD^+ + ADP \overset{\leftrightarrow}{\rightarrow} 3PG + NADH + ATP$
v_6	$3PG + ADP \rightarrow Pyr + ATP$
	Tricarboxylic acid cycle
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}	$SucCoA + GDP + P_i \overset{\leftrightarrow}{\rightarrow} Succ + GTP + CoASH$
v_{12}	$Succ + FAD \overset{\leftrightarrow}{\rightarrow} Fum + FADH_2$
v_{13}	$Fum \overset{\leftrightarrow}{\rightarrow} Mal$
v_{14}	$Mal + NAD^+ \overset{\leftrightarrow}{\rightarrow} Oxal + NADH$
	Pyruvate fates
v_{15}	$Pyr + NADH \overset{\leftrightarrow}{\rightarrow}$ **Lactate** [a]$+ NAD^+$
v_{16}	$Pyr + Glu \overset{\leftrightarrow}{\rightarrow} Ala + \alpha KG$
	Pentose phosphate pathway
v_{17}	$G6P + 2NADP^+ + H_2O \rightarrow Rbl5P + 2NADPH + CO_2$
v_{18}	$Rbl5P \overset{\leftrightarrow}{\rightarrow} R5P$
v_{19}	$Rbl5P \overset{\leftrightarrow}{\rightarrow} X5P$
v_{20}	$X5P + R5P \overset{\leftrightarrow}{\rightarrow} F6P + E4P$
v_{21}	$X5P + E4P \overset{\leftrightarrow}{\rightarrow} G3P + F6P$
	Anaplerotic reaction
v_{22}	$Mal + NAD(P)^+ \overset{\leftrightarrow}{\rightarrow} Pyr + HCO_3^- + NAD(P)H$
	Amino acid metabolism
v_{23}	$Glu + NAD(P)^+ \overset{\leftrightarrow}{\rightarrow} \alpha KG + NH_4^+ + NAD(P)H$
v_{24}	$Oxal + Glu \overset{\leftrightarrow}{\rightarrow} Asp + \alpha KG$
v_{25}	$Gln \rightarrow Glu + NH_4^+$
v_{26}	$Thr + NAD^+ + CoASH \rightarrow Gly + NADH + AcCoA$
v_{27}	$Ser \overset{\leftrightarrow}{\rightarrow} Gly$
v_{28}	$3PG + Glu + NAD^+ \rightarrow Ser + \alpha KG + NADH$
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}	$Lys + 2\alpha KG + 3NAD(P)^+ + FAD^+ \rightarrow \alpha Ka + 2Glu + 3NADPH + FADH_2$
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}	$Asn \overset{\leftrightarrow}{\rightarrow} Asp + NH_4^+$
v_{47}	$Arg \rightarrow Orn + Urea$
v_{48}	$Orn + \alpha KG \overset{\leftrightarrow}{\rightarrow} Glu\gamma SA + Glu$
v_{49}	$Pro \rightarrow Glu\gamma SA$
v_{50}	$Glu\gamma SA + NAD(P)^+ \rightarrow Glu + NAD(P)H$
v_{51}	$His \rightarrow Glu + NH_4^+$
	Urea cycle
v_{52}	$Orn + CarbP \rightarrow Cln$
v_{53}	$Cln + Asp + ATP \rightarrow ArgSucc + AMP$
v_{54}	$ArgSucc \rightarrow Arg + Fum$
	Protein synthesis
v_{55}	$0.023His + 0.053Ile + 0.091Leu + 0.059Lys + 0.023Met + 0.039Phe + 0.059Thr + 0.014Trp + 0.066Val + 0.051Arg$ $+ 0.072Gly + 0.052Pro + 0.032Tyr + 0.078Ala + 0.043Asn + 0.053Asp + 0.019Cys + 0.042Gln + 0.063Glu + 0.068Ser + ATP$ $+ 3GTP \rightarrow Protein + AMP + Pp_i + 3GDP + 3P_i$
	Nucleotide synthesis
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}	$HCO_3^- + NH_4^+ + Asp + 2ATP + NAD^+ \rightarrow Orotate + ADP + NADH$

Table 2 (*Continued*)

Flux	Reaction
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}	$0.285(dATP + dTTP) + 0.215(dGTP + dCTP) \rightarrow DNA$
	Lipid synthesis
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}	$0.5PC + 0.2PE + 0.075PS + 0.075SM + 0.15Cholesterol \rightarrow MembraneLipid$
	Biomass formation
v_{76}	$0.9226Protein + 0.013RNA + 0.0052DNA + 0.0297MembraneLipid \rightarrow Biomass$
	Transport reactions
v_{77}	**Asp_{ext}** [a] $\rightarrow Asp$
v_{78}	$Cys_{ext} \rightarrow Cys$
v_{79}	$Gly \rightarrow$ **Gly_{ext}** [a]
v_{80}	**Ser_{ext}** [a] $\rightarrow Ser$
v_{81}	$Glu \rightarrow$ **Glu_{ext}** [a]
v_{82}	**Tyr_{ext}** [a] $\rightarrow Tyr$
v_{83}	$Ala \rightarrow$ **Ala_{ext}** [a]
v_{84}	**Arg_{ext}** [a] $\rightarrow Arg$
v_{85}	**Asn_{ext}** [a] $\rightarrow Asn$
v_{86}	**Gln_{ext}** [a] $\rightarrow Gln$
v_{87}	$His_{ext} \rightarrow His$
v_{88}	**Ile_{ext}** [a] $\rightarrow Ile$
v_{89}	**Leu_{ext}** [a] $\rightarrow Leu$
v_{90}	**Lys_{ext}** [a] $\rightarrow Lys$
v_{91}	**Met_{ext}** [a] $\rightarrow Met$
v_{92}	**Phe_{ext}** [a] $\rightarrow Phe$
v_{93}	$Pro_{ext} \rightarrow Pro$
v_{94}	**Thr_{ext}** [a] $\rightarrow Thr$
v_{95}	$Trp_{ext} \rightarrow Trp$
v_{96}	**Val_{ext}** [a] $\rightarrow Val$
v_{97}	$Ethanolamine_{ext} \rightarrow Ethanolamine$
v_{98}	$Choline_{ext} \rightarrow Choline$
v_{99}	$NH_4^+ \rightarrow$ **$NH_{4,ext}^+$** [a]
v_{100}	$CO_2 \rightarrow CO_{2,ext}$

[a] Extracellular measured species.
[b] Chosen net direction for reversible reaction.

The metabolism of lipids (reactions v_{69} to v_{75} accounted in this study, assumes that the cellular membrane contains particular percentages of PC, PE, PS, SM and cholesterol which are established on the basis of Hu (2004) and Nelson and Cox (2005) where usual percentages of membrane phospholipids are specified.

3. Experimental data

The experimental data come from CHO-320 cell cultures used in Provost (2006) and Provost and Bastin (2004). The experiments were performed by the 'Unité de Biochimie', Université Catholique de Louvain and have been kindly provided by Professor Y.J. Schneider. 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. Cultures were settled in a working volume of 25 mL in shake-flasks and incubated at 100 rpm in a CO_2 incubator at 37 °C in an atmosphere of 5% CO_2 in air. The cultures were inoculated to reach an initial concentration of 0.3×10^9 cells/L. Cultures were kept for 190 h, with an exponential growth phase of 80 h approximately. These 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, along with the evolution of the biomass inside the bioreactor during the growth phase. It should be noticed that among the 14 additional amino acids measured, Glycine (Gly) and Glutamate (Glu) appear to be produced. The specific uptake and excretion rates are obtained by linear regression of substrates and products during the growth phase. The estimated rates are given in Table 3 and will serve as the data for our MFA.

4. Metabolic flux analysis: principle

The metabolism of CHO cells considered in this analysis is described by the set of $n = 100$ biochemical reactions listed in Table 2. The metabolic network involves the metabolites and biochemical species listed in Appendix A, which are divided in two major groups:

(1) The group of $m = 72$ internal balanced metabolites listed in Table A.2.
(2) The group of unbalanced species which is further divided in two subgroups:

Table 3
Extracellular measurements in mmol $h^{-1} 10^9 cell^{-1}$.

Substrates	Specific uptake rates	Products	Specific excretion rates
Glucose	−0.1871	Lactate	0.3445
Glutamine	−5.0246e^{-2}	NH_4^+	4.5712e^{-2}
Arginine	−2.1417e^{-3}	Glycine	2.2295e^{-3}
Asparagine	−1.1278e^{-3}	Alanine	8.8100e^{-3}
Aspartate	−3.1785e^{-4}	Glutamate	9.5475e^{-4}
Isoleucine	−1.5278e^{-3}		
Leucine	−2.6013e^{-3}		
Lysine	−2.1245e^{-3}		
Methionine	−7.2375e^{-4}		
Phenylalanine	−9.9808e^{-4}		
Serine	−9.2342e^{-4}		
Threonine	−1.1842e^{-3}		
Tyrosine	−7.6104e^{-4}		
Valine	−1.9561e^{-3}		

(a) The intracellular energetic cofactors.
(b) The 26 extracellular metabolites present in the culture medium which are either nutrients or products of excretion (Table A.1).

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 (see Section 3). On the basis of the metabolic network, the flux distributions are 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

$$\mathbf{N}\mathbf{v} = \mathbf{0} \quad \mathbf{v} \geq \mathbf{0} \tag{1}$$

where the $m \times n$ matrix $\mathbf{N}$ 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. A schematic representation showing the locations of the non-zero entries is given in Fig. 2.

An admissible flux distribution $\mathbf{v}$ must satisfy the steady state balance Eq. 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
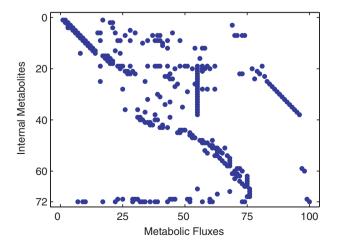
$$\mathbf{v}_m = \mathbf{N}_m \mathbf{v} \tag{2}$$

where $\mathbf{N}_m$ is a proper given $p \times n$ 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 systems (1) and (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

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

The basis vectors $\mathbf{f}_i$ are obtained by applying the software META-TOOL (Pfeiffer et al., 1999; Schuster 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:

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

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

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

Calculating the range of possible values for each metabolic flux is relevant for underdetermined systems which do not have a unique solution. The computation of flux intervals provides a quantitative and qualitative information about the metabolic state of the cell.

5. Metabolic flux analysis: results

We perform a flux analysis for CHO cells on the basis of the underlying metabolic network presented in Table 2.

The first purpose is to characterize the feasible set of solutions using extracellular measurements only and as few additional constraints as possible. 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 measurements are 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 Eq. (2) is defined for a set of external measurements as the one given in Section 3, 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 well posed.

Therefore, in order to have a well-posed system the set of actual measurements is complemented with a measurement of CO_2 taken from Lovrecz and Gray (1994, Fig.4(b)) where a value

Fig. 2. Cartographic representation of matrix $\mathbf{N}$. Dots represent the non-zero entries.

Table 4
Estimates of the uptake or excretion rates of the missing extracellular species in $mmol\,h^{-1}10^9cell^{-1}$.

Extracellular specie	Specific rate
Cysteine	$-9.98e^{-3}$
Proline	$-3.718e^{-2}$
Histidine	$-3.923e^{-2}$
Tryptophan	$-4.946e^{-3}$
Urea	$9.399e^{-2}$
Ethanolamine	$-6.1086e^{-5}$

Fig. 4. Flux distribution intervals for the central metabolism: Glycolysis, *TCA* and *PPP*.

of $0.68\,mmol\,10^{-9}\,cell^{-1}\,h^{-1}$ 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 this section under different assumptions that are successively considered. For this purpose, the 20 experimental measurements are first used to compute the flux intervals. Although most of the obtained intervals are already fairly narrow, it results from this first run that the intervals are unsatisfactory for the cellular macromolecules. Therefore, in a second case study, the system is further constrained with the theoretical assumption of no catabolism of Threonine (Thr), i.e., that Threonine is exclusively used for protein synthesis. With this single additional constraint, we get satisfactory results for the flux analysis with, in particular, a single well defined value for the formation rates of the cellular macromolecules (proteins, DNA, RNA, lipids). We also show that this result is optimal for the criterion of maximizing the biomass production rate.

The second purpose is to assess the impact of the missing extracellular measurements on the quality of the flux analysis. Thus further on, the set of actual measurements is complemented with "pseudo-measurements" for the missing data (Table 4). Firstly, two additional estimated measurements are added, the measurements of Cysteine (Cys) and Proline (Pro) and afterwards, 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. The results show that, with a larger set of extracellular data, the flux analysis can be significantly improved with most estimated intervals (53 out of 74) reduced to a single value and without any additional theoretical constraint being needed.

5.1. Intervals obtained with the initial set of experimental measurements (p = 20)

We first perform the MFA with the experimental data of Table 3 plus the CO_2 measurement. The obtained intervals are presented
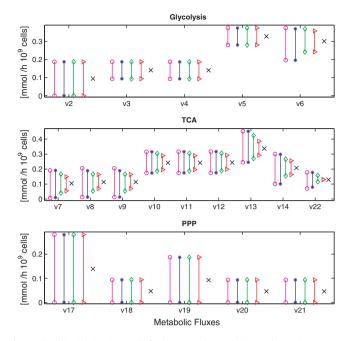
(delimited by dots) in Fig. 3 for biomass and its main macromolecules, Fig. 4 for the central metabolism, Fig. 5 for amino acids, Fig. 6 for nucleotides and Fig. 7 for lipids. From these results the two main conclusions are:

(1) The MFA problem is well-posed and 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. 3), which correspond to the formation of the cellular macromolecules (proteins, RNA, DNA, lipids and biomass).

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 for 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 Threonine.
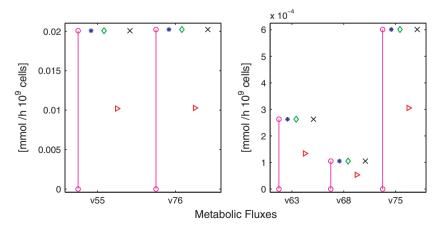

Fig. 3. Flux distribution intervals for macromolecules and biomass synthesis. (○) Initial set of measurements; (∗) initial measurements + no Thr catabolism; (◇) additional measurements of Cys and Pro + no Thr catabolism; (▷) additional measurements of Cys, Pro, His, Trp, Urea and Ethanolamine; (×) optimal solution for the maximization of the biomass from the initial set of measurements.
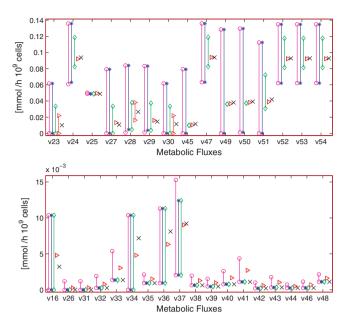
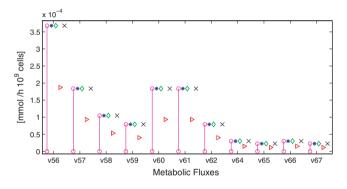**Fig. 5.** Flux distribution intervals for the amino acid metabolism.

Fig. 6. Flux distribution intervals for the nucleotide metabolism.

5.2. Intervals obtained with the initial set of measurements under the assumption of no Threonine catabolism

As an additional modelling assumption, we now assume that Threonine is exclusively used for protein formation in reaction v_{55}. This assumption implies that Threonine is not used for catabolic purposes and that fluxes v_{26} and v_{31} (Table 2) are set to zero. There are two main reasons that make this a well motivated assumption:

(1) The maximum possible production rate of proteins is given by the essential amino acid with the lowest ratio between its uptake rate and its stoichiometric coefficient for protein synthesis. Among all essential amino acids, Threonine is precisely the amino acid with this lower ratio (i.e., Threonine is the most limiting amino acid).

(2) The intervals for v_{26} and v_{31} in the previous results include effectively zero as a possible solution and are very small in comparison to the other amino-acid intervals that also include zero (v_{21}, v_{23}, v_{27}, v_{30}, v_{34}, v_{45}, v_{49}, v_{51}).

The flux intervals obtained in this case are delimited by stars in all figures. The following conclusions can be drawn:

(1) The flux intervals of the central metabolism as well as those of metabolic reactions directly connected to it, are not significantly modified (see Figs. 4 and 5).

(2) For the production of the cellular macromolecules (proteins, RNA, DNA and lipids) the intervals are now reduced to a single plausible value (see Fig. 3). In other words, with the initial set of data and the additional assumption that Threonine is an essential amino acid exclusively used for protein formation, the MFA (although globally underdetermined) allows to predict uniquely the specific production rates of these macromolecules which are not directly measured. Additionally, the flux intervals obtained for the pathways of nucleotides and lipids are reduced to a single value as well (see Figs. 6 and 7).

(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. 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 (or even Quadratic) Programming. In this approach, the goal is to compute solutions that optimize some behavioral optimization objective. The most typical examples of considered optimization objectives are maximization of cell growth rate, maximization of ATP synthesis or minimization of substrate utilization (see for instance Altamirano et al., 2006 and the references therein). 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 set of data from Table 3. This is a nice example of a situation were the two approaches for reducing the range of flux distributions are completely equivalent:
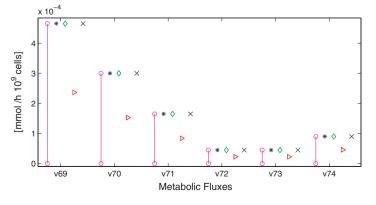

Fig. 7. Flux distribution intervals for the lipid metabolism.

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 (including obviously the biomass production rate v_{76}). For the other reactions, it is interesting to notice that the optimal fluxes are located almost at the center of the concerned intervals.

5.3. Intervals obtained with 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 5.1. The estimated rates of Cys and Pro are presented in Table 4.

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.

We see here that in spite of the addition of further constraints the size of the flux intervals of the central metabolism is not reduced. This is explained by the fact that Glycolysis and *PPP* are set in parallel, and thus are not distinguishable from extracellular measurements only. The assimilation of *G6P* could occur in the Glycolysis or in the *PPP* indistinctly, and thus their flux intervals are in counterbalance.

5.4. Intervals that could be obtained with an even larger set of extra-cellular measurements (p = 26) but no assumption

We now consider that a larger set of measurements is available and we relax the assumption of no Thr catabolism so as to assess the impact on the size of the flux intervals and on the quality of the flux analysis. For this purpose, the previous set of measurements is complemented with four additional "pseudo-measurements" for Histidine, Tryptophan, Urea and Ethanolamine (see Table 4). All these species could possibly be measured in the culture medium in more extensive experimental studies. These pseudo-measurements are just arbitrarily taken from an admissible flux distribution located in the middle of the solution polytope obtained in Section 5.1 as described above (see Section 5.3)

In this case the MFA is again well-posed and we have the following conclusions:

(1) In this case, the solution gives a biomass formation rate which is smaller than the maximum value of the previous Section 5.2. 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

Table 5
Number of basis vectors defining the space of solutions.

Case study	f_i
Initial set of measurements	144
Initial set of measurements + no Thr catabolism	32
Additional measurements of Cys and Pro + no Thr catabolism	12
Additional measurements of Cys, Pro, His, Trp, Urea and Ethanolamine	4

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 fluxes are exactly determined while others remain as very small intervals with a length which is less than 10% of the maximum value. 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 fluxes are now totally determined.

(2) Regarding the flux 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: Pyr + $NAD^+ \rightarrow$ AcCoA + CO_2 + NADH
- v_{19}: Rbl5P $\overset{\leftrightarrow}{\rightarrow}$ X5P
- v_{20}: X5P + R5P $\overset{\leftrightarrow}{\rightarrow}$ F6P + E4P
- v_{21}: X5P + E4P $\overset{\leftrightarrow}{\rightarrow}$ G3P + F6P
- v_{23}: Glu + $NAD(P)^+ \overset{\leftrightarrow}{\rightarrow} \alpha KG + NH_4^+$ + NAD(P)H
- v_{30}: Ser $\rightarrow$ Pyr + NH_4^+

The presence of zero within the flux intervals does not necessarily mean that these reactions could occur in the reverse direction. This condition can also be explained by the fact that certain reactions are in counterbalance. Similarly as it occurs for the central metabolism, some reactions occur in parallel in certain paths synthesizing the same intermediary metabolite(s) and cannot therefore be completely distinguished from the available data.

An interesting issue concerns the number of basis vectors computed for each one of the analyzed systems. The number of basis vectors calculated decreases considerably with each additional constraint applied. For instance, from 114 basis vectors obtained for the initial mass balance system analyzed (Section 5.1) by applying the simple theoretical assumption of no catabolism of Threonine, the number of basis vectors is notably reduced to 32 vectors. In the last case analyzed (Section 5.4) with the largest set of measurements, the set of solutions is defined by only 4 basis vectors, meaning that the space of solution is now quite restricted. The number of basis vectors calculated for each case study is presented in Table 5.

6. Testing the direction of possibly reversible reactions

From the previous section, it appears that some computed flux intervals include zero as a feasible solution, implying that the corresponding reactions could possibly operate in the reverse direction. In this section, we analyze this issue from a metabolic structure viewpoint.

The network describing the metabolism of CHO-320 cells involves both irreversible and reversible reactions. Up to now we have considered that the net steady-state direction during the exponential growth is known beforehand and can therefore be imposed in the statement of the MFA problem. For instance, the reversible reactions of Glycolysis and *PPP* should run in the direction which allows glucose assimilation and nucleic acid synthesis, respectively, to actually achieve cellular growth.
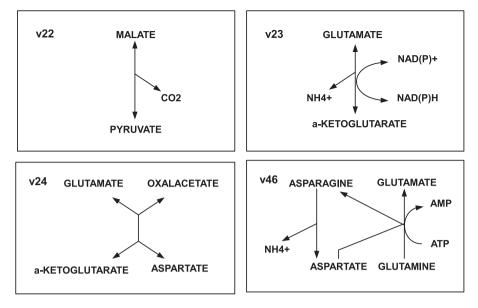
Fig. 8. Reversible reactions considered in the analysis.

Nevertheless, there are some reversible reactions whose net direction cannot be guessed in advance only on the basis of the qualitative metabolic behaviour of the cells. It is the case for the four reactions in Fig. 8, which can run in both directions depending on the need for some metabolites in particular reactions or pathways.

The MFA presented in the previous subsections has been made for a particular choice on the net direction of these four reversible reactions (as they appear in Table 2). Nevertheless, other network configurations might be obtained by alternating the net direction of reactions v_{22}, v_{23}, v_{24} and v_{46}. Hence, we shall test whether a solution set to the flux analysis exists (i.e., if the system is *well posed*) when these reactions are reversed. It must be stressed that this analysis is performed using only the initial data set in Table 3.

Among the 16 possible network structures that can be obtained by changing the direction of the above-mentioned reactions, only two of them have a non-empty solution set, i.e., only two network configurations give rise to a *well posed* problem, admitting a solution which satisfies the constraints imposed by the extracellular measurements. One of them is obviously the configuration considered to calculate the solution set of the previous subsections. The second admissible configuration is characterized by the following four reactions:

$v_{22} : Malate \rightarrow Pyruvate + CO_2$
$v_{23\,rev} : \alpha KG + NH_4^+ + NAD(P)H \rightarrow Glu + NAD(P)^+$
$v_{24} : Oxaloacetate + Glu \rightarrow Asp + \alpha KG$
$v_{46} : Asn \rightarrow Asp + NH_4^+$

This flux distribution implies the occurrence of reaction v_{23} in the reverse direction, from α-ketoglutarate to glutamate, which has generally been reported to be feasible in cultures under high ammonia concentrations. Normally, this reaction produces α-ketoglutarate from glutamate, as a second step of the metabolic pathways for glutamine degradation Schneider et al. (1996). Also in other kind of mammalian cells, specifically Hybridoma cell cultures, it has been demonstrated that under ammonia-stress conditions, the reaction catalyzed by glutamate dehydrogenase (reaction v_{25}) goes in the reverse direction, while control cells transform glutamate in α-ketoglutarate and ammonia (Bonarius et al., 1998). In our CHO cell culture, ammonia is constantly produced and accumulated, but its concentration during the growth phase is probably not sufficient to stimulate the shift of direction in v_{23}, even if it is mathematically possible. In Bonarius et al. (1998) the ammonia-stress condition is given by 10 mM of ammonia, while, in our culture at the end of the growth phase the concentration only reaches 5 mM.

In summary, it is seen that by systematically considering all the possible configurations, the network structure allows only two feasible distributions in agreement with the experimental observations. Otherwise, an *ill posed* problem emerges, for which no solution exists. We believe that we can interpret the results of this analysis as a strong evidence of the consistency of the chosen metabolic network and the results obtained in Section 5.

7. Balance of energetic co-factors

As it is clearly documented in the literature (Bonarius et al., 1996, 1997; Henry et al., 2005; Stephanopoulos et al., 1998; Xie and Wang, 1994, 1996b), energetic cofactors *ATP*, $NAD(P)H$ and $FADH_2$ cannot be considered as internal balanced metabolites. The reason is that an important basic function of the metabolism is also to provide energy for mechanisms not represented by the network, like for instance the turnover of macromolecules and other so-called futile cycles.

Here, we use the optimal solution obtained in Section 5.2 to estimate the energy production.

Table 6
ATP production rate due to *ATP*, $NAD(P)H$ and $FADH_2$, units in mmol/h 10^9 cell.

Energetic cofactor	Production		Consumption		ATP net production
	Production rate	ATP equivalents	Consumption rate	ATP equivalents	
ATP/GTP	0.8695	0.8695	0.7041	0.7041	0.1654
NAD(P)H	1.5001	4.5004	0.3631	1.0894	3.4110
FADH$_2$	0.2463	0.4926	–	–	0.4926
		ATP_{TOTAL}			4.0691

Assuming that 1 mol of *NAD(P)H* yields 3 mol of *ATP* in the respiratory chain (P/O), 1 mol of $FADH_2$ yields 2 mol of *ATP*, and 1 mol of *GTP* (or eventually *CTP*) yields 1 mol of *ATP* (Henry et al., 2005; Schneider et al., 1996; Xie and Wang, 1994), the equivalent ATP production rates given in Table 6 are obtained.

As compared with the results found in the literature (Henry et al., 2005; Xie and Wang, 1996b), the net ATP production through our metabolic network for CHO cells appears to be far more important than the ATP production for Hybridoma cells.

However, the general qualitative trend is similar (but with a larger magnitude) to that observed in Xie and Wang (1996b) for Hybridomas. For instance, NAD(P)H is the main source of energy, around 84% of the total ATP production. It also appears that the *TCA* cycle provides the major part of the energy, around 76%, a percentage close to that given in Xie and Wang (1996b) where the *TCA* cycle provides 51–68% of the total ATP production for Hybridoma cells. Similar results were obtained in Henry et al. (2005) also for Hybridomas.

8. 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 under the constraint of the measurements of the time evolution of a number of culture components. With commonly available measurements in today's laboratory practice, the mass balance system remains underdetermined due to an insufficient number of measurements. The analysis of the solution space of this underdetermined system allows to define admissible flux ranges for each metabolic reaction in the network, giving both a qualitative and quantitative idea of the metabolic state of the cells.

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 their extreme points.

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. Nevertheless, the size of certain flux intervals cannot be reduced due to the existence of parallel pathways with fluxes in counterbalance.

A unique flux distribution maximizing the biomass production rate has been calculated for the initial set of measurements which appears to be equivalent to the solution space obtained when considering that Threonine is only used for protein synthesis. The optimal fluxes are exactly identical to the uniquely determined fluxes obtained under the assumption of no catabolism of Threonine. This is a nice example of two equivalent approaches for reducing the range of flux distributions: either to add theoretical linear constraints or to select an optimization criterion.

Acknowledgment

Appendix A. List of biochemical species

See Tables A.1 and A.2.

Table A.1
Extracellular species in matrix $\mathbf{N}_m$.

Row occupied in matrix $\mathbf{N}_m(p \times n)$	Extracellular metabolite	Row occupied in matrix $\mathbf{N}_m(p \times n)$	Extracellular metabolite
1	Glucose	14	Methionine$_{ext}$
2	Glutamine$_{ext}$	15	Lactate
3	Serine$_{ext}$	16	Alanine$_{ext}$
4	Asparagine$_{ext}$	17	Glycine$_{ext}$
5	Aspartate$_{ext}$	18	Glutamate$_{ext}$
6	Arginine$_{ext}$	19	$NH_{4\ ext}^+$
7	Tyrosine$_{ext}$	20	$CO_{2\,ext}$
8	Threonine$_{ext}$	21	Proline$_{ext}$
9	Lysine$_{ext}$	22	Cysteine$_{ext}$
10	Valine$_{ext}$	23	Tryptophan$_{ext}$
11	Isoleucine$_{ext}$	24	Histidine$_{ext}$
12	Leucine$_{ext}$	25	Urea
13	Phenylalanine$_{ext}$	26	Ethanolamine$_{ext}$

Table A.2
Internal metabolites in matrix $\mathbf{N}$.

Row occupied in matrix $\mathbf{N}(m \times n)$	Internal metabolite	Row occupied in matrix $\mathbf{N}(m \times n)$	Internal metabolite
1	Glucose6P	37	Tryptophan
2	Fructose6P	38	Valine
3	DihydroxyacetoneP	39	α-Ketobutyrate
4	Glyceraldehyde3P	40	PropionylCoA
5	3Phosphoglycerate	41	α-Ketoadipate
6	Pyruvate	42	AcetoacetylCoA
7	AcetylCoA	43	Acetoacetate
8	Citrate	44	Glutamate-γ-semialdehyde
9	α-Ketoglutarate	45	Ornitine
10	SuccinylCoA	46	Citrulline
11	Succinate	47	Argininosuccinate
12	Fumarate	48	PRPP
13	Malate	49	IMP
14	Oxaloacetate	50	Orotate
15	Ribulose5P	51	UTP$_{RN}$
16	Ribose5P	52	ATP$_{RN}$
17	Xylulose5P	53	GTP$_{RN}$
18	Erythrose4P	54	CTP$_{RN}$
19	Glutamate	55	dATP
20	Aspartate	56	dGTP
21	Glycine	57	dTTP
22	Serine	58	dCTP
23	Tyrosine	59	Ethanolamine
24	Cysteine	60	Choline
25	Alanine	61	Glycerol3P
26	Arginine	62	Phosphatidylethanolamine
27	Asparagine	63	Phosphatidylcholine
28	Glutamine	64	Phosphatidylserine
29	Histidine	65	Sphingomyelin
30	Isoleucine	66	Cholesterol
31	Leucine	67	Protein
32	Lysine	68	DNA
33	Methionine	69	RNA
34	Phenylalanine	70	MembraneLipid
35	Proline	71	NH_4^+
36	Threonine	72	CO_2

References

Altamirano, C., Illanes, A., Becerra, S., Cairó, J., Gòdia, F., 2006. Considerations on the lactate consumption by CHO cells in the presence of galactose. Journal of Biotechnology 125, 547–556.

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

Boghigian, B. A., Seth, G., Kiss, R., Pfeifer, B. A., 2009. Metabolic flux analysis and pharmaceutical production. Metabolic Engineering. doi:10.1016/j.ymben.2009.10.004.

Bonarius, H.P.J., Hatzimanikatis, V., Meesters, K., Goojier, C., Schimd, G., Tramper, J., 1996. Metabolic flux analysis of hybridoma cells in different culture media using mass balances. Biotechnology and Bioengineering 50, 299–318.

Bonarius, H.P.J., Houtman, J.H., de Gooijer, C.D., JohannesTramper, Schmid, G., 1998. Activity of glutamate dehydrogenase is increased in ammonia-stressed hybridoma cells. Biotechnology and Bioengineering 57, 447–453.

Bonarius, H.P.J., Houtman, J.H., Schmid, G., de Gooijer, C.D., Tramper, J., 2000. Metabolic flux analysis of hybridoma cells under oxidative and reductive stress using mas balances. Cytotechnology 32, 97–107.

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

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

Hu, W.-S., 2004. Cell Biology for Process Engineering. Tech. Rep., Cellular Bioprocess Technology, University of Minnesota.

Kao, F.-T., Puck, T.T., 1967. Genetics of somatic mammalian cells. IV. Properties of Chinese hamster cell mutants with respect to the requirement for proline. Genetics 55, 513–524.

KEGG, 2008. Kyoto Encyclopedia of Genes and Genomes. Online. http://www.genome.ad.jp/kegg/pathway/map/map01100.html.

Klamt, S., Schuster, S., 2002. Calculating as many fluxes as possible in underdetermined metabolic networks. Molecular Biology Reports 29, 243–248.

Klamt, S., Schuster, S., Gilles, E.D., 2002. Calculability analysis in underdetermined metabolic networks illustrated by a model of the central metabolism in purple nonsulfur bacteria. Biotechnology and Bioengineering 77 (7), 734–751.

Llaneras, F., Picó, J., 2007. An interval approach for dealing with flux distributions and elementary modes activity patterns. Journal of Theoretical Biology 246, 290–308.

Lovrecz, G., Gray, P., 1994. Use of on-line gas analysis to monitor recombinant. Mammalian Cell cultures, 167–175.

Nadeau, I., Sabatié, J., Koehl, M., Perrier, M., Kamen, A., 2000. Human 293 cell metabolism in low glutamine-supplied culture: interpretation of metabolic changes through metabolic flux analysis. Metabolic Engineering 2, 277–292.

Nelson, D., Cox, M., 2005. Lehninger, Principles of Biochemistry, fourth edition.

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

Provost, A., 2006. Metabolic design of dynamic bioreaction models. Ph.D. Thesis, Université Catholique de Louvain.

Provost, A., Bastin, G., 2003. Dynamical metabolic modelling of biological processes. In: Fourth International Symposium on Mathematical Modelling, February.

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

Provost, A., Bastin, G., Schneider, Y., 2007. From metabolic networks to minimal dynamic bioreaction models. In: 10th International IFAC Symposium on Computer Applications in Biotechnology, June.

Savinell, J.M., Palsson, B.O., 1992a. Optimal selection of metabolic fluxes for in vivo measurement. I. Development of mathematical methods. Journal of Theoretical Biology 155, 201–214.

Savinell, J.M., Palsson, B.O., 1992b. Optimal selection of metabolic fluxes for in vivo measurement. II. Application to *Escherichia coli* and hybridoma cell metabolism. Journal of Theoretical Biology 155, 215–242.

Schneider, M., Marison, I.W., von Stockar, U., 1996. The importance of ammonia in mammalian cell culture. Journal of Biotechnology 46, 161–185.

Schuster, S., Dandekar, T., Fell, D.A., 1999. Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. TIBTECH 17, 53–60.

Seoka, N., 1988. Directional mutation preassure and neutral molecular evolution. Proceedings of the National Academy of Sciences of the United States of America 85, 2653–2657.

Sidorenko, Y., Wahl, A., Dauner, M., Genzel, Y., Reichl, U., 2008. Comparison of metabolic flux distributions for MDCK cell growth in glutamine- and pyruvate-containing media. Biotechnology Progress 24, 311–320.

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

Sueoka, N., 1961. Correlation between base composition of deoxyribonucleic acid and amino acid composition of protein. Proceedings of the National Academy of Sciences of the United States of America 47, 1141–1149.

Sueoka, N., 1962. On genetic basis of variation and heterogeneity of DNA base composition. Proceedings of the National Academy of Sciences of the United States of America 48, 582–592.

van der Heijden, R., Heijnen, J., Hellinga, C., Romein, B., Luyben, K., 1994. Linear constraint relations in biochemical reaction systems. I. Classification of the calculability and the balanceability of conversion rates. Biotechnology and Bioengineering 43, 3–10.

Wahl, A., Sidorenko, Y., Dauner, M., Genzel, Y., Reichl, U., 2008. Metabolic flux model for an anchorage-dependent MDCK cell line. Biotechnology and Bioengineering 101, 135–152.

Xie, L., Wang, D.I., 1994. Applications of improved stoichiometric model in medium design and fed-batch cultivation of animal cells in bioreactor. Cytotechnology 15, 17–29.

Xie, L., Wang, D.I., 1996a. Material balance studies on animal cell metabolism using stoichiometrically based reaction network. Biotechnology and Bioengineering 52, 579–590.

Xie, L., Wang, D.I., 1996b. Energy metabolism and ATP balance in animall cell cultivation using a stoichiometrically based reaction network. Biotechnology and Bioengineering 52, 591–601.

Zupke, C., Stephanopoulos, G., 1995. Intracellular flux analysis in hybridomas using mass balances and in vitro ^{13}C NMR. Biotechnology and Bioengineering 45, 292–303.