Multi-objective experimental design for ^{13}C-based metabolic flux analysis

Jeroen Bouvin[a], Simon Cajot[a,1], Pieter-Jan D'Huys[b], Jerry Ampofo-Asiama[c,1], Jozef Anné[d], Jan Van Impe[e], Annemie Geeraerd[c], Kristel Bernaerts[a,*],

[a]*Bio- & chemical systems Technology, Reactor Engineering and Safety Section, Department of Chemical Engineering, KU Leuven, Celestijnenlaan 200F, 3001 Leuven, Belgium*
[b]*Department of Industrial Science and Technology, UC Leuven-Limburg, Agoralaan Building B, 3590 Diepenbeek, Belgium*
[c]*Division of Mechatronics, Biostatistics and Sensors (MeBioS), Department of Biosystems, KU Leuven, Willem de Croylaan 42, 3001 Leuven, Belgium*
[d]*Laboratory of Molecular Bacteriology, Department of Microbiology and Immunology, KU Leuven, Herestraat 49, 3000 Leuven, Belgium*
[e]*Chemical and Biochemical Process Technology and Control section, Technology Campus Ghent, Gebroeders De Smetstraat 1, 9000 Gent, Belgium*

Abstract

13C-based metabolic flux analysis is an excellent technique to resolve fluxes in the central carbon metabolism but costs can be significant when using specialized tracers. This work presents a framework for cost-effective design of ^{13}C-tracer experiments, illustrated on two different networks.

Linear and non-linear optimal input mixtures are computed for networks for *Streptomyces lividans* and a carcinoma cell line. If only glucose tracers are considered as labelled substrate for a carcinoma cell line or *S. lividans*, the best parameter estimation accuracy is obtained by mixtures containing

*Corresponding author
Email address: kristel.bernaerts@cit.kuleuven.be (Kristel Bernaerts)
[1]former member

high amounts of 1,2-$^{13}C_2$ glucose combined with uniformly labelled glucose. Experimental designs are evaluated based on a linear (D-criterion) and non-linear approach (S-criterion). Both approaches generate almost the same input mixture, however, the linear approach is favoured due to its low computational effort. The high amount of 1,2-$^{13}C_2$ glucose in the optimal designs coincides with a high experimental cost, which is further enhanced when labelling is introduced in glutamine and aspartate tracers. Multi-objective optimization gives the possibility to assess experimental quality and cost at the same time and can reveal excellent compromise experiments. For example, the combination of 100% 1,2-$^{13}C_2$ glucose with 100% position one labelled glutamine and the combination of 100% 1,2-$^{13}C_2$ glucose with 100% uniformly labelled glutamine perform equally well for the carcinoma cell line, but the first mixture offers a decrease in cost of $ 120 per ml-scale cell culture experiment.

We demonstrated the validity of a multi-objective linear approach to perform optimal experimental designs for the non-linear problem of 13C-metabolic flux analysis. Tools and a workflow are provided to perform multi-objective design. The effortless calculation of the D-criterion can be exploited to perform high-throughput screening of possible ^{13}C-tracers, while the illustrated benefit of multi-objective design should stimulate its application within the field of 13C-based metabolic flux analysis.

Keywords: 13C-based metabolic flux analysis, multi-objective optimal experimental design, *Streptomyces lividans*, central carbon metabolism, Carcinoma cell line, cost-effective experimental design

---

1. Introduction

During the last decades, 13C-based metabolic flux analysis (13C-MFA) has developed into an powerful tool in cell engineering and bioprocess optimization. Complementary to transcriptomics and proteomics, 13C-MFA offers the possibility to in vivo quantify intracellular fluxes in different metabolic phenotypes of cultured cell (Antoniewicz, 2013; Young, 2014). Several research works on fundamental intracellular flux quantification as well as on biotechnological applications (e.g., Ahn and Antoniewicz, 2013; Driouch et al., 2012; Leighty and Antoniewicz, 2012) apply 13C-MFA but its wide-spread use remains limited because of its computational challenges but possibly also due to the expenses related to high 13C-tracers costs.

While most 13C-MFA studies make use of common and relatively cheap labelled substrates (e.g., position one labelled glucose, uniformly labelled glucose), the increasing availability of various labelled carbon sources opens up new opportunities in 13C-MFA. Alternative 13C-tracers can increase flux estimation quality or reduce practical flux identifiability problems. Metallo et al. (2009) and Nargund and Sriram (2013), for example, selected 1,2-$^{13}C_2$-glucose as an excellent tracer for resolving the phosphoglucoisomerase flux and all metabolic fluxes in the central carbon metabolism of mammalian and plant cells, respectively. In a follow-up study on mammalian cells, Walther et al. (2012) applied a genetic algorithm to optimize mixtures of glucose and glutamine tracers, which resulted in an optimal input mixture of 1,2-$^{13}C_2$-glucose and uniformly labelled glutamine. The use of more customized labelled substrates brings along an increase in experimental cost, e.g., 1,2-$^{13}C_2$-glucose costs three times more than uniformly labelled glucose. A clear view

on the experimental cost of optimal 13C-tracer experiments with respect to information content is, however, lacking and could impose a bottleneck for its widespread acceptance in the biotech research community. Multi-objective optimal design which optimizes for both cost-effectiveness and parameter estimation accuracy can be helpful here but has never been reported, as far as the authors are aware.

The current paper focusses on cost-effective optimal design for 13C-based metabolic flux analysis. Since both linear and non-linear approaches of optimal experimental design are used for 13C-MFA, both methods are evaluated in this paper. The classical optimal experimental design maximizes or minimizes a scalar function of the Fisher Information Matrix, which is a linear approximation of the cost surface around its optimum. The novel optimal experimental design approach accounts for the non-linear nature of the 13C-MFA parameter estimation problem. This approach calculates non-linear confidence intervals of intracellular fluxes and summarizes their accuracy in the precision score, as proposed by Metallo et al. (2009). Within the knowledge of the authors, no study has ever compared the linear and non-linear approaches with respect to the resulting optimal designs.

Two different software platforms, namely influx_s (Sokol et al., 2012) and 13C-FLUX2 (Weitzel et al., 2013), are used and the methods are illustrated for two distinct networks: a network of a mammalian cell and a bacterial cell, i.e., a lung carcinoma cell line and *Streptomyces lividans*, respectively. The network of the carcinoma cell line is taken from Metallo et al. (2009)

and Walther et al. (2012). An accurate description of the metabolism of such cancer cells can help with identification of possible targets for treatment development. Carcinoma cells are highly compartmentalized, grow on complex media and require at least two carbon substrates.

In contrary, *S. lividans* is non-compartmentalized and can easily grow on one carbon substrate. 13C-based metabolic flux studies on *S. lividans* and the accompanying optimal input design have not yet been reported in literature, notwithstanding their industrial importance. *S. lividans* is a Gram-positive bacterium, mostly found in the soil. Streptomycetes are well known for their production of a wide variety of antibiotics and hydrolytic enzymes (Anné et al., 2012; D'Huys et al., 2011). These products are easily obtainable at high concentration in the extracellular medium due to the inherent potent secretion capacity of Streptomycetes (Anné et al., 2012). The latter advantage has resulted in studies of Streptomycetes as a host for heterologous protein production. Vrancken and Anné (2009) show an extensive list of large and small proteins with industrial and therapeutic applications, which are successfully secreted by *S. lividans*. However, the secretion of heterologous products causes a so-called metabolic burden, often leading to unsatisfactory yields (Anné et al., 2012). Metabolic engineering can help here to relax this metabolic burden and improve protein yields. Lule et al. (2012) show how over-expression of phosphoenolpyruvate carboxykinase in *S. lividans* TK24 leads to an increased yield of heterologous human tumour necrosis factor-alpha. To understand the origin of this positive effect as well as to find other targets for metabolic engineering, metabolic flux analysis is the best tool to use. Although preliminary flux balance analysis has already been performed

for *S. lividans* TK24 (D'Huys et al., 2012), 13C-metabolic flux analysis is necessary to increase the knowledge concerning intracellular flux distribution when using different kinds of substrates. A first step is to determine the optimal glucose input (mixture). In addition, the impact on input designs should be studied when using additional substrates in *S. lividans*, since research demonstrated a preferential uptake of glutamate and aspartate over all other amino acids (D'Huys et al., 2011).

For the above-mentioned networks, the present work aims to find optimal and cost-effective mixtures of labelled substrates. The message of this paper is not the specific input design as such, but the methodology for linear and non-linear optimal design in combination with cost optimization. Such a multi-objective approach is new in the field of metabolic flux analysis. To achieve this goal, an optimal mixture of glucose and glutamine tracers will be determined for a lung carcinoma cell line by optimizing a scalar function of the Fisher Information Matrix and a non-linear criterion, defined by Metallo et al. (2009). Advantages and disadvantages of both approaches will be discussed. Besides the results of specific tracers mixtures, a complete overview will be shown, which enables a clearer understanding of the effect of specific tracers on flux estimation accuracy. In a second part, the optimization methodology will be repeated on *S. lividans*, fed with only glucose or mixtures of glucose and aspartate tracers. For both cases, a multi-objective optimization will make the trade-off between information content and experimental cost and will determine a range of suitable compromises.

2. Materials and Methods

Theoretical background and computational implementation details are explained in this section. Figure 1(A) summarizes the different elements in the overall multi-objective optimal experimental design framework. Specific inputs are required to set up a complete metabolic network model (e.g., nominal fluxes, measurement type) which is further used to calculate linear and non-linear designs for optimal parameter estimation (Section 2.1 and 2.2). Subsequently, the combination of experimental cost with information content is embedded in the multi-objective framework (Section 2.3). Section 2.4 and 2.5 specify the applied network models and 13C-tracer inputs. To conclude, Section 2.6 gives details on the software used, the implementation of the multi-objective design framework as well as the supplementary files provided with this paper.

2.1. Optimal experimental design: a linear approach

The Fisher Information Matrix (FIM) is the key factor in linear optimal experimental design (OED) (e.g., Walter and Pronzato, 1997). The FIM is a linear approximation of the cost surface in the neighbourhood of the true parameters. It combines information on the sensitivity of the model outputs to small variations in the model parameters and on the uncertainty of the output measurements. For 13C-MFA, the model outputs are the net substrate uptake fluxes and the mass isotopomer measurements, while the model parameters are the free, independent fluxes. An important property of the FIM is that the inverse of this matrix approximates the Cramér Rao lower bound of the parameter variancecovariance matrix, when evaluated for the set

of true parameter values v^* (Walter and Pronzato, 1997). Since the inverse of the FIM approximates the covariance matrix of the estimated parameters, the diagonal of the FIM can be used to compute the confidence intervals of the estimated fluxes. The optimal design is found by maximisation/minimisation of a scalar function of the FIM. In the current work, the determinant of the FIM is maximized (D-criterion), which aims at maximisation of the global parameter estimation accuracy by minimizing the volume of the uncertainty region. The linear optimal experimental design approach yields the following equation in the case of 13C-MFA (e.g., Araúzo-Bravo and Shimizu, 2003):

$$\max \quad \det(\text{FIM}) \tag{1}$$

$$\text{with} \quad \text{FIM}(v^0, \mathbf{u}) \;=\; \text{Cov}(v^0)^{-1} = \left(\frac{\partial M}{\partial v}\right)_{v=v^0} \cdot \Sigma_M^{-1} \cdot \left(\frac{\partial M}{\partial v}\right)_{v=v^0}^{T} \tag{2}$$

where v^0 are the nominal values for the independent fluxes, $\mathbf{u}$ is the mixture design, M are the measured model outputs and Σ_M is the variance-covariance matrix of the output measurements.

In optimal experimental design for 13C-MFA experiments, the variables that will influence the FIM are the substrates, the type of measurements, the amount of measurements and the nominal values for the independent fluxes. The type (e.g., MS and/or NMR, free and/or proteinogenic amino acids) and amount of measurements are assumed to be fixed and known *a priori*. All published results are derived for GC-MS measurements, although the considered software platforms (Section 2.6) also accept NMR measurements. The possible substrates which are taken into account will be discussed in

Section 2.5. Due to the non-linear nature of the network model the FIM is also influenced by the values of the free, independent fluxes v, which are not known *a priori*. This requires a good first guess for values of the free fluxes, i.e., the nominal fluxes v^0. Previous research has already demonstrated the insensitivity of the optimal design with respect to small changes of the values in v^0 (e.g., Wiechert et al., 2001) and was confirmed for the case studies in this study (data not shown).

2.2. Optimal experimental design: a non-linear approach

Non-linear optimal experimental design accounts for the non-linear nature of 13C-MFA. First, non-linear confidence intervals are computed via Monte Carlo simulations. Afterwards, these confidence intervals are used to calculate a criterion value, which assesses the global accuracy of the experimental design.

In analogy with the linear approach, nominal values are chosen for the independent fluxes. Nominal values are used to calculate the corresponding nominal measurement values. In each Monte Carlo run, new measurement values are randomly sampled from normal distributions around their nominal values and the independent fluxes are estimated. The known standard error of each measurement is used as standard deviation for the normal distribution. After N Monte Carlo runs (and N parameter estimations) the 95% confidence interval can be derived for each individual flux by taking the 2.5% lower and 97.5% upper bound. In the current research N equals 200, which represents an acceptable compromise between duration of simulations and accuracy on the confidence intervals.

In order to compare different experimental designs, the S-criterion or the

so-called precision score, as defined by Metallo et al. (2009), is calculated. This criterion assesses the width of the confidence intervals. The search for the optimal design thus results in the following optimisation problem:

$$\max \quad S \tag{3}$$

$$\text{with} \quad S = \sum_i w_i \cdot \exp\left(-\frac{r_i}{\beta}\right) \tag{4}$$

$$r_i = \min\left(\frac{u_i}{|v_i|}, \frac{v_i}{|v_i|} + \alpha\right) - \max\left(\frac{l_i}{|v_i|}, \frac{v_i}{|v_i|} - \alpha\right) \tag{5}$$

where v_i, u_i and l_i are the mean value, the lower bound (2.5%) and the upper bound (97.5%) of the estimated flux i, respectively. α is a cut-off parameter that diminishes the influence of badly determined fluxes on the precision score (Metallo et al., 2009). The resulting r_i is a normalized range for the flux i. These normalized ranges are combined in the formula for the precision score S, where w_i and β are a weighing parameter for flux i and an overall scaling parameter, respectively. w_i can be one or zero. In the current work, the weighing parameters w_i of all free independent fluxes are set to 1, which results in an overall precision score that can vary from 0 (i.e. completely unidentifiable network) to a value equal to the number of free fluxes (i.e. perfectly identifiable network). The α and β values were set to 3. To get a full picture of the optimization space, the design criteria are not only calculated for specific predefined mixtures, but also for a grid, which spans the total space of possible substrate mixtures. This grid enables the identification of regions of optimal designs and indicates how the information content alters for varying input mixtures.

2.3. Multi-objective formulation

The multi-objective optimization problem for cost-effective 13C-MFA takes the following form:

$$\min \quad \{f_1(\mathbf{u}), f_2(\mathbf{u})\} \tag{6}$$

$$s.t \quad \mathbf{u} \in U, \tag{7}$$

$$\text{with} \quad f_1(\mathbf{u}) \equiv \det[FIM\left(v^0, \mathbf{u}\right)]^{-1} \quad \text{or} \quad S^{-1} \tag{8}$$

$$f_2(\mathbf{u}) \equiv \sum_{i=1}^{n} u_i C_i \tag{9}$$

The decision variable $\mathbf{u}$ is the vector containing the mixture fractions of the applied substrate tracers. U is the feasible space of these input mixtures and C represents the cost of a tracer per gram. In this paper, the multi-objective formulation contains two objectives, i.e. minimisation of the global parameter estimation uncertainty (expressed by the inverse of the D- or S-criterion) and minimisation of part of the consumable cost of the experiment (expressed as the cost per 1 gram of glucose). Other objectives as well as combinations of objectives can however be made. Whereas the decision variable vector is the main focus in single-objective optimization, this focus shifts towards the objective space in the multi-objective context. This is mainly due to the fact that usually no single optimal solution exists in a multiple objective optimization. Instead, the concept of Pareto optimality is introduced to determine the optimal solutions (Miettinen, 1999):

A decision vector $\mathbf{u}^* \in \mathbf{U}$ is Pareto optimal if there does not exist another decision vector $\mathbf{u} \in \mathbf{U}$ such that $f_i(\mathbf{u}) \leq f_i(\mathbf{u}^*)$ for all $i = 1, \ldots, k$ and $f_j(\mathbf{u}) < f_j(\mathbf{u}^*)$ for at least one index j. This means that a decision vector is Pareto optimal if there exists no other decision vector that improves at least

one objective without having a negative effect on another one. The resulting set of Pareto optimal solutions forms the so-called Pareto optimal set or Pareto front and returns, from a mathematical point of view, equally optimal solutions. A decision maker (a person or group), with a clear understanding of the problem, is now needed to select a specific solution.

To obtain the Pareto optimal set in the current study, the ϵ-constraint method is used as described in Miettinen (1999):

$$\min f_l(\mathbf{u}) \tag{10}$$

$$s.t. \quad f_j(\mathbf{u}) \leq \epsilon_j \text{ for all } j = 1, \ldots, k, \quad j \neq l, \tag{11}$$

$$\mathbf{u} \in U, \tag{12}$$

where $l \in \{1, \ldots, k\}$. By solving this problem for a fixed range of ϵ_j-values, the complete Pareto front can be calculated. For the bi-objective study, this means an optimization of the D (or S) -criterion, while keeping the cost constrained, or vice versa.

2.4. Network models

The reaction network for the central metabolism of the carcinoma cell line is obtained from Metallo et al. (2009) and Walther et al. (2012). This network contains the glycolysis pathway, the pentose phosphate pathway, the tricarboxylic cycle (TCA), amino acid synthesis and anaplerotic reactions. The model counts 27 intracellular metabolites and 46 fluxes, of which some are implemented as bidirectional. The designation for each flux can be found in Supplementary file 1. The bidirectional fluxes are presented by their net and exchange flux as described in Wiechert et al. (2001). Throughout the paper

net and exchange fluxes will be denoted by an n or x at the end of the flux name, respectively. Twenty fluxes were identified as free, independent fluxes. Their nominal values can be found in the additional files (Supplementary file 2). Type, number and precision of measurements are taken in agreement with Walther et al. (2012). A network for the central metabolism of *S. lividans* TK24 was set up based on D'Huys et al. (2012). Pathways included are the glycolysis pathway, pentose phosphate pathway, TCA cycle, anaplerotic and gluconeogenic reactions and nitrogen metabolism. Biomass precursor effluxes are calculated from the biomass composition of *Streptomyces coelicolor*, as applied in Borodina et al. (2005). The final model counts 35 metabolites and 71 fluxes of which some are implemented as bidirectional. The number of free fluxes is 18, of which 6 are net fluxes and 12 are exchange fluxes. A complete overview of the implemented network with the designation for each flux and metabolite, information on the carbon mapping, the bidirectionality of the reactions, the nominal flux values, and the presumed measurements can be found in Supplementary file 1.

2.5. Input mixture composition

Considered labelled glucose tracers in both models are uniformly labelled (U-GLC), position one labelled (1-GLC), unlabelled (0-GLC) and position one and two labelled glucose (1,2-GLC). Various mixtures of these substrates were applied, while the total glucose uptake remained the same. Applied glutamine tracers for the carcinoma cell line are uniformly labelled (U-GLN), position one labelled (1-GLN) and unlabelled glutamine (0-GLN). Various mixtures of these substrates were simulated, while the total uptake was constrained to 30% of the glucose uptake. In the case of an additional aspartate

uptake for *S. lividans*, uniformly labelled (U-ASP), position one labelled (1-ASP) and unlabelled aspartate (0-ASP) tracers are considered. When considering aspartate as additional substrate, the uptake flux was constrained to 10% of the glucose uptake, which is in agreement with the experimental observations of D'Huys et al. (2012). The prices of all substrates are summarized in Table 1.

Table 1: Prices and product numbers of all the substrates that are used in the simulations.

Substrate	**Product number**	**Price per gram**
0-GLC	G8270-1KG [a]	$ 0.04
1-GLC	CLM-420-1 [b]	$ 212
U-GLC	CLM-1396-1 [b]	$ 194
1,2-GLC	CLM-504-1 [b]	$ 775
1-ASP	489972-100MG [a]	$ 5340
0-ASP	A9256-100G [a]	$ 0.2
U-ASP	CLM-1801-1 [b]	$ 2450
1-GLN	CLM-3612-1 [b]	$ 1328
0-GLN	G3126-100G [a]	$ 0.61
U-GLN	CLM-1822-0.5 [b]	$ 5312

[a] Sigma Aldrich (Website consulted at 26/03/2014)

[b] Cambridge Isotope Laboratories (Website consulted at 26/03/2014)

2.6. Implementation and software platforms

All simulations were performed in two different 13C-MFA software environments, i.e. 13C-FLUX2 (Weitzel et al., 2013) and influx_s (Sokol et al., 2012). This choice has different reasons. First of all, by offering methodology implementations on different platforms, a broader range of research groups

can be reached, since the threshold to start optimal designs is lowered. An important fact here, is that design calculations on the carcinoma cell line, performed by Metallo et al. (2009) and Walther et al. (2012), were not performed in either 13C-FLUX2 or influx_s. The current research, therefore, offers two additional methodology implementations. Researchers can choose the software they are familiar with. Secondly, not all software platforms are equally efficient or user-friendly in respect to the desired calculations. Finally, input files of influx_s (FTBL-files) can easily be converted to input files for 13C-FLUX2 (FML-files), since 13C-FLUX2 offers a built-in feature to convert FTBL-files.

Some notable differences between the software packages read as follows. Both software platforms use different definitions for exchange fluxes. In influx_s all exchange fluxes are scaled to values between 0 and 1, which is not the case in 13C-FLUX2. In order to gather comparable results (also with the results of Metallo et al. (2009) and Walther et al. (2012)) exchange fluxes were not scaled. This required additional recalculations of the influx_s results. In contrary, Monte Carlo runs for non-linear calculation of the confidence intervals are a built-in feature in influx_s. This is not the case in 13C-FLUX2 and necessitates additional implementations. For both software platforms, the complete implementation for experimental design led to a comprehensive, user friendly, implementation which combines perl scripts, matlab-scripts and third-party scripts to attain both linear and non-linear results. A schematic overview of the non-linear implementation for 13C-FLUX2 for one specific input design is summarized in Figure 1(B). Figure 1(A) situates this specific

implementation within the complete framework of multi-objective experimental design for 13C-MFA, which is established in this work. The framework starts from a complete network model, either supplied as a FML- or FTBL-files. This network model summarizes all possible user-defined inputs like network structure, nominal fluxes, measurement types and numbers, considered substrate tracers and their corresponding cost (green boxes). By starting the implementation from this complete network model, flexibility with respect to user-defined inputs is ensured. Flexibility is also ensured within the optimal design simulations. By uncoupling the cost and performance calculations, multi-objective design can easily be extended to other criteria, like for example other widely used scalar function of the FIM (Walter and Pronzato, 1997). For each step within the general work flow, Figure 1(A) suggests some possible alterations. The different implementations of the information content calculation (i.e. D- and S-criterion in 13C-FLUX2 and influx_s) represent a major part of the presented work. All scripts which enable the reproduction of the published results are available at: https://perswww.kuleuven.be/∼u0006977/MathematicalBiosciences(Bouvin_etal).

For both the carcinoma cell line and *S. lividans*, the model reaction networks were implemented in the FTBL- and FML-format. Once the complete network is implemented, together with specification of the nominal values of the free fluxes, the whole network is fixed and all dependent reactions and measurements (model outputs) can be calculated. This is called a forward simulation. The FIM can also be calculated in the forward simulation.

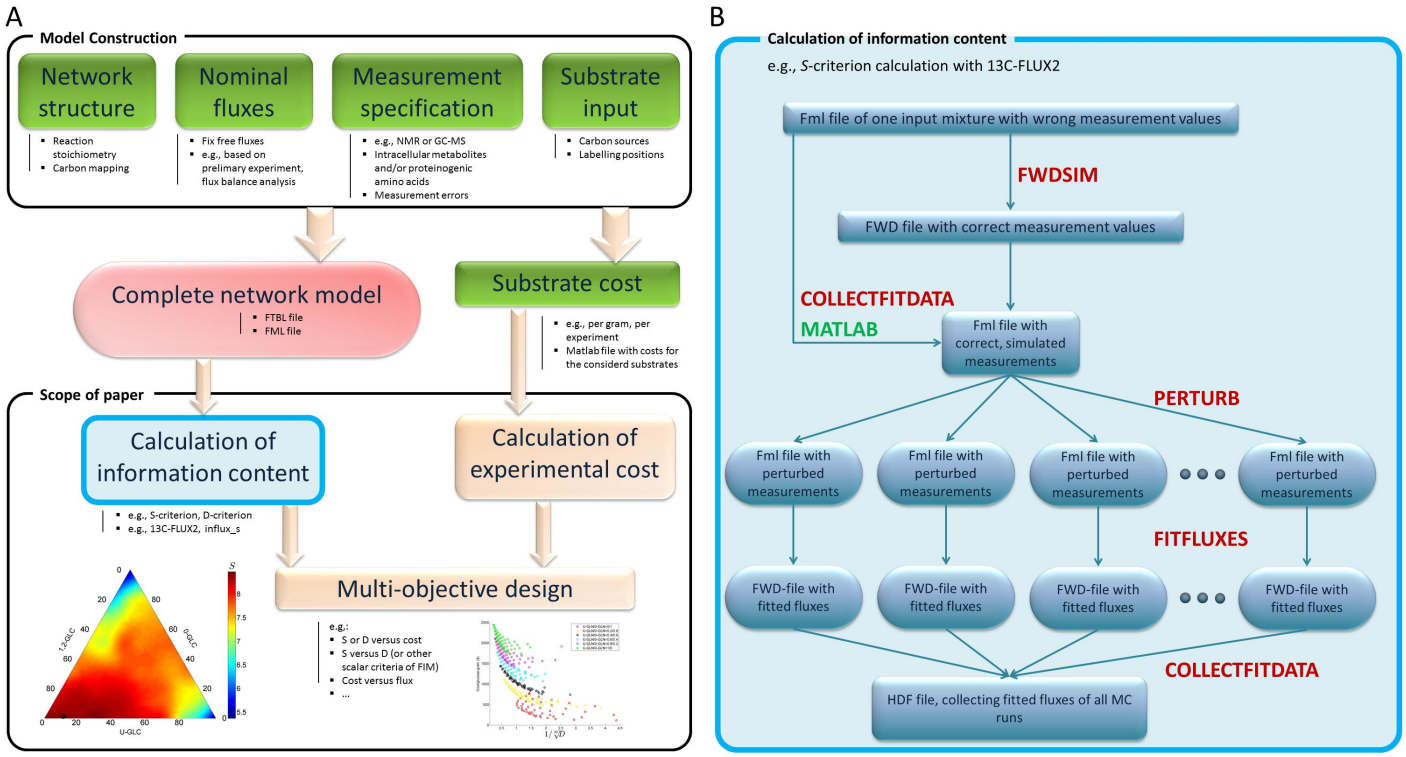

Figure 1: A) General flow scheme for implementation of a multi-objective optimal design for 13C-MFA. The green boxes indicate user-supplied inputs. B) Schematic overview of the implementation of non-linear optimal design in 13C-FLUX2 for one specific input design. In a first step, measurements are simulated based on the nominal values and the input mixture via a forward simulation (FWDSIM). The correct, simulated measurements are put in the FML-file and perturbed N-times for N Monte Carlo runs. This perturbation is accomplished by randomly sampling new measurement values from normal distributions around their nominal values. The known standard error of each measurement is used as standard deviation for the normal distribution. Fluxes are estimated (FITFLUXES) and all data are collected in one HDF-file. This HDF-file is processed in MATLAB to calculate the upper and lower bound of all estimated fluxes. The red words are built-in methods of 13C-FLUX2.

3. Results and discussion

3.1. Linear versus non-linear optimal experimental design: a comparison based on glucose tracers

In order to compare the results of D-optimal (linear) and S-optimal (non-linear) experimental design, the information content of input mixture designs is calculated for two exemplary metabolic networks. Mixtures of 1,2-GLC, U-GLC and 0-GLC are fed to the networks of carcinoma cells and *S. lividans*. The carcinoma cell network is also fed with unlabelled glutamine, which is an essential substrate. Classically, 13C-MFA studies apply mixtures of U-GLC, 0-GLC and 1-GLC, but Metallo et al. (2009) and Nargund and Sriram (2013) have shown that 1,2-GLC significantly augments the parameter estimation quality in mammalian and plant cell networks. Figure 2 depicts the mixture design triangles for 1,2-GLC, U-GLC and 0-GLC based on the linear and non-linear approach, and it can immediately be noticed that both designs reproduce very similar mixture triangles. The dark red coloured area indicates the input designs with the highest information content. For D-optimal design, a clear optimum is situated at a mixture of 56% 1,2-GLC and 44% U-GLC for carcinoma cells and at a mixture of 67% 1,2-GLC and 33% U-GLC for *S. lividans*. For S-optimal design, the area of high S-values is more spread out, especially for *S. lividans*. The S-optimal design for the carcinoma cell network (60/40 1,2-GLC/U-GLC) is consistent with the D-optimal design, taking into account the different grid sizes. For *S. lividans*, local optima are detected, i.e. 90/10 1,2-GLC/U-GLC and 70/30 1,2-GLC/U-GLC, the latter of which closely resembles the D-optimal design. The moderate resolution of the mixture triangle, the smeared out area of high S-values, and the mod-

erate number of Monte Carlo iterations per condition makes the non-linear approach prone to local optima. Levelling out of S-values can be related to the cut-off factor α, which is an arbitrary value in the definition of the precision score S (see Equation 5) and diminishes the influence of badly determined fluxes on the S-criterion. For example, increasing the amount of 1,2-GLC can significantly increase some confidence intervals of fluxes in the TCA (Metallo et al., 2009), but this is probably smoothed out by the α factor. Apart from the effect of the cut-off factor, S-optimal design is computationally intensive and cannot detect practically unidentifiable fluxes. An experiment with a mixture of 50% 0-GLC and 50% U-GLC, for example, is unable to determine the exchange flux of the glucose-6-phosphate isomerization step in the carconima network. This incapacity is reflected by a zero row in the FIM and a corresponding zero value for the D-criterion, while the S-value still scores high. In other words, if the D-criterion scores high, a good overall estimation of all fluxes is guaranteed.

Based on these analyses, it can be concluded that D-optimal and S-optimal experimental design return comparable optimal input mixture designs. This emphasizes the usefulness of linear optimal design of 13C-MFA, although this experimental fluxomics technique is non-linear of nature. Besides the computational advantage, calculations performed on identical networks always result in the same D-criterion value regardless of the software environment in which simulations are performed. Such a good agreement between linear and non-linear optimal experimental design was already postulated by Millard et al. (2014), yet was not illustrated.

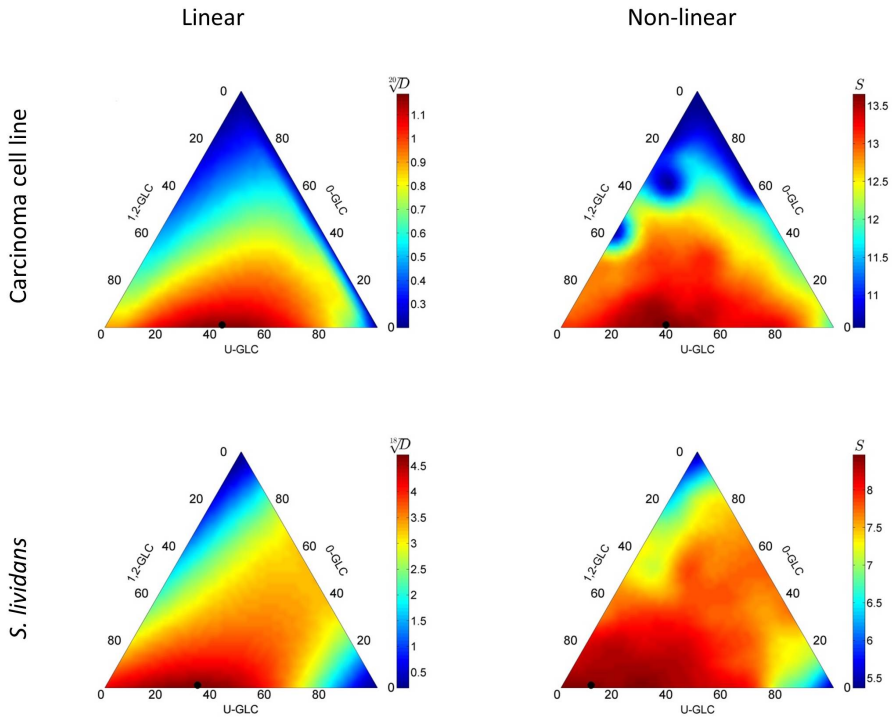

Figure 2: Optimal mixture of 1,2-GLC,U-GLC and 0-GLC for carcinoma cell line and *S. lividans*. The colormaps give an overview of the change in experimental quality when the mixture composition changes. The red regions are the most informative experiments. Both the linear and non-linear approach result in a similar lay-out of the colormaps. The optimal mixture is indicated with a black dot. The grid sizes are different in the linear and non-linear colormap. In the linear case, mixture compositions changes with steps of 1%, while in the non-linear case compositions change with steps of 10% in order to limit calculation time.

3.2. Linear versus non-linear optimal experimental design: tracers versus cost

Figure 3 evaluates the information content of a series of input mixtures fed to the carcinoma cell network and lists the associated cost of each input

design per gram of glucose. It should be emphasized that this unit of cost per gram of glucose is chosen as a reference value for tracer comparison. When only a single culture experiment is considered as described in Metallo et al. (2009), the real financial cost is one order of magnitude lower. Repetitions of optimal experiments would again multiply the costs. The considered set of experimental inputs includes pure traditional tracers (e.g., uniformly labelled glucose and position one labelled glucose) as well as 50/50 mixtures, which are commonly used in literature (e.g., Wiechert et al., 2001; Fischer et al., 2004; Möllney et al., 1999). Three mixtures from Figure 2 are also included, i.e. 100% 1,2-GLC, 50/50 1,2-GLC/U-GLC and the S-optimal mixture of 60/40 1,2-GLC/U-GLC. The set is further extended with the well performing mixture of 100% 1,2-GLC with 100% U-GLN (Walther et al., 2012; Metallo et al., 2009). The impact of U-GLN is also assessed by using mixtures of 100% 1-GLC and 100% U-GLN and a mixture of 100% 0-GLC and 100% U-GLN. This set already enables a comparison within a wide spectrum of possible tracers.

Figure 3 clearly illustrates the out-performance of the mixture 1,2-GLC/U-GLN (100/100), both in the linear and the non-linear design approach, but it also emphasizes the tremendous increase of substrate costs. Second best experiments are the mixtures of 1-GLC/U-GLN (100/100) followed by the optimized 1,2-GLC/U-GLC (60/40) with 100% unlabelled glutamine. Costs can drastically change depending on the input mixture, and it is the choice of the researcher to select the applied input mixture balancing between information content and experimental costs. Selecting the experiment with

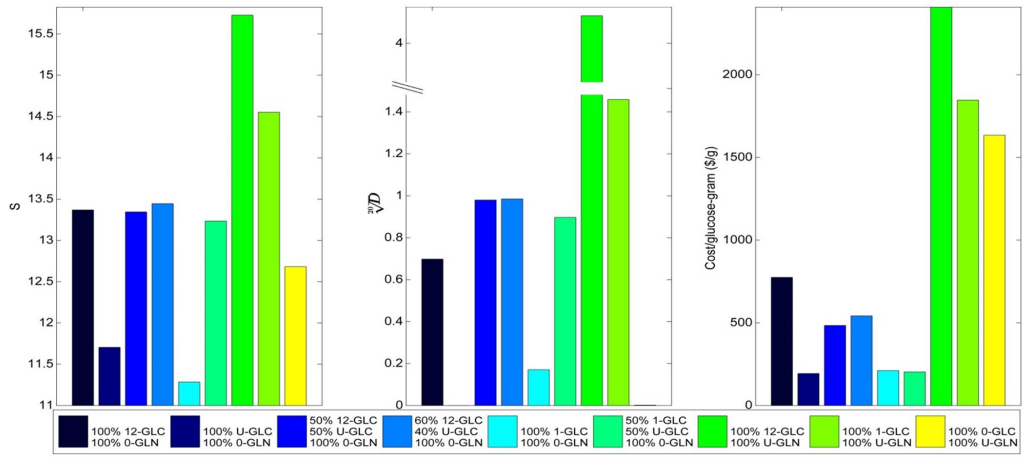

Figure 3: Information content and experimental cost for various sets of labelled substrates, fed to the carcinoma cell line. S-criterion (left), D-criterion (middle) and cost per glucose-gram (right) is shown for common and recently identified substrate mixtures, fed to the network of carcinoma cell line.

the highest information and cost is one option, while the other option could be to replicate the second best experiment in order to account for biological variability and to increase the information content. In order to help scientists to take such crucial decisions, a multi-objective optimal design approach can be exploited. Screening for tracers which optimize both experimental quality and cost could offer better alternatives.

Mixtures of 1-GLC/U-GLN and mixtures of 1,2-GLC/U-GLC have not yet been reported in literature. Previous work only focused on identifying tracers which maximize the quality of the parameter estimation with no constraints on cost. The best performing mixture of 1,2-GLC/U-GLN was identified by such search methods, which is confirmed in the current paper, but the coinciding feed cost of approximately $ 2500 per glucose-gram was never mentioned as a possible bottleneck in future 13C-MFA. Seeing that the mixtures

of 1-GLC/U-GLN and 1,2-GLC/U-GLC score well in experimental quality and offer a decrease in cost (especially the 1,2-GLC/U-GLC mixture), these designs could surface when considering cost as an extra optimization criterion.

Furthermore, 13C-MFA would benefit from a high-throughput screening of potential tracers and their mixture regions, which could enable a better understanding of the concept of optimal design in the field of 13C-MFA. It could be argued that the attempt for a high-throughput screening would imply a great, time-consuming effort, seeing the non-linear confidence-intervals. However, as illustrated in Figure 2 and 3, the D-criterion calculation results in the same relative ranking of the considered input designs, is calculated with minor effort and (as discussed earlier) guarantees a good overall estimations of all fluxes. The latter aspect is reinforced in Figure 3 for tracers which can not resolve the entire flux map (like U-GLC combined with 0-GLN and U-GLN combined with 0-GLC). Experiments with U-GLC for example are essentially the same as experiments with 0-GLC in the upper glycolysis pathway and the pentose phosphate pathway of the carinoma cell line. 0-GLN can not reach the upper part of the network (see network specifications in Metallo et al. (2009) and Walther et al. (2012)) which means that all intracellular metabolites in the upper part of the network will be fully labelled, resulting in unidentifiable fluxes. Similarly, experiments with U-GLN are proven to accurately predict fluxes in the TCA-cycle, but not in the glycolyis or the pentose phosphate pathway. A prerequisite in the calculation of the covariance matrix is that all fluxes are structural identifiable (Wiechert et al., 2001). Evaluation of the D-criterion is therefore not possible in the

latter two mixture designs, but S-values are obtained due to the smoothing property of the α cut-off factor. This statement can be verified by taking a look at the non-linear 95% confidence intervals of the free fluxes, summarized in the additional files (Supplementary file 2). Mixtures of U-GLC combined with 0-GLN and U-GLN combined with 0-GLC give rise to very large confidence intervals for a number of fluxes (e.g., v10.n, v2.n, v10.x, v12.x, v4.x) compared to the other mixtures, considered in Figure 3. The relatively high S-values can therefore unmistakeably be attributed to the α cut-off factor.

The confidence intervals also illustrate the potency of specific tracers to elucidate fluxes in different pathways as stated in previous publications (e.g., Metallo et al., 2009; Nargund and Sriram, 2013). Introducing U-GLN clearly reduces the confidence intervals in the TCA (e.g., v17.x and v18.x), while introducing 1,2-GLC has a big influence on fluxes in the pentose phosphate pathway and the glycolysis (e.g,. v2.n, v10.n, v10.x and v11.x). The same conclusions can be drawn from the linear confidence intervals, which are also summarized in the Supplementary file 2. As expected, these linear confidence interval tend to overestimate the confidence intervals.

3.3. Multi-objective optimal experimental design

Multi-objective optimization jointly optimizes two objective functions, in this case maximal information content and minimal experiment cost. Such multi-objective approach is first demonstrated on the same exemplary networks and input mixtures as in Figure 2. The D-criterion is used to quantify the performance of each input design. Figure 4 (left plots) depicts the Pareto front for each network. Grey boxes mark points which do not obey the Pareto

optimality definition and, therefore, do not belong to the Pareto front. Remaining points are all Pareto optimal and, therefore, mathematically spoken equally optimal. Non-Pareto optimal points are found because the ϵ-value in the problem statement (Equation 10) is implemented as equality constraints instead of inequality constraints. For completeness, the Pareto front is drawn onto the mixture triangle showing the composition changes along the Pareto line (rights plots in Figure 4).

The Pareto optimal line now allows the user to make a motivated choice balancing global parameter estimation quality and experimental cost. Figure 4 (left plots) shows that maximum D-value and minimum cost are conflicting objectives. In general, use of D-optimal mixtures allies with a very high cost. Starting in the optimal mixture and gradually replacing 1,2-GLC fractions with U-GLC systematically reduces cost, but strongly decreases the D-value. Opposed to the most expensive input mixture (100% 1,2-GLC), which is not Pareto optimal, a Pareto optimal mixture with equivalent D-value but lower costs exists. For the carinoma cell line this is a mixture of 20% 1,2-GLC, 62% U-GLC and 18% 0-GLC, while for *S. lividans* this is a mixture of 43% 1,2-GLC, 53% U-GLC and 4% 0-GLC.

3.4. Multi-objective and multi-substrate optimal experimental design

Intracellular metabolites from cells growing on multiple substrates can be labelled from different carbon sources. A multi-substrate optimal experimental design optimizing the labelling ratios for all carbon sources and all possible labelling compositions can thus be adopted. To illustrate this, the multi-objective design for the carcinoma cell network (see Figure 4, top left)

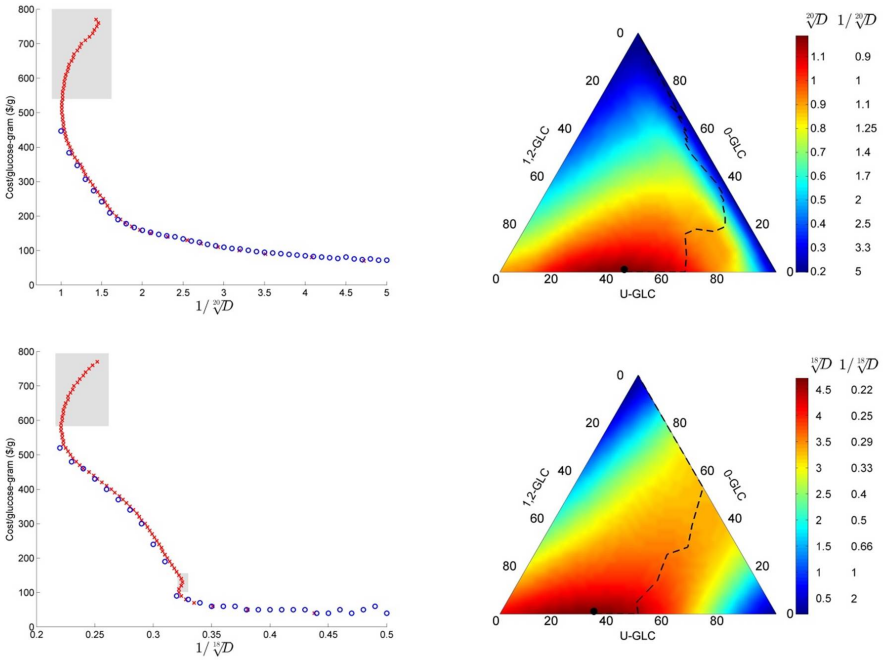

Figure 4: Result of multi-objective optimization between cost ($) and D-criterion for the network of the carcinoma cell line (top) and *S. lividans* (bottom). The red crosses are the result of the ϵ-constraint method when cost-values are constraint, while the blue circles represent the optimization when the D-criterion is constraint. Due to this hard constraints, experiments are identified which do not obey to the definition of the Pareto front (grey boxes).

is expanded to include additional glucose and glutamine tracers with different labelling positions. Selected results are shown in Figure 5. Mixtures triangles for three labelled substrates are calculated while systematically changing the fraction of two additional substrates. Different color codes are used for different mixture triangles, with the red color depicting the mixture triangle from Figure 2 (top left). Figure 5(A) illustrates the tremendous increase in substrates cost associated with increasing ratios of uniformly labelled glu-

tamine (U-GLN) to unlabelled glutamine (0-GLN), but it also significantly increases the information content. Best performing substrate mixtures combine 100% 1,2-GLC with fractions of U-GLN, and ideally with 100% U-GLN. U-GLN reproduces informative labelling in the TCA metabolites enabling accurate elucidation of fluxes in the TCA cycle. In this case, the Pareto front depicts the trade-off between information loss and cost, but it does not offer a better alternative for 100% 1,2-GLC/100% U-GLN.

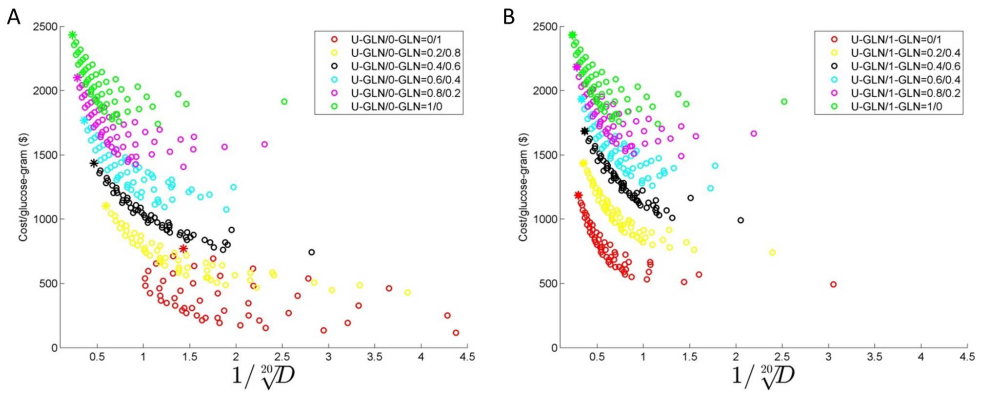

Figure 5: Multi-objective and multi-substrate design on the network of the carcinoma cell line. Figure A summarizes the D-criterion and the cost of all possible mixtures of 1,2-GLC, U-GLC and 0-GLC, combined with possible mixtures of U-GLN and 0-GLN. Figure B summarizes the D-criterion and the cost of all possible mixtures of 1,2-GLC, U-GLC and 0-GLC, combined with possible mixtures of U-GLN and 1-GLN. Mixtures with the same ratio U-GLN/0-GLN or U-GLN/1-GLN are marked in the same color. 100% 1,2-GLC is marked with an *.

In Figure 5(B), mixtures of uniformly labelled glutamine and position one labelled glutamine are added. Two input mixtures return almost the same highest global parameter estimation quality, but have significantly different

costs. The mixture of 100% 1,2-GLC and 100% 1-GLN offers high experimental quality and low cost. The D-criterion only drops from 4.3 to 3.4 as compared to 100% 1,2-GLC with 100% U-GLN, while the cost drops from approximately $ 2500 towards $ 1200 per glucose-gram. From both an economical and experimental view, Figure 5(B) also reveals the fruitless effort of using mixtures of U-GLN and 1-GLN. Mixtures of both lowers the D-criterion, compared to the use of either pure 1-GLN and U-GLN, and belong, according to the definition, not to the Pareto front.
In order to re-strengthen the claim that optimal design results based on the D-criterion are valid, the S-criterion is calculated for the mixture of 100% 1,2-GLC with 100% 1-GLN. The resulting S-value of 15.57 closely approaches the highest S-value (15.73) in Figure 3. Although these results point at the relevant contribution of the 1-GLN tracer, it could not been identified as an informative tracer by the genetic algorithm of Walther et al. (2012).

As discussed earlier, *S. lividans* can, in contrast with the carcinoma cell line, grow on a single carbon substrate. When searching for the optimal tracers, feeding with one or two substrates should therefore be treated as two different design and estimation problems. Figure 2 already summarized the results when applying common glucose tracers as single substrate. When considering *S. lividans* experiments grown on two substrates, additional mixtures of 0-ASP, 1-ASP and U-ASP are considered. Different cost-experimental quality overviews are calculated, of which the most interesting is shown in Figure 6. Without performing any optimal designs, extrapolating the previous multi-objective results of the carcinoma cell line would result in a mixture

of 100% 1,2-GLC combined with 100% 1-ASP. For the current network, this would result in a non-optimal experiment with the maximum cost ($ 1350). Instead, the highest global parameter estimation quality is obtained when using a mixture of 90/10 1,2-GLC/U-GLC combined with 100% U-ASP. The cost of this experiment is approximately $ 950.

These conclusions support the need and merit of multi-objective and multi-substrate optimal experimental design when performing 13C-MFA on a new network and should stimulate 13C-MFA researchers to perform such designs, seeing the easy and quick calculation of the D-criterion.

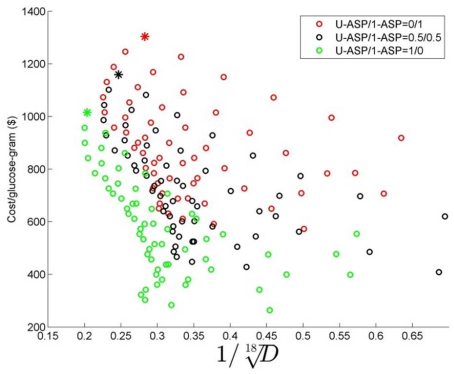

Figure 6: Multi-objective and multi-substrate design on the network of *S. lividans*. The figure summarizes the D-criterion and the cost of all possible mixtures of 1,2-GLC, U-GLC and 0-GLC, combined with possible mixtures of U-ASP and 1-ASP. Mixtures with the same ratio U-ASP/1-ASP are marked in the same color. 100% 1,2-GLC is marked with an *.

3.5. Alternative software implementation

Implementation of the models and confidence interval calculations were performed in both influx_s and 13C-FLUX2. Further data analysis was

performed in MATLAB. Calculations of the D-criterion were for both software platforms exactly the same. Non-linear calculations however diverted, due to the numerical nature of calculations, differentiation in optimizer and differences in simulation set-up. The diversion is especially notable for badly identifiable fluxes. 13C-FLUX2 appears less sensitive and still generates relatively small confidence intervals, resulting in even more misleading values of the S-criterion. For the carcinoma cell line, 100% U-GLN combined with 100% 0-GLC, for example, outperformed a 50/50 mixture of 1-GLC/U-GLC, which is not in accordance with the D-criterion and previous published results (Walther et al., 2012). For this reason, published S-values are a result of calculations with influx_s. An overview of the corresponding S-values, calculated in 13C-FLUX2, can be found in Supplementary file 2, Figure S4.

4. Conclusion

This study reports on cost-effective experimental designs for 13C-metabolic flux analysis, implemented on the central carbon metabolism of *S. lividans* and a carcinoma cell line. In a first part, the search for the optimal design was implemented based on a linear and a non-linear approach. In both cases, validation is based on a single scalar value, i.e the D-criterion and the S-criterion, respectively. The same optimal design is found, independently of the chosen approach, which opens the door for a high-throughput screening of labelled substrates based on the computationally effortless calculation of the linear D-criterion, once the network has been characterized.

In a second part this screening possibility was combined with minimization of experimental cost. Previous studies, focusing on identifying the optimal

mixtures, do not emphasize the financial cost of experiments with respect to parameter estimation quality, although the new tracers cause a significant increase in required budget. When comparing different mixtures of 1,2-GLC, U-GLC and 0-GLC, the optimal mixture in both networks consisted of 1,2-GLC and U-GLC, i.e 56/44 1,2-GLC/U-GLC for the carinoma cell line and 67/33 1,2-GLC/U-GLC for the network of *S. lividans*. Both mixtures generate a noticeable decrease in cost as compared to 100% 1,2-GLC, which was identified in previous research as the most effective tracer. Moreover, the multi-objective approach offers a way to evaluate the information loss when budget is constraint.

In the last part, the dimensionality of the problem is increased by extending the set of possible substrate types. The most striking result was obtained in the case of the carcinoma cell line, were a mixture of 1,2-GLC and 1-GLN renders the same global accuracy for the parameter estimation, but costs approximately $ 130 per culture experiment less (i.e. more than a 50% reduction) than the high performant mixture of 1,2-GLC and U-GLN identified by Walther et al. (2012). In addition, a mere extrapolation of this result to the network of *S. lividans* would result in an inferior experiment, since a mixture of U-ASP with a high amount of 1,2-GLC outperforms 1-ASP with 1,2-GLC in both price as experimental quality. The latter remark highlights the importance for an *a priori* experimental design when studying a new organism or strain.

The presented framework for multi-objective optimal design for 13C-MFA is very flexible and can be adopted to variable networks, phenotypes, substrates,

criteria, mixture composition, etc. Labelling purity of the used tracers can be accounted for in both software platforms. If this is not taken into account during the design, this information should at least be determined and used during flux estimation.

Follow-up studies are recommended to make use of the beneficial properties of the D-criterion when designing optimal experiments. Recently, Walther et al. (2012) developed a genetic algorithm to find the optimal labelling in which non-linear calculations were used to evaluate the designs. Since the D-criterion offers the same qualitative evaluation and requires less computational effort, the genetic algorithm could easily be extended with extra objectives (like cost) and more substrates.

Possible variations in the search for the optimal input design can be found in extending the multi-objective approach by other criteria or other design approaches. If one specific flux is targeted, its confidence interval can, for example, be used as an extra criterion. Another possibility is the combination of multiple scalar criteria of the FIM, as illustrated by Telen et al. (2012). Furthermore, combining the design approach based on the EMU basis vector methodology, as applied by Crown et al. (2012), with financial constraints could yield extra rational labelling rules in the search for optimal experimental designs in 13C-based metabolic flux analysis.

Ethics

The authors declare that no experiments have been performed as part of the research for this manuscript.

Availability of supporting data

Source code is available at:
https://perswww.kuleuven.be/∼u0006977/MathematicalBiosciences(Bouvin_etal)

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Jeroen Bouvin is supported by a PhD grant of the Agency for Innovation by Science and Technology (IWT). Jerry Ampofo-Asiama is supported by the Research Council of the Katholieke Universiteit Leuven (DBOF/08/033). The research leading to these results has received funding from the European Commission's Seventh Framework Programme (FP7) under the grant agreement STREPSYNTH (project n° 613877). The scientific responsibility is assumed by its authors.

References

Ahn, W. S., Antoniewicz, M. R., Jan. 2013. Parallel labeling experiments with [1,2-(13)C]glucose and [U-(13)C]glutamine provide new insights into CHO cell metabolism. Metab. Eng. 15, 34–47.

Anné, J., Maldonado, B., Van Impe, J., Van Mellaert, L., Bernaerts, K., 2012. Recombinant protein production and streptomycetes. J. Biotechnol. 158 (4), 159–167.

Antoniewicz, M. R., 2013. 13c metabolic flux analysis: optimal design of isotopic labeling experiments. Curr. Opin. Biotechn. 24 (6), 1116 – 1121.

Araúzo-Bravo, M. J., Shimizu, K., 2003. An improved method for statistical analysis of metabolic flux analysis using isotopomer mapping matrices with analytical expressions. J. Biotechnol. 105 (12), 117 – 133.

Borodina, I., Krabben, P., Nielsen, J., Jun. 2005. Genome-scale analysis of Streptomyces coelicolor A3(2) metabolism. Genome Res. 15 (6), 820–829.

Crown, S. B., Ahn, W. S., Antoniewicz, M. R., Jan. 2012. Rational design of ^{13}C-labeling experiments for metabolic flux analysis in mammalian cells. BMC Syst. Biol. 6 (1), 43.

D'Huys, P.-J., Lule, I., Van Hove, S., Vercammen, D., Wouters, C., Bernaerts, K., Anné, J., Van Impe, J. F. M., 2011. Amino acid uptake profiling of wild type and recombinant Streptomyces lividans TK24 batch fermentations. J. Biotechnol. 152 (4), 132–143.

D'Huys, P.-J., Lule, I., Vercammen, D., Anné, J., Van Impe, J. F. M., Bernaerts, K., Sep. 2012. Genome-scale metabolic flux analysis of Streptomyces lividans growing on a complex medium. J. Biotechnol. 161 (1), 1–13.

Driouch, H., Melzer, G., Wittmann, C., 2012. Integration of in vivo and in silico metabolic fluxes for improvement of recombinant protein production. Metab. Eng. 14 (1), 47–58.

Fischer, E., Zamboni, N., Sauer, U., Feb. 2004. High-throughput metabolic flux analysis based on gas chromatographymass spectrometry derived 13C constraints. Anal. Biochem. 325 (2), 308–316.

Leighty, R. W., Antoniewicz, M. R., 2012. Parallel labeling experiments with [U-^{13}C]glucose validate E. coli metabolic network model for ^{13}C metabolic flux analysis. Metab. Eng. 14 (5), 533 – 541.

Lule, I., Maldonado, B., D'Huys, P.-J., Van Mellaert, L., Van Impe, J., Bernaerts, K., Anné, J., Oct. 2012. On the influence of overexpression of phosphoenolpyruvate carboxykinase in Streptomyces lividans on growth and production of human tumour necrosis factor-alpha. Appl. Microbiol. Biotechnol. 96 (2), 367–372.

Metallo, C. M., Walther, J. L., Stephanopoulos, G., 2009. Evaluation of ^{13}C isotopic tracers for metabolic flux analysis in mammalian cells. J. Biotechnol. 144 (3), 167 – 174.

Miettinen, K., 1999. Nonlinear Multiobjective Optimization. Kluwer Academic Publishers, Boston.

Millard, P., Sokol, S., Letisse, F., Portais, J.-C., 2014. Isodesign: A software for optimizing the design of 13c-metabolic flux analysis experiments. Biotechol. Bioeng. 111 (1), 202–208.

Möllney, M., Wiechert, W., Kownatzki, D., de Graaf, A. A., Jan. 1999. Bidirectional reaction steps in metabolic networks: IV. Optimal design of isotopomer labeling experiments. Biotechnol. Bioeng. 66 (2), 86–103.

Nargund, S., Sriram, G., Jan. 2013. Designer labels for plant metabolism: statistical design of isotope labeling experiments for improved quantification of flux in complex plant metabolic networks. Mol. BioSyst. 9 (1), 99–112.

Sokol, S., Millard, P., Portais, J.-C., 2012. influx_s: increasing numerical stability and precision for metabolic flux analysis in isotope labelling experiments. Bioinformatics 28 (5), 687–693.

Telen, D., Logist, F., Van Derlinden, E., Tack, I., Van Impe, J., Aug. 2012. Optimal experiment design for dynamic bioprocesses: a multi-objective approach. Chem. Eng. Sci. 78, 82–97.

Vrancken, K., Anné, J., 2009. Secretory production of recombinant proteins by Streptomyces. Future Microbiol. 4 (2), 181–188.

Walter, E., Pronzato, L., 1997. Identification of parametric models from experimental data. Springer, Berlin.

Walther, J., Metallo, C., Zhang, J., Stephanopoulos, G., 2012. Optimization of ^{13}C isotopic tracers for metabolic flux analysis in mammalian cells. Metab. Eng. 14, 162–171.

Weitzel, M., Nöh, K., Dalman, T., Niedenführ, S., Stute, B., Wiechert, W., Jan. 2013. 13CFLUX2–high-performance software suite for ^{13}C-metabolic flux analysis. Bioinformatics 29 (1), 143–145.

Wiechert, W., Möllney, M., Petersen, S., de Graaf, A. A., 2001. A universal framework for 13C metabolic flux analysis. Metab. Eng. 3, 265–283.

Young, J. D., 2014. 13c metabolic flux analysis of recombinant expression hosts. Current Opinion in Biotechnology 30 (0), 238 – 245, chemical biotechnology Pharmaceutical biotechnology.

Appendix. Supplementary Materials

Supplementary file 1 — Network of S. lividans and carcinoma cell line

Full implementation details of the network of *S. lividans* (metabolic reactions, carbon mapping, measurements, constraints) are provided.

Supplementary file 2 — Summary of linear and non-linear confidence intervals of the carcinoma network and comparison of S-criterion obtained via influx_s and 13C-FLUX2.