Bidirectional Reaction Steps in Metabolic Networks
Part IV: Optimal Design of Isotopomer Labeling Experiments

Michael Möllney, Wolfgang Wiechert, Dirk Kownatzki
IMR, Department of Simulation
University of Siegen, 57068 Siegen, Germany

Albert A. de Graaf
Institute of Biotechnology
Research Center Jülich, 52425 Jülich, Germany

Accepted for publication in "Biotechnology an Bioengineering"
July 5, 1999

Revised Manuscript

Running title: Bidirectional Reaction Steps in Metabolic Networks (Part IV)

Corresponding author: Wolfgang Wiechert
IMR, Department of Simulation
University of Siegen
57068 Siegen
Germany
Phone: -271 / 740-47 27, Fax: -23 65
E-mail: wiechert@simtec.imr.mb.uni-siegen.de

Acknowledgements: This project is partially funded by the Deutsche Forschungsgemeinschaft

Contents

Abstract: This contribution generalizes the statistical tools for the evaluation of carbon labeling experiments that have been developed for the case of positional enrichment systems in Part II of this series to the general case of isotopomer systems. For this purpose a new generalized measurement equation is introduced that can describe all kinds of measured data like positional enrichments, relative ^{13}C NMR multiplet intensities or mass isotopomer fractions produced with MS instruments. Then, to facilitate the specification of the various measurement procedures available, a new flexible textual notation is introduced from which the complicated generalized measurement equations are automatically generated. Based on these measurement equations, a statistically optimal flux estimator is established and parameter covariance matrices for the flux estimation are computed. Having implemented these tools different kinds of labeling experiments can be compared by using statistical quality measures. A general framework for the optimal design of carbon labeling experiments is established on the basis of this method. As an example it is applied to the *Corynebacterium* network from Part II extended by various NMR and MS measurements. In particular, the positional enrichment, multiplet or mass isotopomer measurements with the greatest information content for flux estimation are computed (measurement design) and various differently labeled input substrates are compared with respect to flux estimation (input design). It is discussed in detail how the measurement procedure influences the estimation quality of specific fluxes like the pentose phosphate pathway influx.

Keywords: metabolic flux analysis, isotopomer measurements, cumomers, optimal experimental design

1 Introduction

In Part III of this series (Wiechert *et al.*, 1999) (henceforth simply called Part III) the model for positional carbon labeling systems developed in Part I (Wiechert & de Graaf, 1997) was extended to general isotopomer labeling systems. For this purpose the concepts of cumomers, cumomer fractions and cumomer networks were introduced which turned out to be a very elegant tool to describe and analyze isotopomer labeling experiments. Part IV now deals with the extension of the statistical methods developed in Part II (Wiechert *et al.*, 1997a) to isotopomers. The main goal of this contribution is a detailed quantitative comparison of various experimental labeling techniques with respect to the achievable flux information. This problem of optimal experimental design has been under discussion in recent years but has never been precisely answered on objective criteria.

1.1 *Isotopomer measurements*

In principle isotopomer fractions (or more precisely certain linear combinations of them) can be measured with two techniques: nuclear magnetic resonance (NMR) and mass spectrometry (MS). The different measurement techniques are now characterized in more detail with respect to the formulation of the measurement equations that will be introduced below. In particular, emphasis is laid on the statistical quality of each data source:

1. The 1-dimensional ^{1}H-NMR measurement of positional ^{13}C enrichment (PE) was extensively discussed in Parts I and II (see also (Marx *et al.*, 1996)). This method requires purified substances because it is difficult to disentangle the large number of peaks arising for a substance mixture. On the other hand, this method produces the highest accuracy achievable with NMR. The reason is that ^{1}H-NMR has a much higher sensitivity than ^{13}C-NMR and that this technique requires no additional calibration measurements because both ^{13}C- and ^{12}C-bounded protons are distinguishable in one and the same spectrum (Figure 1a). Thus the peak areas corresponding to different carbon isotopes can be directly normalized to the PE scale.

2. The quantification of isotopomers by 1-dimensional ^{13}C-NMR relies on the presence of multiplet peaks (MP) - i.e. doublets, triplets or doublets of doublets - in the ^{13}C-spectrum caused by two or more neighboring labeled carbon atoms. Figure 1b shows some typical MP spectra of alanine, which is a measurable metabolic successor of pyruvate. The precise correspondence between the isotopomers and the MP peaks they produce can be derived once the coupling constants between all neighboring carbon atoms are known.

   Unfortunately, several different isotopomers may produce identical multiplets. For example (see Figure 1b) the alanine isotopomers Ala#000, Ala#100, Ala#001 and Ala#101 cannot be distinguished in a mixture. Generally, the available measurements can only be used to distinguish certain groups of isotopomers. Moreover, ^{12}C carbons cannot be detected by ^{13}C-NMR so that the total amount of carbon atoms remains unknown unless further independent measurements are carried out. Consequently, a normalization of peak areas to a percentage scale involves considerable time and effort and for that reason ratios of peak areas are usually used for data evaluation (Jeffrey *et al.*, 1991; Künnecke *et al.*, 1993; Szyperski, 1995). Only if additional ^{1}H-NMR measurements are performed can the complete isotopomer distribution be quantified (de Graaf *et al.*, 1996). Finally, ^{13}C-NMR is rather insensitive so that it takes a comparatively long measurement time to produce high-resolution spectra with adequate signal to noise ratios.

3. The 2-dimensional ^{1}H-^{13}C-NMR (2D-NMR) technique essentially allows the ^{13}C quantification of isotopomers to be performed with the sensitivity of ^{1}H-NMR. The ^{13}C multiplet resonances thus emerge in a 2-dimensional coordinate system dispersed according to the chemical shift of the bonded protons. The advantage of this method is that peak overlaps are almost completely eliminated in two dimensions so that a mixture of many compounds can be analyzed in a single experiment (Szyperski, 1995). This represents considerable progress because there is no need for separation of amino acids as in (Marx *et al.*, 1996). The only drawback is that the resolution in the ^{13}C dimension is somewhat limited.

4. Finally, some isotopomer measurements can be obtained with mass spectrometry (MS) (Inbar & Lapidot, 1987; Donato *et al.*, 1993; Christensen & Nielsen, 1999), which is usually combined with gas chromatography as GC-MS or liquid chromatography as LC-MS (Figure 1c). This also enables mixtures to be analyzed without prior separation of the components. MS can only separate isotopomers with different molecular mass (so-called *mass isotopomers* (Lee *et al.*, 1991)). Thus a metabolite with n carbon atoms essentially enables $n + 1$ mass isotopomers (MI) to be measured. Moreover, also the isotopomers of compound fragments produced by thermal cleavage in the MS instrument can be detected. This considerably increases the amount of measured data. A signal normalization to a percentage scale is necessary if not all MS peaks corresponding to one compound are quantified (Lee, 1993). In principle, the sensitivity and precision of MS is unrivaled offering a chance to directly measure various intracellular compounds that have been hitherto inaccessible (Christensen & Nielsen, 1999).

   However, care must be taken in the evaluation of MI spectra. Firstly, the naturally stable isotopes of O, H, N, S, P also contribute to the MI signals thus significantly biasing the measured peak areas. However, this effect

can be corrected after the measurement (Lee *et al.*, 1991; Lee *et al.*, 1992) and thus will be ignored later on. Secondly, the biological experiment is never stationary with such a high precision and certain corrections like the washout correction in (Marx *et al.*, 1996; Wiechert, 1996) have to be considered. Summarizing, a MS signal must never be assumed to be as precise as the instrument alone because the whole sampling and measurement process must be considered.

As will be shown in Section "Isotopomer Measurements", all these measurement procedures yield – up to normalization to a percentage scale – certain linear combinations of isotopomer fractions. In the following the term *isotopomer measurement* is used for such a measurable linear combination. The combination of PE, MP and MI data sometimes enables all isotopomer fractions of a certain metabolite to be quantified. The principal computational details of this measurement disentangling process have already been sketched in (Wiechert & de Graaf, 1996).

1.2 *Types of labeling experiments*

Carbon labeling experiments are still expensive with respect to manpower and material. Consequently, it is important to get maximum information about the unknown intracellular fluxes with minimal effort. Apart from choosing one of the newly developed measurement techniques described above the experimenter can compose the input substrate to achieve optimal results. In particular isotopomers can be drawn quite freely from the following four sources:

Unlabeled substrates (up to natural ^{13}C abundance) are always cheaply available as a mixture component.

Specifically labeled substrates are ^{13}C labeled at a single specified carbon position. For example [1-^{13}C]-glucose (henceforth abbreviated as 1- glucose) has been successfully used in Part II.

Multiply specifically labeled substrates like [5,6-$^{13}C_2$]-glucose (henceforth abbreviated as 5,6-glucose) are rather expensive and the extent to which the use of such substrates can help to increase the flux information remains an open question.

Uniformly labeled substrates , i.e. compounds labeled at each carbon atom position are available at a price comparable to singly labeled substrates. For example [1,2,3,4,5,6- $^{13}C_6$]-glucose (henceforth denoted as U-glucose) has been successfully used in (Szyperski, 1995; Sauer *et al.*, 1997).

Clearly, a purely unlabeled as well as a purely uniformly labeled substrate will yield no flux information at all. Thus they must always be mixed with another species. In (Szyperski, 1995) and (Schmidt *et al.*, 1998a) different mixtures of uniformly labeled, unlabeled and 1-labeled glucose have been used with good results. In general, mixtures of input isotopomers are promising candidates to obtain statistically well determined flux estimates (Schmidt *et al.*, 1998b). However, no quantitative results are currently available to decide which mixture is best for a certain experiment. This question is treated in detail in this contribution.

1.3 *Statistical analysis*

As in the case of positional carbon labeling, flux estimation from isotopomer measurements is best performed by using a least squares minimization approach (cf. Part II). This method is able to simultaneously exploit all the available measurement data and thus yields a maximum of statistical information. However, to achieve optimal results the statistical properties of the measurements have to be properly modeled. Apart from some application-specific examples (Sauer *et al.*, 1997) based on uniformly labeled input substrates the least squares method was first applied in full generality in (Schmidt *et al.*, 1998b). However, the problem of ill-conditionedness in the case of large exchange fluxes (see Part II) could not be generally solved in (Schmidt *et al.*, 1997).

A general tool for the statistical analysis of isotopomer labeling experiments based on analytical derivatives like that established in Part II for positional carbon labeling systems is not yet available. Statistical analysis has only been applied in some specific cases (Lee, 1993). To achieve reliable results a powerful mathematical formalism is required to express the variety of NMR and MS isotopomer measurements available. This generalization of the measurement matrices from Part II was first sketched in (Wiechert & de Graaf, 1996; Wiechert, 1996) but without the necessary normalization operation that was later described by (Schmidt *et al.*, 1998b). However, it will be shown in Appendix A that the statistical properties of the measurements were suboptimally exploited in the latter approach.

The most important statistical quality measure that is computed in a regression analysis is the covariance matrix of the estimated parameters. Based on this matrix, confidence regions can be computed and possibly occurring linear dependencies between the estimated parameters can be analyzed (see Part II). In combination with the parameter sensitivity matrix the causes of linear dependencies can then be investigated in more detail (Chatterjee & Hadi, 1988). The first rough estimate of this matrix was achieved by (Schmidt *et al.*, 1998a) by using a Monte Carlo approach. However this approach is computationally much too expensive to be used as a general procedure in optimal experimental design where covariance matrices have to be repeatedly computed.

1.4 *Optimal experimental design*

As has become clear so far, there is now a large variety of possible measurement techniques for carbon labeling experiments. The right choice of the experimental conditions depends on the flux information that is achieved and also on the the cost of the experiment. Currently, no systematic methods based on quantitative criteria are available to compare the different experimental approaches for flux determination.

Clearly, such a criterion should in some way be derived from the ellipsoidal confidence region that can be computed from the parameter covariance matrix for a single experiment. This parameter confidence region depends on the following experimental conditions that can be influenced by the experimenter:

Measured pools, i.e. the set of all metabolites extracted from cell mass.

Measured data, i.e. positional enrichments (PEs), multiplets (MPs) or mass isotopomers (MIs).

Input substrate, i.e. the composition of the substrate pool from a given set of available isotopomers (unlabeled, singly, multiply or uniformly labeled).

In the theory of optimal experimental design (Pázman, 1986), various methods have been developed to compare the information yield of different experiments. To this end several scalar criteria (like the A-, D-, E-, or G-criterion) have been derived to compare the shapes of the respective confidence regions in the parameter space. The D-criterion has become the most popular in the past decade (Munack, 1989; Takors *et al.*, 1997). It essentially measures the volume of the confidence ellipsoid and is closely related to the well-known Fisher information. For this reason it is the most natural criterion because it does not only consider the single parameter confidence intervals but also the correlations among the parameters which may lead to a needle-shaped confidence ellipsoid.

A disadvantage of a scalar criterion is that all aspects of estimation precision are condensed into one single number. If the user is interested only in certain parameters the criterion should be modified. For example, the experimenter might be interested in one single flux (e.g. the pentose phosphate pathway influx), in a subset of all fluxes (e.g. all anaplerotic fluxes or all net fluxes) or in a flux that is linearly dependent on the free fluxes. To meet these requirements the D-criterion can be restricted to a set of linear combinations of the free fluxes as defined by a matrix $\mathbf{L}$ which leads to the $D_{\mathbf{L}}$- criterion (Pázman, 1986).

A well-known conceptual problem for nonlinear optimal experimental design is the dependency of the parameter covariance matrix on the true values of the unknown parameters. This means that an experiment has actually to be performed in order to compute the quality criterion. At first glance this would mean that experimental design methods are useless for flux determination. As far as the choice of measurements is concerned the problem of a priori flux knowledge is not that critical because the quality of the flux estimate can always be refined a posteriori by making additional measurements. This leads to an iterative design procedure similar to that described in (Takors *et al.*, 1997). On the other hand, there is no chance of correcting a wrong input substrate mixture a posteriori without a complete repetition of the experiment. The usual way to escape this vicious circle is to assume a value for the unknown flux parameters a priori and to rely on the assumption that the comparison of two experiments by the scalar criterion is rather insensitive with respect to the assumed flux values. As will be shown below this robustness assumption can be easily tested by performing parameter variation studies.

1.5 *Aims of this contribution*

The following aims of the present contribution are derived from the preceding discussion:

1. Establishment of general measurement equations that can describe all types of available measurements including their normalization to a percentage scale and introduction of a flexible textual notation to facilitate the specification of all different measurement combinations.

2. Generalization of the statistical model from Part II to isotopomer experiments based on the preparatory work from Part III and extension to general isotopomer experiments of the software tools already available for the automatic generation of model equations.

3. Implementation of an algorithm for flux estimation and computation of parameter covariance matrices based on the analytically computed gradients from Part III thus minimizing the computational effort.

4. Application of statistical methods from regression theory to achieve a quantitative judgement of the achievable flux information by the different types of labeling experiments and measurements; establishment of some general theorems and procedures for optimal experimental design of labeling experiments.

5. Demonstration of all the developed methods by a complex example (i.e. the central metabolism of *Corynebacterium glutamicum*). In particular the influence of the measurement procedure on the quality of the pentose phosphate pathway influx estimate is discussed.

In summary, the different methods that are currently propagated for metabolic flux analysis will be quantitatively compared thus establishing a rational platform for the optimal experimental design of labeling experiments.

2 Isotopomer Measurements

With the concepts presented in Part III all components of the general statistical model presented in Part II can now be generalized to isotopomer systems. The most important difference between positional and isotopomer labeling systems is that the state vector $\mathbf{x}$ formerly composed of all positional enrichments is now extended to all cumomer fractions. The positional enrichments can be easily recovered as the 1-cumomer subvector ${}^1\mathbf{x}$ of $\mathbf{x}$. Moreover, there is a simple linear transformation $\mathbf{T}$ between isotopomer and cumomer fractions. In Part III it has been shown how the cumomer fractions can be computed as a function $\mathbf{\Gamma}$ of the forward and backward fluxes in the system.

The new measurement equations for isotopomer systems must now link the isotopomer measurement data vector $\mathbf{y}$ with the new labeling state vector $\mathbf{x}$. At the same time, the normalization operation for peak areas must be included appropriately in the model. This is the aim of the current section.

2.1 *Generalized measurement matrices*

In order to model the relation between the labeling state vector $\mathbf{x}$ and the measurement data vector $\mathbf{y}$ the concept of measurement matrices as introduced in Part II has to be extended in such a way that PE, MP and MI data sets can be represented by using the same formalism. Two steps have to be taken in that direction:

1. Measurement matrices are now allowed to be general matrices where the rows are no longer required to be unit vectors (as was the case in Part II).

2. In general MP and MI fractions corresponding to one compound are only measured up to a scaling factor. This must be accounted for by introducing scaling parameters which will lead to a bilinear structure of the measurement equations.

One measurement equation has to be written for each group of measurement data with a common scaling factor. For example, all MP areas corresponding to one carbon atom position are scaled by the same factor. Thus in general the measurement equation for the kth group of measurements has the form

$$\mathbf{y}_k = \omega_k \cdot \mathbf{M}'_{\mathbf{y},k} \cdot \mathbf{x} + \varepsilon_{\mathbf{y},k} \tag{1}$$

with the labeling state vector $\mathbf{x}$, the measurement data vector $\mathbf{y}_k$ for the kth group, a certain measurement matrix $\mathbf{M}'_{\mathbf{y},k}$, an (unknown) scaling parameter ω_k, and the measurement errors $\varepsilon_{\mathbf{y},k}$. The $'$-decoration of the matrix $\mathbf{M}$ is used because this is a preliminary definition of the measurement matrix that will be modified later on. Some examples of the matrix structures that can arise for common experimental methods are now given.

2.2 *Examples of measurement matrix composition*

To illustrate the composition of the measurement matrices consider an alanine molecule Ala with its 3 carbon atoms. Assume for the moment that Ala is the only metabolite in the system, so that the labeling state vector is given by the cumomer fraction vector (written in binary and positional notation) as

$$\begin{aligned} \mathbf{x} &= (ala_{xxx}, \quad ala_{xx1}, \quad ala_{x1x}, \quad ala_{x11}, \quad ala_{1xx}, \quad ala_{1x1}, \quad ala_{11x}, \quad ala_{111})^T \\ &= (ala, \quad ala_3, \quad ala_2, \quad ala_{23}, \quad ala_1, \quad ala_{13}, \quad ala_{12}, \quad ala_{123})^T \end{aligned}$$

or alternatively by the isotopomer fraction vector

$$\overline{\mathbf{x}} = (ala_{000}, \quad ala_{001}, \quad ala_{010}, \quad ala_{011}, \quad ala_{100}, \quad ala_{101}, \quad ala_{110}, \quad ala_{111})^T \quad .$$

Using the isotopomer-cumomer transformation matrix $\mathbf{T}_3$ presented in Part III these vectors are linked by $\mathbf{x} = \mathbf{T}_3 \cdot \overline{\mathbf{x}}$. Now the three different types of isotopomer measurements are treated:

1. Looking more thoroughly at the origin of PE data than in Part II it should be recognized that each PE measurement is actually composed of the area of the corresponding ^{12}C- and ^{13}C-related resonance peaks in the ^{1}H-NMR spectrum (Figure 1a). As an example, the second carbon atom position of alanine is considered. Then the quantities ala_2 and $1 - ala_2 = ala_{xxx} - ala_{x1x}$ are both measured up to a scaling factor ω_{AlaH2}. This leads to the first measurement matrix

$$\mathbf{M}'_{\text{AlaH2}} = \begin{pmatrix} . & . & 1 & . & . & . & . & . \\ 1 & . & -1 & . & . & . & . & . \end{pmatrix} \tag{2}$$

   where dots represent zero entries. Alternatively, the measurements can be written with isotopomer fractions $\bar{\mathbf{x}}$ by using the transformed matrix:

$$\overline{\mathbf{M}}'_{\text{AlaH2}} = \mathbf{M}'_{\text{AlaH2}} \cdot \mathbf{T}_3 = \begin{pmatrix} . & . & 1 & 1 & . & . & 1 & 1 \\ 1 & 1 & . & . & 1 & 1 & . & . \end{pmatrix}$$

   Likewise, the matrix $\overline{\mathbf{M}}'_{\text{AlaH3}}$ is composed with the corresponding scaling parameter ω_{AlaH3}.

2. Assume now that MPs can be measured by ^{13}C-NMR as shown in Figure 1b. Consider for example the singlet, doublet A, doublet B and doublet of doublets signal of the second carbon atom. The corresponding measurement matrix can be most easily formulated by using the isotopomer fraction coordinates:

$$\overline{\mathbf{M}}'_{\text{AlaC2}} = \begin{pmatrix} . & . & 1 & . & . & . & . & . \\ . & . & . & . & . & . & 1 & . \\ . & . & . & 1 & . & . & . & . \\ . & . & . & . & . & . & . & 1 \end{pmatrix}$$

From this it is immediately obtained:

$$\mathbf{M}'_{\text{AlaC2}} = \overline{\mathbf{M}}'_{\text{AlaC2}} \cdot \mathbf{T}_3^{-1} = \begin{pmatrix} . & . & 1 & -1 & . & . & -1 & 1 \\ . & . & . & . & . & . & 1 & -1 \\ . & . & . & 1 & . & . & . & -1 \\ . & . & . & . & . & . & . & 1 \end{pmatrix} \tag{3}$$

Likewise the measurement matrix $\mathbf{M}'_{\text{AlaC3}}$ is obtained. It has two rows because only one singlet and one doublet occur (see Figure 1b). The corresponding scaling parameters are $\omega_{\text{AlaC2}}, \omega_{\text{AlaC3}}$.

3. At last, the mass spectrometric measurement is described (Figure 1c). After the natural isotope correction (Lee *et al.*, 1992) there remain four MI peaks corresponding to the molecules of Ala that are 0, 1, 2, or 3 times ^{13}C labeled. This immediately yields the measurement matrix

$$\overline{\mathbf{M}}'_{\text{AlaM}} = \begin{pmatrix} 1 & . & . & . & . & . & . & . \\ . & 1 & 1 & . & 1 & . & . & . \\ . & . & . & 1 & . & 1 & 1 & . \\ . & . & . & . & . & . & . & 1 \end{pmatrix} \tag{4}$$

with respect to isotopomer fractions and the matrix

$$\mathbf{M}'_{\text{AlaM}} = \overline{\mathbf{M}}'_{\text{AlaM}} \cdot \mathbf{T}_3^{-1} \tag{5}$$

with respect to cumomer fractions together with a scaling factor ω_{AlaM}.

Finally, it is assumed that the 2,3-fragment of alanine is also measured. The corresponding measurement matrix then is:

$$\overline{\mathbf{M}}'_{\text{Ala23M}} = \begin{pmatrix} 1 & . & . & . & 1 & . & . & . \\ . & 1 & 1 & . & . & 1 & 1 & . \\ . & . & . & 1 & . & . & . & 1 \end{pmatrix} \tag{6}$$

2.3 *General structure of measurement equations*

If all the measurements described above are available, the Equations (2,3,5,6) given as examples yield the overall measurement equation (compare to Equation 1):

$$\begin{pmatrix} \mathbf{y}_{\text{AlaH2}} \\ \mathbf{y}_{\text{AlaH3}} \\ \mathbf{y}_{\text{AlaC2}} \\ \vdots \\ \mathbf{y}_{\text{Ala23M}} \end{pmatrix} = \begin{pmatrix} \omega_{\text{AlaH2}} \cdot \mathbf{M}'_{\text{AlaH2}} \\ \omega_{\text{AlaH3}} \cdot \mathbf{M}'_{\text{AlaH3}} \\ \omega_{\text{AlaC2}} \cdot \mathbf{M}'_{\text{AlaC2}} \\ \vdots \\ \omega_{\text{Ala23M}} \cdot \mathbf{M}'_{\text{Ala23M}} \end{pmatrix} \cdot \mathbf{x} + \begin{pmatrix} \varepsilon_{\text{AlaH2}} \\ \varepsilon_{\text{AlaH3}} \\ \varepsilon_{\text{AlaC2}} \\ \vdots \\ \varepsilon_{\text{Ala23M}} \end{pmatrix}$$

Remember that this equation was built upon the simplifying assumption that alanine is the only metabolite in the system. In general the vector $\mathbf{x}$ has also entries $\mathbf{x}_i$ for all cumomers of metabolites other than alanine so that the dimension of the preliminary measurement matrices given above doesn't match the dimension of $\mathbf{x}$.

Assume, for example, that Ser and Asp are also present in the system and that additional measurement matrices $\mathbf{M}'_{\text{SerC2}}, \mathbf{M}'_{\text{SerC3}}$ and $\mathbf{M}'_{\text{AspM}}$ have been constructed as described above. Then the cumomer fractions are

$$\mathbf{x} = (ser_{xxx}, \ldots, ser_{111}, ala_{xxx}, \ldots, ala_{111}, asp_{xxx}, \ldots, asp_{111})^T$$

and the general structure of the measurement equation is given by

$$\underbrace{\begin{pmatrix} \mathbf{y}_{\text{SerC2}} \\ \mathbf{y}_{\text{SerC3}} \\ \mathbf{y}_{\text{AlaH2}} \\ \vdots \\ \mathbf{y}_{\text{Ala23M}} \\ \mathbf{y}_{\text{AspM}} \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} \omega_{\text{SerC2}} \cdot \mathbf{M}'_{\text{SerC2}} & \mathbf{0} & \mathbf{0} \\ \omega_{\text{SerC3}} \cdot \mathbf{M}'_{\text{SerC3}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \omega_{\text{AlaH2}} \cdot \mathbf{M}'_{\text{AlaH2}} & \mathbf{0} \\ \vdots & \vdots & \vdots \\ \mathbf{0} & \omega_{\text{Ala23M}} \cdot \mathbf{M}'_{\text{Ala23M}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \omega_{\text{AspM}} \cdot \mathbf{M}'_{\text{AspM}} \end{pmatrix}}_{\mathbf{M}_{\mathbf{y}}(\omega)} \cdot \mathbf{x} + \underbrace{\begin{pmatrix} \varepsilon_{\text{SerC2}} \\ \varepsilon_{\text{SerC3}} \\ \varepsilon_{\text{AlaH2}} \\ \vdots \\ \varepsilon_{\text{Ala23M}} \\ \varepsilon_{\text{AspM}} \end{pmatrix}}_{\varepsilon_{\mathbf{y}}} \tag{7}$$

where $\mathbf{0}$ denotes zero matrices (of varying dimensions) and

$$\omega = (\omega_{\text{SerC2}}, \omega_{\text{SerC3}}, \omega_{\text{AlaH2}}, \ldots, \omega_{\text{Ala23M}}, \omega_{\text{AspM}})^T$$

is the vector of all scaling factors. Now the final measurement matrices

$$\mathbf{M}_{\mathbf{y},i}, \quad i = \text{SerC2, SerC32, AlaH2, ..., AspM}$$

are defined as

$$\mathbf{M}_{\text{SerC2}} = \begin{pmatrix} \mathbf{M}'_{\text{SerC2}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \; \mathbf{M}_{\text{SerC3}} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{M}'_{\text{SerC3}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \; \mathbf{M}_{\text{AlaH2}} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}'_{\text{AlaH2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \ldots$$

Here the $\mathbf{0}$-matrices are sized as in Equation (7) so that all $\mathbf{M}_{\mathbf{y},i}$ have the same dimensions. The measurement Equation (7) can then be concisely written in the form

$$\mathbf{y} = \mathbf{M}_{\mathbf{y}}(\omega) \cdot \mathbf{x} + \varepsilon_{\mathbf{y}} = \left(\sum_{i=1}^{p} \omega_i \cdot \mathbf{M}_{\mathbf{y},i} \right) \cdot \mathbf{x} + \varepsilon_{\mathbf{y}} \tag{8}$$

where p is the number of data groups resulting from the application of the different measurement techniques to the different compounds in the system.

2.4 *A flexible notation for isotopomer measurements*

The general measurement Equation (8) contains many high-dimensional but sparsely populated generalized measurement matrices $\mathbf{M}_{\mathbf{y},i}$. From the user's viewpoint, matrix input into a software system is always tedious work and a primary source of errors. For this reason an intuitive and easy to handle textual notation for the various types of measurements will be presented from which the measurement matrices $\mathbf{M}_{\mathbf{y},i}$ are automatically generated. This strategy has already been used for the generation of the carbon balance equations in Part I.

Generally, to specify a measurement equation together with the corresponding measured data the user has to supply a measurement name, a specification of the measured quantity and a measured value with standard deviation. Then all measurement values associated with the same scaling factor have to be put into one measurement group. For example to generate $\mathbf{M}_{\text{AlaC2}}$ it can be specified:

```
GROUP Ala_C2 :=
{
    Ala_Sing_2      := Ala#010= 0.157 +- 0.005
    Ala_DoubA_2     := Ala#110= 0.285 +- 0.005
    Ala_DoubB_2     := Ala#011= 0.122 +- 0.005
    Ala_DDoub_2     := Ala#111= 0.053 +- 0.005
}
```

The measurement value is given without scaling, i.e. in the case of MPs the peak areas are directly supplied. The measurement specification on the right side of each equation can be written by using one of the following notations:

1. The combined binary isotopomer/cumomer notation (e.g. Ala#0X1) allows each of the letters 0, 1 or X to be used with the meaning introduced in Part III. This combines the binary isotopomer notation (where only the symbols 0, 1 are allowed) and the binary cumomer notation (where only the symbols X, 1 are allowed).

2. The positional cumomer notation (e.g. Ala#13) can be used alternatively, which keeps the notation consistent with the PE notation in Part II.

3. The fragment weight notation (e.g. Ala#23(2)) is used to specify MS measurements. It is composed of the observed fragment given by the respective carbon atom positions 23 of the original molecule and the bracketed number (2) of the measured MI peak starting with zero for the lightest fraction.

Additionally, any linear combination of such terms can be specified giving the user full flexibility to express arbitrary measurement configurations. For example, the opposite carbon atoms numbered 5 and 9 in the benzene ring of phenylalanine cannot be disentangled by NMR, which is expressed (in the case of ^{1}H-NMR) by the PE sum

```
Phe_H2 := Phe#5+Phe#9    .
```

Obviously, this notation enables all the measurement types to be expressed in a very comprehensive way and in most cases only one term has to be written. Thus the risk of an erroneous input is rather small.

3 Statistical Evaluation and Experimental Design

As becomes evident from the previous section, an important structural difference between the statistical models in Parts II and IV is that another set of parameters had to be introduced by the vector of scaling factors ω in the measurement equations. This causes another model nonlinearity. The optimal estimation of the additional parameters is discussed in Appendix A. It is shown there that the approach of (Schmidt *et al.*, 1998a) which treats the additional parameters separately from the flux parameters leads to suboptimal results and can be improved by estimating all parameters simultaneously from one least squares minimization. As will be shown below the general treatment of these additional parameters is straightforward within the general framework. All the necessary software tools for model generation, equation solving, flux fitting and statistical analysis described below have been implemented by the authors so that the framework can be immediately put to work.

3.1 *The general statistical model*

All computational steps of the model from Part II can now be reestablished starting with the free flux vector Θ. The following mappings were introduced in Parts I-III:

$\mathbf{\Psi}$	computes the numerical flux coordinates $\mathbf{v}^{\text{net}}, \mathbf{v}^{\text{xch}[0,1]}$ from the free fluxes Θ
$\mathbf{\Phi}^{[0,1]}$	transforms $\mathbf{v}^{\text{net}}, \mathbf{v}^{\text{xch}[0,1]}$ to the application flux coordinates $\mathbf{v}^{\text{net}}, \mathbf{v}^{\text{xch}}$
$\mathbf{\Phi}$	transforms $\mathbf{v}^{\text{net}}, \mathbf{v}^{\text{xch}}$ to the natural flux coordinates $\mathbf{v}^{\rightarrow}, \mathbf{v}^{\leftarrow}$
$\mathbf{\Gamma}$	computes the cumomer labeling state vector $\mathbf{x}$ from $\mathbf{v}^{\rightarrow}, \mathbf{v}^{\leftarrow}$
$\mathbf{M}_w$	relates the measured net fluxes $\mathbf{w}$ to $\mathbf{v}^{\rightarrow}, \mathbf{v}^{\leftarrow}$
$\mathbf{M}_y\,(\omega)$	relates the isotopomer measurements $\mathbf{y}$ to $\mathbf{x}$

Compared to the model from Part II, only the structure of the mapping $\mathbf{\Gamma}$ and that of the measurement matrix $\mathbf{M}_y\,(\omega)$ has to be changed in this sequence. Consequently, the complete input-output relation from Part II has now been reestablished ($\circ$ denotes the concatenation of maps):

$$\begin{pmatrix} \mathbf{w} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} (\mathbf{M_w}, -\mathbf{M_w}) \cdot & \mathbf{\Phi} \circ \mathbf{\Phi}^{[0,1]} \circ \mathbf{\Psi}\,(\Theta) \\ \mathbf{M_y}\,(\omega) & \cdot \mathbf{\Gamma} \circ \mathbf{\Phi} \circ \mathbf{\Phi}^{[0,1]} \circ \mathbf{\Psi}\,(\Theta) \end{pmatrix} + \begin{pmatrix} \varepsilon_{\mathbf{w}} \\ \varepsilon_{\mathbf{y}} \end{pmatrix} \stackrel{def}{=} \begin{pmatrix} \mathbf{F_w}\,(\Theta) \\ \mathbf{F_y}\,(\Theta, \omega) \end{pmatrix} + \begin{pmatrix} \varepsilon_{\mathbf{w}} \\ \varepsilon_{\mathbf{y}} \end{pmatrix} \tag{9}$$

This is the general statistical regression model for isotopomer labeling experiments. Having computed a flux estimate $\hat{\mathbf{\Theta}}$ and a scaling factor estimate $\hat{\omega}$ by solving the weighted least squares problem

$$\min_{\Theta, \omega} ||\mathbf{F_w}\,(\Theta) - \mathbf{w}||^2_{\mathbf{\Sigma_w}} + ||\mathbf{F_y}\,(\Theta, \omega) - \mathbf{y}||^2_{\mathbf{\Sigma_y}} \tag{10}$$

the combined covariance matrix of the free fluxes and the scaling factors can be estimated by

$$\text{Cov}\,(\hat{\mathbf{\Theta}}, \hat{\omega})^{-1} = \begin{pmatrix} \frac{\partial \mathbf{F_w}}{\partial \Theta} & \mathbf{0} \\ \frac{\partial \mathbf{F_y}}{\partial \Theta} & \frac{\partial \mathbf{F_y}}{\partial \omega} \end{pmatrix}^T \cdot \begin{pmatrix} \mathbf{\Sigma_w}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma_y}^{-1} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial \mathbf{F_w}}{\partial \Theta} & \mathbf{0} \\ \frac{\partial \mathbf{F_y}}{\partial \Theta} & \frac{\partial \mathbf{F_y}}{\partial \omega} \end{pmatrix} \quad . \tag{11}$$

It is shown in Appendix B how the derivatives $\partial \mathbf{F_y} / \partial \Theta$ and $\partial \mathbf{F_y} / \partial \omega$ can be computed analytically by using matrix calculus. Usually the auxiliary parameters ω will be of no interest in practical application. For this reason the desired covariance matrix $\text{Cov}\,(\hat{\mathbf{\Theta}})$ can be immediately obtained by taking the upper left submatrix of $\text{Cov}\,(\hat{\mathbf{\Theta}}, \hat{\omega})$ with dimension $\dim\,\Theta \times \dim\,\Theta$.

Once the covariance matrix has been computed all the statistical procedures introduced in Part II, like the computation of single parameter and simultaneous confidence regions, can be carried out. This has been done recently for several different isotopomer experiments and will be presented elsewhere (Petersen *et al.*, 1998). The following chapters concentrate on the optimal design of labeling experiments.

3.2 *Optimal experimental design*

As has been explained in the Introduction, the covariance matrix $\text{Cov}\,(\hat{\mathbf{\Theta}})$ depends on some parameters that can be influenced by the experimenter:

1. the isotopomer mixture composing the input substrate as specified by $\mathbf{x}^{\text{inp}}$,
2. the type and amount of measurements as specified by $\mathbf{M_y}$.

Clearly, these parameters cannot be arbitrarily chosen, because there are various constraints set by the experimental facilities and the commercial availability of isotopomers. The realizable measurement matrices and mixtures are henceforth called *feasible* design parameters.

In order to compare two or more designs the D-optimality criterion is used in this text as explained in the Introduction. However, other criteria (Pázman, 1986) can be directly built into this framework because none of the computational procedures specifically depends on the structure of the D-criterion.

Choosing the D-criterion means that the squared volumes of the confidence ellipsoids for a given significance level are compared between two experimental designs. This squared volume is given (up to a constant that does not depend on the design parameters) by

$$D\left(\mathbf{x}^{\text{inp}}, \mathbf{M_y}; \Theta\right) = \det \text{Cov}\left(\mathbf{x}^{\text{inp}}, \mathbf{M_y}; \Theta\right) \quad .$$

This D-optimality criterion has to be minimized with respect to the space of all feasible input mixtures $\mathbf{x}^{\text{inp}}$ and measurement matrices $\mathbf{M_y}$:

$$D_{\text{opt}}\left(\Theta\right) = \min_{\mathbf{x}^{\text{inp}}, \mathbf{M_y} \text{ feasible}} D\left(\mathbf{x}^{\text{inp}}, \mathbf{M_y}, \Theta\right)$$

Obviously, D_{opt} depends on the true value of Θ. Thus an approximate value for the parameter vector Θ has to be assumed a priori. In the following it will be shown that this procedure is well-justified because the optimal design is rather insensitive with respect to even large variations in Θ.

As an extension to D-optimality, a restriction to a set of linear combinations of the free fluxes can be made by minimizing the more general $D_{\mathbf{L}}$-criterion

$$D_{\mathbf{L}} = \det \text{Cov}\left(\mathbf{L} \cdot \Theta\right) = \mathbf{L} \cdot \text{Cov}\left(\Theta\right) \cdot \mathbf{L}^T \tag{12}$$

with an appropriate matrix $\mathbf{L}$. This is called a partial optimal design (Pázman, 1986). For example $\mathbf{L}$ might select a certain free flux, a group of free fluxes or a flux that is linearly dependent on the free fluxes. More details on the numerical computation of the D-criterion are given in Appendix B. The practical use of the optimal design procedure will be illustrated below by an example.

4 An Example

All the concepts presented above are now put into practice by using the network of central metabolic pathways of *Corynebacterium glutamicum* previously illustrated in Part II. This saves a lengthy specification of networks, parameters and assumptions and only the new details need to be given here.

4.1 *Metabolic network and free fluxes*

The system contains a total of 23 free fluxes and 13 of these fluxes (i.e. the biomass effluxes) were directly measured. It is not surprising that the confidence intervals of the corresponding 13 estimated values are almost the same as those of the corresponding measurements. Consequently, labeling data have almost no influence on the estimation quality of these fluxes so that the biomass effluxes are not interesting for an experimental design study. For this reason the following investigation concentrates on the remaining 3 *essential net fluxes*

ppp_1^{net}	pentose phosphate pathway influx
lp_2^{net}	lysine production via the ddh pathway
gc^{net}	glyoxylate shunt flux

and the 7 *essential exchange fluxes*

$gly_1^{\text{xch[0,1]}}, gly_3^{\text{xch[0,1]}}$	in the glycolytic pathway
$ppp_2^{\text{xch[0,1]}}, ppp_3^{\text{xch[0,1]}}, ppp_4^{\text{xch[0,1]}}$	in the pentose phosphate pathway
$ac_1^{\text{xch[0,1]}}$	in the anaplerotic section
$cac_4^{\text{xch[0,1]}}$	in the citric acid cycle .

As in Part II the anaplerotic section is represented by only one flux between a lumped PEP- Pyr pool and a lumped OAA-Mal pool. The reason is that without this assumption certain nonidentifiable fluxes would occur (see Part II and (Wiechert, 1995)) which would make it difficult to compare the different approaches. The detailed treatment of the anaplerotic flux identifiability problem is well within the scope of the methods presented here but will be carried out in a later publication.

It is assumed that the free fluxes estimated in Part II (under lysine-producing conditions) are indeed the true flux values. Starting with this assumption it will now be investigated which type of experiment would have been best to estimate these fluxes. Particular interest is attached to the pentose phosphate pathway influx estimate ppp_1^{net} which was only determined within a 90% confidence interval of ± 12 (relative to the glucose uptake rate 100), which may not be good enough to discriminate two different genetically engineered strains from each other.

In Part II it was not necessary to model the amino acid biosynthesis pathways in detail because each PE measurement could be immediately related to a precursor carbon atom in central metabolism. This is no longer true for MP and MI measurements because some amino acids (like phenylalanine, valine or leucine) are composed of more than one precursor molecule. In this situation an MP or MI fraction cannot be directly traced back to the single precursors. For this reason the amino acid synthesis pathways for serine, alanine, phenylalanine, valine, leucine, isoleucine, proline, threonine and aspartate have been incorporated into the network.

4.2 *Measurements*

Three sources of isotopomer measurement data with typical measurement errors are now compared for the *C. glutamicum* example. Some details on the origin of the assumed data sets are given now:

PE data set: The positional enrichments in Table 1 are assumed to be measured by ^{1}H-NMR. They were taken from Part II but here the various measurement errors are assigned to typical values rather than the specific outcomes from Part II. However the standard deviations are always in the order of magnitude of those values published in Part II. This results in a total of 26 PE values.

MP data set: The multiplet data set from Table 1 obtained by 2D-NMR is essentially taken from (Szyperski, 1995) but has been extended by some recently published data (Schmidt *et al.*, 1998a; Petersen *et al.*, 1998). It should be noticed that some of the measurements (like those for Ri5p#1) can only be obtained from RNA/DNA hydrolysate. However they have been included to obtain a fair comparison. The complete set of all published MP signals contains many obvious redundancies, e.g. when the pyruvate successors alanine, valine, leucine and isoleucine are all measured. These redundancies would also arise if positional enrichments were considered. Because the basic attitude of the present investigation is to concentrate on the essential information these redundant data sources are not taken into consideration. This results in 68 MP data values within 21 groups, i.e. the net number of possibly informative values after MP scaling is 68-21=47.

Unfortunately, no measurement standard deviations for multiplet peaks are given in the literature although these parameters can be derived from the signal to noise ratio. These were therefore estimated from the published spectra. If in doubt, a best case assumption for the MP measurements was made by assigning rather optimistic values to the standard deviations. Overall, the values have error bars in the same order of magnitude as those assumed for the PE peaks, which is also justified by the remarks made in the introduction. The influence of these assumptions on the quality of the estimated fluxes will be investigated in detail in the last Section before the Conclusion by performing a variation study.

MI data set: The most extensive mass isotopomer data set currently known to the authors is published in (Christensen & Nielsen, 1999) and has been kindly supplied by Bjarke Christensen from Lyngby, Denmark (Table 2). Knowledge of which molecular fragments can currently be observed by MS is taken from this source. Again only those MI data values that are not obviously redundant are considered. This produces 90 MI data values in 20 groups, i.e. the net number of possibly informative values is 70.

As explained in the Introduction MS is a virtually exact instrument. From some repeated runs in (Christensen & Nielsen, 1999) it can be deduced that the measurement error of the MS instrument alone is about 0.2% on the fractional enrichment scale. However due to the mentioned error sources (washout correction, sample preparation, isotope effects) a larger error of 0.4% is assumed here for all MI values. This might be a rather pessimistic assumption and the sensitivity of the flux estimation quality with respect to this assumption is studied later in the last section before the Conclusion.

4.3 *Interpretation of D-optimality criteria*

A further remark should be made on the presentation of the results. If only a few experiments have to be compared the standard deviations for all 10 essential flux estimates are presented by using bar charts. In the case of an experimental design study with many different experiments this is not possible and the D-optimality values of the respective designs are used. Usually these are very small numbers which are hard to interpret in an intuitive manner. In the example discussed below the D-values have magnitudes in the order of 10^{-120}. Improvements in the design can quickly change this value by a factor of 10^{10}, which is quite impressive at first glance but should be interpreted in the right way.

In order to obtain an easier intuitive interpretation of these D-values the following transformations have been made:

1. The focus is restricted to the essential flux estimates by inspecting the respective 10×10 submatrix of Cov (Θ) in Equation (11). This results in a partial optimal design criterion $D_{\mathbf{L}}$ where $\mathbf{L}$ is the projection to the essential coordinates (see Equation (12)).

2. Imagine that the resulting 10-dimensional confidence ellipsoid is a sphere. Then the radius of this sphere is given up to a constant factor by $\sqrt[20]{D}$, i.e. the geometric mean of the principal axis length. This value can be interpreted as an average confidence interval length over all estimated fluxes.

3. Finally, a *reference experiment* is defined which is the ^{1}H-NMR experiment from Part II with 1-glucose as input and the error bars from Table 1. Its partial D-optimality value is $D_{\text{ref}} = 1.5 \cdot 10^{-24}$, i.e. the average

axis length is $\sqrt[20]{\overline{D}_{\text{ref}}} = 0.06$. All D-values are then related to this reference, i.e. the quality measures that are finally used in Figure 4 and 6 are:

$$\overline{D} = \sqrt[20]{D/D_{\text{ref}}} \qquad \text{and} \qquad \overline{I} = 1/\overline{D} \quad .$$

As a numerical example the information value $\overline{I} = 1.5$ means that the corresponding experiment will yield an average 1.5-fold shrinking of the parameter confidence region compared to the reference experiment. Roughly speaking the experiment is 1.5 times more informative than the reference. Clearly, the same effect is also produced if only one axis shrinks by a factor of $1.5^{20} = 3325$, which can only be discovered by inspecting the respective error bar plots.

4.4 *Guiding problems*

The following investigations were guided by some central questions:

1. Which of the three data sources gives the best results?
2. What is the optimal input substrate mixture in each case?
3. Which experiment yields the best estimation of the pentose phosphate pathway influx and of the split lysine production pathway ratio?

5 Optimal Input for Positional Enrichment Measurements

The investigation is started with the comparison of experimental designs when only PE data are available, as was the case in Part II. In this special situation some theoretical results can be proven that do not apply to the general case.

5.1 *A theoretical result*

Clearly, in this situation only the PEs within an input mixture influence the outcome of the experiment. By denoting with ${}^1\mathbf{x}^{\text{inp}}$ the subvector of PEs in the input cumomer fraction vector $\mathbf{x}^{\text{inp}}$ (see Part III) this means that

$$D\left(\mathbf{x}^{\text{inp}}, \mathbf{M}_{\mathbf{y}}^{\text{PE}}; \Theta\right) = D\left({}^1\mathbf{x}^{\text{inp}}, \mathbf{M}_{\mathbf{y}}^{\text{PE}}; \Theta\right) \quad .$$

An additional condition for the following Theorem is that each value of ${}^1\mathbf{x}^{\text{inp}}$ is feasible, i.e. arbitrary fractions $\mathbf{0} \leq {}^1\mathbf{x}^{\text{inp}} \leq \mathbf{1}$ for all carbon atoms ${}^1\mathbf{x}_i^{\text{inp}}$ can be realized in practice. Under these conditions it has been proven in (Kownatzki, 1998):

Theorem: For a fixed PE measurement matrix $\mathbf{M}_{\mathbf{y}}^{\text{PE}}$ it holds:

1. $D\left({}^1\mathbf{x}^{\text{inp}}, \mathbf{M}_{\mathbf{y}}^{\text{PE}}; \Theta\right) = D\left(\mathbf{1} - {}^1\mathbf{x}^{\text{inp}}, \mathbf{M}_{\mathbf{y}}^{\text{PE}}; \Theta\right)$ with the vector $\mathbf{1}$ composed from all ones.
2. $D\left(\lambda \cdot {}^1\mathbf{x}^{\text{inp}}, \mathbf{M}_{\mathbf{y}}^{\text{PE}}; \Theta\right)$ increases monotonously with respect to $\lambda \geq 0$.
3. $D\left(\lambda \cdot \mathbf{1}, \mathbf{M}_{\mathbf{y}}^{\text{PE}}; \Theta\right) = \infty$ for $0 \leq \lambda \leq 1$.

These statements have a simple intuitive explanation:

1. If each ^{13}C in the system is replaced by a ^{12}C (i.e. ${}^1\mathbf{x}^{\text{inp}}$ is replaced by $\mathbf{1} - {}^1\mathbf{x}^{\text{inp}}$) then nothing changes because ^{1}H-NMR always measures both fractions. Both experiments are subsequently called *mirror images* of each other.
2. It can easily be shown that the PE subvector ${}^1\mathbf{x}$ of $\mathbf{x}$ depends linearly on ${}^1\mathbf{x}^{\text{inp}}$. The more labeled material is put into the system (i.e. λ is raised) the higher is the intracellular positional labeling and thus all measurement sensitivities are improved.
3. If all positions of the input substrate have the same PE λ (i.e. ${}^1\mathbf{x}^{\text{inp}} = \lambda \cdot \mathbf{1}$) then all intracellular carbon atom pools will finally have the same PE λ whatever the fluxes in the system are. Thus there is no flux information in this experiment. This is e.g. the case with a mixture of unlabeled and uniformly labeled substrate as used in (Szyperski, 1995).

For an input substrate with only two carbon atoms the statements of this Theorem are illustrated in Figure 2a. It becomes clear that in this special situation there are exactly two optimal experiments with the input PEs $(1,0)^T$ and $(0,1)^T$ which are mirror images of each other. In higher dimensions the optimal solution is still restricted to a small subset of the design cube $\mathbf{0} \leq {}^1\mathbf{x}^{\text{inp}} \leq \mathbf{1}$. More precisely it follows from the Theorem:

Corollary: For PE measurements it holds:

1. If ${}^1\mathbf{x}^{\text{inp}}$ is optimal then one of its coordinate entries must be 1 and another must be 0.
2. Adding unlabeled or uniformly labeled isotopomers to any given input substrate mixture always decreases the available flux information.

In three dimensions this restricts ${}^1\mathbf{x}^{\text{inp}}$ to either $(\lambda,0,1)^T$, $(0,\lambda,1)^T$ or $(0,1,\lambda)^T$ with $0 \leq \lambda \leq 1$ or one of the mirror images (Figure 2b).

5.2 *Optimal input for unrestricted input substrate mixture*

Although the Theorem does not completely solve the optimal design problem it suggests good starting points and constraints in the search for the optimal solution by a nonlinear optimization algorithm. In particular the corners of the higher dimensional design cube $\mathbf{0} \leq {}^1\mathbf{x}^{\text{inp}} \leq \mathbf{1}$ are good candidates. Removing the mirror images and the informationless corners this means that $2^{n-1}-1$ D-values have to be computed where n is the number of carbon atoms in the input. For a glucose input ($n=6$) the results are shown in Figure 3. The computational effort to produce this Figure is only in the order of a few minutes, which shows the efficiency of the numerical methods presented in Part III.

By successively starting with each of the cube corners an optimization algorithm has a good chance of finding a global optimum. Interestingly, for all examples where this was done it turned out that the optimal design was actually one of the cube corners even when the starting value was randomly chosen (see also Figure 4). This "corner rule" can serve as a rule of thumb for future computations.

Figure 3 shows that except for some very badly designed experiments the choice of the input substrate can indeed improve the average confidence interval by a factor of 3. The best substrate turns out to be 2,3,5-glucose (or its 1,4,6-labeled mirror image), which unfortunately is not available in practice like most other substrates with higher molecular weights. This problem will be discussed in the next section. In general there is a rough correlation between the improvement factor and the number of labeled carbon atoms as can also be seen in Figure 3.

The best singly labeled substrate is 2-glucose, which yields a 1.4 times improved result compared to the reference input 1-glucose. Even 3-glucose produces a 1.2-fold improvement. However, from a practical viewpoint these improvements are not dramatic so that at least almost the best realizable experiment was carried out in Part II.

The last entries in Figure 3 show that a total of 16 possible input substrates are completely ruled out because some or all of the parameter standard deviations are tremendously large. They are all characterized by three labeled (or unlabeled) carbon atoms at position 1,2,3. Because the 4,5,6-fragment passes unaltered through the whole pentose sugar metabolism and the glycolysis as well it becomes clear why these substrates perform so badly.

5.3 *Optimal feasible input*

The investigation of the design cube corners is always a good starting point for an experimental design study because it gives an impression of what can be achieved under the best possible conditions. However, in order to compute a practically meaningful design the input vector ${}^1\mathbf{x}^{\text{inp}}$ must be restricted to the set of feasible inputs. This is given by mixtures of a limited number of available isotopomers. Unlabeled and uniformly labeled substrates are always ruled out as mixture components by the Corollary. For this reason the mixture of 1-, 2- and 3-glucose (which turned out to be the most informative singly labeled substrates) is now considered as an example.

A table of D-values for the set of all possible three component mixtures can be easily computed. The result is graphically represented in Figure 4 by a mixture diagram. It turns out that the information optima are exactly all three corners of the mixture triangle. A small local optimum can be found on the 1-glucose-3-glucose edge of the mixture triangle. This is clear from the Theorem because 100% 1,3-glucose yields a 1.5-fold improvement. Likewise, the global minimum at the triangle center yields zero information because 1,2,3-glucose is uninformative.

As a result it can be concluded that the mixing of input substrates is not a good idea for PE experiments. On the other hand, this shows that the investigation of all design cube corners as done before is a good starting point to find the best practically realizable input substrate.

5.4 *Robustness of the design*

It has been pointed out before that nonlinear optimal experimental designs always depend on an a priori guess for the true parameter Θ. The extent to which this guess influences the optimal solution has been investigated. To this

end, all free fluxes were sequentially varied within $\pm 50\%$ in each coordinate direction, which is quite a large range and in practice will undoubtedly cover the true parameters. As a first result it turned out that the D-optimality values indeed depend on the free fluxes. Quantitatively speaking the essential confidence intervals change their size up to a factor of 3 when the free fluxes are varied (data not shown).

Fortunately the relative $\overline{D}$-criteria remain approximately constant whatever the varied fluxes are. This means that the ranking of the different experiments with respect to the achievable flux information always remains the same except for minor position changes in the ranking that are not practically relevant. Thus it can be assumed based on this empirical finding that optimal designs for labeling experiments are very robust with respect to the parameter guess. This at last justifies the practical application of the experimental design method.

5.5 *Partial optimal designs*

Up to now only the $\overline{D}$-criterion has been used as an average information measure over all the essential fluxes. If a specific flux or a certain section of the metabolic network is of interest, the confidence intervals have to be inspected in more detail. This is now done for the single parameter standard deviation of the pentose phosphate pathway influx.

Figure 3 also shows the improvement of this specific error bar compared to the reference experiment. Although there is a rough correlation between the overall performance and the ppp_1^{net} standard deviation it can be seen that there is one outstanding experiment with 1,2,4-glucose yielding a 5.3-fold improvement. The best feasible experiment with 2-glucose yields a 3.8-fold improvement. Thus the problem of an unsatisfactory pentose phosphate pathway influx estimate would have been overcome with this substrate labeling.

Finally, Figure 5 shows a detailed comparison of standard deviations for some interesting PE experiments. It becomes clear that certain substrates specifically enhance the estimation quality of certain fluxes. As soon as more than one flux is required to be well estimated in addition to ppp_1^{net} the best design is almost exactly the same as the best design for all essential fluxes. In particular this holds true if emphasis is laid only on the 3 net fluxes or on the 7 exchange fluxes in the system. This once more demonstrates that net flux estimation cannot be decoupled from the estimation of exchange fluxes, although the latter are not of primary importance in flux analysis.

6 Optimal Input for General Isotopomer Measurements

In the case of general isotopomer experiments the restriction to a relatively low-dimensional design cube can no longer be made. Thus an inspection of the input design space by computing a table of values is not practicable. For this reason the inputs should be restricted to the set of feasible inputs by specifying input substrate mixture components. The primary analysis of the PE experiment may help in the choice of the mixture candidates. In the present example these candidates are the reference isotopomer 1-glucose, the unlabeled and the uniformly labeled isotopomer. 1-glucose was chosen for a better comparison with the literature although the 2-glucose input performed slightly better in the PE example. For example a 90:10 mixture of unlabeled and uniformly labeled glucose was used for the MP measurements in (Szyperski, 1995) while a 90:10 mixture of 1-labeled and uniformly labeled glucose was used in (Schmidt *et al.*, 1998a). Such a mixture should also be a good substrate for mass isotopomer measurements which require a sufficient amount of labeled carbon atoms to produce sensitive signals.

6.1 *Comparison of measurement techniques*

The $\overline{D}$-criterion was computed for the MP and MI data set and for all mixtures of the three components (Figure 6). For the PE experiment it is already known by the Corollary that a pure 1-glucose input will perform best. Inspecting the plots yields some interesting observations:

1. It is intuitively clear that MP measurements can identify the fluxes because this is already achieved with PE measurements. On the other hand there is no intuitive feeling about the potential of MI data because MS is not sensitive to specific carbon atom positions within the intracellular molecules. Figure 6b now proves that MI measurements alone are sufficient to obtain a complete flux quantification.

2. The reference experiment can be improved with MP measurements as well as with MI measurements. However the best MP experiment (improvement factor 1.3) and the best MI experiment (improvement factor 1.6) yield almost the same improvement as the 2-glucose PE experiment (factor 1.4). Thus there is no practically relevant difference between all three experimental methods.

3. The best MP experiment (the 50:5:45 mixture of uniformly labeled, unlabeled and 1-labeled glucose) and the best MI experiment (40:12:48) perform quite similarly which indicates that these two measurement techniques can be well combined. On the other hand, the PE optimum is always on the edges of the triangles (Figure 4). This indicates that the additional incorporation of PE data into an MP or MI experiment will not significantly improve the results.

4. It is not a good idea to use a mixture containing only unlabeled and uniformly labeled glucose with MP experiments. The 0:10:90 mixture used in (Szyperski, 1995) will enlarge the error bars of the reference experiment by a factor of 2.5 while the more expensive optimal 0:60:40 mixture still has a factor of 1.3. On the other hand the experiment from (Schmidt *et al.*, 1998a) with a 90:0:10 mixture lies "on the other side" of the optimum and also yields 1.3. The optimum is 50:5:45, which improves the error bars by a factor of 0.8.

5. Interestingly the MI experiment almost fails to identify the fluxes if only unlabeled and fully labeled glucose is used. This will be explained in the next paragraph by a detailed inspection of the error bars.

6.2 *Detailed comparison of error bars*

Figures 7 and 8 show a detailed comparison of some specific experiments taken from the mixture plots in Figure 6. The following results can be found:

1. The MP data set (Figure 7) with a mixture containing only unlabeled and fully labeled glucose almost fails to identify the pentose phosphate pathway fluxes, which is the main reason for the poor overall performance.

2. The MI data set (Figure 8) with a mixture of unlabeled and fully labeled glucose yields extremely poor results because the distinction between the two lysine production pathways completely fails. This can immediately be understood because these pathways run in parallel with a different fate of the carbon atoms coming from pyruvate and oxalo acetate. These positional changes of the carbon atoms can be well observed with the positionally sensitive PE and MP experiments (Sonntag *et al.*, 1993) while MS cannot detect a positional change of labeled atoms because this does not influence the molecular weight.

3. The best MI experiment produces at least a poor (but still practically useless) discrimination between the two lysine producing pathways. The reason is that not the complete lysine molecule was measured but only the 2-6-fragment. Thus there is at least some low sensitivity to positional changes of carbon atoms (Figure 8).

4. Interestingly, the best MI discrimination of the lysine pathways is obtained with a pure 1-labeled input. However the pentose phosphate pathway is badly determined with this substrate.

If the best MP experiment is compared with the best MI experiment and the best feasible PE experiment as is done in Figure 9 it becomes clear that each of the experiments has its specific strength:

1. If only the pentose phosphate pathway is of interest the PE experiment performs better than both the MP and MI experiments.

2. MI experiments are unable to distinguish the split lysine pathways but produce the best exchange flux estimates.

3. MP experiments are worst when the pentose phosphate influx and the glycolytic exchange fluxes have to be estimated.

6.3 *Influence of assumed measurement errors*

As has been said before, measurement standard deviations for the MP and MI measurements could not be obtained from literature and thus had to be assumed. To investigate the influence of the assumed values all standard errors for MPs or MIs were simultaneously varied by a factor between 0.5 and 2.0. The result is that the overall quality changes with approximately the same factor. For example if the assumed error bars are halved, the estimation quality doubles (which is not clear a priori because the biomass efflux errors are still kept fixed).

Because the error bars for the PE and MP data sets are rather typical from the already available data sets it is unlikely that large improvements can be achieved with these types of experiments compared to the results presented here. On the other hand, it is possible that the real potential of MI data is higher. Thus if the MI measurement error bars are assumed of half the size – which is believed to be the upper bound by the authors (cf. also (Christensen & Nielsen, 1999)) – the MI experiment will outperform all other experiments. However, the discrimination of the lysine fluxes is still not sufficient and, moreover, the outcome is quite sensitive to the input mixture composition.

7 Conclusions

In this contribution the statistical and computational foundations for evaluating general isotopomer experiments have been laid and a universal framework for the design of experiments has been established and implemented as a software system. The statistically optimal formulation of the measurement equations turned out to be a

critical point which has not been recognized hitherto. The framework now offers a large variety of possibilities for experimental design as was demonstrated by the example.

The investigation of the example enabled the different experimental setups to be quantitatively compared. Without the extensive performance of simulation studies, as in the example discussed above, it is impossible to intuitively judge the potential of carbon labeling experiments. These studies could be carried out because the flexibility of the developed software system and the computational efficiency of the implemented algorithms is extremely high (see Appendix of Part III). Although the true flux values must be assumed in a design study it turned out that this method is of practical relevance because the outcome is insensitive with respect to the assumptions made.

All three experimental techniques (^{1}H-NMR, ^{13}C-NMR or MS) can in principle identify the intracellular fluxes and none of these approaches is generally superior to the others. On the other hand, it turned out that this discussion is not fruitful for future developments because there are always certain fluxes (e.g. pentose phosphate pathway influx or split lysine production pathways) which are best identified by one of the three methods. This shows that each method will have its own application field in the future and a combination may be the only way in some situations.

The input substrate composition is very critical for the design of an informative labeling experiment (especially for MI data). Moreover the experimental design may lead to unexpected results. For example, the inferiority of 1-^{13}C glucose to other singly labeled species in the PE experiment was not recognized before. In many situations the achieved information improvement well justifies the cost of some more expensive labeled substrates. On the other hand the combination of the measurement techniques is not always advisable. While MP and MI experiments will complement each other the PE experiment is rather antagonistic to the others. This means that an informative experiment for MP and MI is an uninformative one for PE and vice versa.

For the design of an optimal experiment for a certain application problem no general recipes can be given. In contrast, the best experiment depends on what the focus of interest is. In particular the optimization criterion has to be chosen appropriately. More precisely, an experimental design must always be an iterative procedure. An optimality criterion is defined first which is then optimized by a complete inspection of the PE design cube edges (Figure 3), the whole design space (Figure 4) or by a numerical optimization procedure. This gives a rough impression about the performance of different experiments. The achieved single parameter confidence intervals then reveal their specific strengths and weaknesses. Their inspection will lead to a modified criterion (e.g. by giving up the discrimination of the two lysine production fluxes) and possibly to another optimal mixture. From this result the whole procedure might start anew.

An important problem for the future will be the assignment of reasonable measurement error bars for MS experiments. To this end more theoretical and experimental studies are required. Even if it turns out that MI data really have an outstanding quality it must be pointed out that the assumption of very small error bars will almost surely produce strong inconsistency problems for fitting the large data sets. These inconsistencies stem from the simplifying assumptions that are still present in the network formulation and will undoubtedly make data evaluation more difficult.

With the extension of the statistical framework for carbon labeling experiments from Parts I and II the theory, the experimental procedures, and the evaluation methods have now reached a certain state of maturity. Software tools to make flux determination by labeling experiments an affordable routine procedure have been supplied. Some further work has to be done from the viewpoint of user interfacing because the complexity of the whole experiment specification can currently not be handled by an inexperienced person. For this reason an integrated environment with expert components that help the user to specify the details of his experiment is now under development. A prototype of this system is now running in a spreadsheet environment. Finally, the iterative experimental design procedure described above will be supported within the environment.

A Optimal Estimation of Peak Scaling Factors

In (Schmidt *et al.*, 1998a) an estimation procedure for the peak scaling factors has been suggested which is different from the one given by Equation (10). Both approaches are now compared. For simplicity it is assumed here that only one metabolite pool is present in the system and only one group of measurement data with the same scaling factor is present. Additionally, only one flux v has to be estimated in this simple example. The general measurement model from Equation (8) then reduces to:

$$\mathbf{y} = \omega \cdot \mathbf{M_y} \cdot \mathbf{x}\,(v) + \varepsilon_\mathbf{y} \quad \text{with} \quad \text{Cov}\,(\varepsilon_\mathbf{y}) = \mathbf{\Sigma_y} \tag{13}$$

with a scalar factor ω. The difference between the approaches to estimate v from the data vector $\mathbf{y}$ can now be made clear:

1. In (Schmidt *et al.*, 1998a) the parameters ω and v are estimated in succession. To this end, assume for the moment that v is already known. In this case ω can be easily estimated from any single equation in the Equation system (13) by dividing a measured value $\mathbf{y}_j$ by its predicted unscaled value $[\mathbf{M_y} \cdot \mathbf{x}\,(v)]_j$. However this would lead to a strong propagation of the corresponding measurement error $[\varepsilon_\mathbf{y}]_j$. For this reason an averaging over all values was performed thus producing the estimate:

$$\hat{\omega} = \hat{\omega}\,(v) \approx \frac{\sum_j [\mathbf{y}]_j}{\sum_j [\mathbf{M_y x}\,(v)]_j} = \frac{\mathbf{1}^T \cdot \mathbf{y}}{\mathbf{1}^T \cdot \mathbf{M_y} \cdot \mathbf{x}\,(v)}$$

   Here the vector $\mathbf{1}^T = (1, 1, \ldots, 1)^T$ is introduced to formalize the sum operation. Replacing ω in Equation (13) by its estimate and dividing by $\mathbf{1}^T \cdot \mathbf{y}$ the following is now obtained:

$$\frac{\mathbf{y}}{\mathbf{1}^T \cdot \mathbf{y}} = \frac{\mathbf{M_y} \cdot \mathbf{x}\,(v)}{\mathbf{1}^T \cdot \mathbf{M_y} \cdot \mathbf{x}\,(v)} + \frac{\varepsilon_\mathbf{y}}{\mathbf{1}^T \cdot \mathbf{y}}$$

   This transformed model normalizes both the measurements $\mathbf{y}$ and the predicted measurements $\mathbf{M}_\mathbf{y}\,\mathbf{x}\,(v)$ by dividing by their respective sum so that the resulting equation no longer contains ω. At the same time the standard deviations of the error terms are rescaled by a random variable which is the critical operation in this approach. Assuming that the variance of the mean value $\mathbf{1}^T \cdot \mathbf{y}$ can be neglected compared to $\varepsilon_\mathbf{y}$ it approximately holds:

$$\text{Cov}\left(\frac{\varepsilon_\mathbf{y}}{\mathbf{1}^T \cdot \mathbf{y}}\right) \approx \frac{1}{(\mathbf{1}^T \cdot \mathbf{y})^2} \cdot \Sigma_\mathbf{y}$$

   A least squares fit then leads to an estimate for v:

$$\hat{v} \quad = \quad \arg\,\min_v\, ||\frac{\mathbf{y}}{\mathbf{1}^T \cdot \mathbf{y}} - \frac{\mathbf{M_y} \cdot \mathbf{x}\,(v)}{\mathbf{1}^T \cdot \mathbf{M_y} \cdot \mathbf{x}\,(v)}||^2_{\Sigma_\mathbf{y}/(\mathbf{1}^T \cdot \mathbf{y})^2} \tag{14}$$

   with the weighted norm $||\xi||^2_\Sigma = \xi^T\,\Sigma^{-1}\,\xi$ that has been already used in Part II.

2. The second approach is much more straightforward, but on the other hand involves more computational effort. Using a standard nonlinear regression approach both parameters ω and v are estimated simultaneously by fitting the data:

$$\begin{pmatrix} \hat{\omega} \\ \hat{v} \end{pmatrix} \quad = \quad \arg\,\min_{\omega, v}\, ||\mathbf{y} - \omega \cdot \mathbf{M_y} \cdot \mathbf{x}\,(v)||^2_{\Sigma_\mathbf{y}} \tag{15}$$

An in-depth analysis of the two alternative flux estimators cannot be presented here because too many mathematical preliminaries would be required. In general the well-known Gauss-Markov theorem (Arnold, 1990) states that (up to linearization) the general regression approach from Equation (15) will always produce the best possible estimate. Thus it has to be expected that the estimate from Equation (15) will always perform better than the one from Equation (14).

To quantify the difference Monte Carlo simulation runs were performed by generating 10000 different experiments with normally distributed measurement errors ε. The resulting estimates for both approaches were then computed for each run. Afterwards the two estimators were characterized by the sample mean and standard deviation over all runs. This procedure was carried out for different assumptions concerning the measurement standard errors. In each case both approaches were almost unbiased but had quite different variances. The approach from Equation (15) was always better than that from Equation (14), by a factor of 1.2 in typical situations and 10.0 in exceptionally bad situations depending on the assumed error variances. The difference is most significant when bad signal-to-noise ratios are present and this is even true when only a few such imprecise values are present. This situation occurs quite often in a labeling experiment when some peak sizes are close to zero and thus can hardly be distinguished from the ground noise.

In summary, it is always advisable to take the more rigorous approach from Equation (15) to avoid possible misbehavior of the estimation procedure. The consequence is that the scaling factors ω_i must always be incorporated as auxiliary variables into the model and estimated together with the free fluxes.

B Numerical and Implementation Details

The compiler program for the automatic generation of all system matrices and vectors from the textual input that was described in Part I and II has been extended to isotopomer systems. The main task was to generate the isotopomer transition matrices and the measurement matrices. A sparse matrix implementation was used to implement all the 2- and 3-dimensional matrix structures. Internally the program does all computations by using the cumomer coordinate system. These coordinates are transformed back to isotopomer fractions if this is necessary for interfacing with the user.

The numerical solution of the central cumomer balance equations is achieved by applying the cascaded solution algorithm presented in the Appendix of Part III. All linear equations are solved by using a QR factorization thereby achieving a numerically stable solution even for very large exchange fluxes. It should be mentioned that a solution can even be achieved for the limiting case of an infinite exchange flux because the preconditioning method that was developed in (Wiechert, 1996) can be directly applied to the cascaded system. For the complex central metabolism example the computation of $\mathbf{x}$ from $\mathbf{v}$ takes about 2 seconds on a 166 MHz Pentium processor.

Finally, the different mappings in Equation (9) have to be derived. Most of these computations are the same as in Part II. It has already been explained in the Appendix of Part III how $\partial\mathbf{x}/\partial\mathbf{v}$ can be computed with very low computational effort (i.e. below 1 second). Additionally, the general measurement Equation (8) must be derived with respect to $\mathbf{x}$ and ω. The result is:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \mathbf{M}_{\mathbf{y}}(\omega) \qquad \text{and} \qquad \frac{\partial \mathbf{y}}{\partial \omega_i} = \mathbf{M}_{\mathbf{y},i} \cdot \mathbf{x} \quad .$$

The complete computational procedure has been tested with many different checks which all were passed successfully:

1. comparison with the analytical solution for the example system in Part III,
2. comparison with the outcome of two completely independent implementations from (Schmidt *et al.*, 1997) and (Wurzel, 1997) for the pentose phosphate pathway network,
3. comparison with the outcome of an iterative solution algorithm based on (Wiechert *et al.*, 1997b),
4. alternative computation of derivatives by numerical differentiation.

After the parameter covariance matrix Cov (Θ) has been computed, the D-criterion is obtained from the singular value decomposition that is also the basis for principal axis computation. A range-restricted evolutionary algorithm (Bäck, 1996) was used for searching the global optimum in the design cube for the PE example.

References

Arnold, S.F. 1990. *Mathematical Statistics*. Ellis Horwood.

Bäck, T. 1996. *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. Oxford University Press.

Chatterjee, S., & Hadi, A.S. 1988. *Sensitivity Analysis in Linear Regression*. Probability and Mathematical Statistics. Wiley.

Christensen, B., & Nielsen, J. 1999. Isotopomer Analysis using GC-MS. *Journal of Metabolic Engineerig*. In press.

de Graaf, A.A., Wendisch, V., & Sahm, H. 1996. Full ^{13}C isotopomer quantitation of metabolic intermediates using ^{1}H NMR. *Page 135 of: Int. Conf. on Magnetic Resonance in Biological Systems, August 18-23, Keystone, Colorado*.

Donato, L. Di, Rosiers, C. Des, Montgomery, J.A., David, F., Garneau, M., & Brunengraber, H. 1993. Rates of Gluconeogenesis and Citric Acid Cycle in Perfused Livers, Assessed from the Mass Spectrometric Assay of the ^{13}C Labeling Pattern of Glutamate. *The Journal of Biological Chemistry*, **268**(6), 4170–4180.

Inbar, L., & Lapidot, A. 1987. ^{13}C-NMR, ^{1}H-NMR and gas-chromatography mass-spectrometry studies of the biosynthesis of ^{13}C-enriched L-lysine by *Brevibacterium flavum*. *Eur. J. Biochem.*, **162**, 621–633.

Jeffrey, F.M.H., Rajagopal, A., Malloy, C.R., & Sherry, A.D. 1991. ^{13}C-NMR: a simple yet comprehensive method for analysis of intermediary metabolism. *TIBS*, **16**, 5–10.

Kownatzki, D. 1998. *Numerische Verfahren für die Versuchsplanung bei metabolischen ^{13}C-Markierungsexperimenten*. Diploma dissertation, University of Siegen.

Künnecke, B., Cerdan, S., & Seelig, J. 1993. Cerebral Metabolism of (1,2-$^{13}C_2$)Glucose and (U-$^{13}C_4$)3-Hydroxybutyrate in Rat Brain as Detected by ^{13}C NMR Spectroscopy. *NMR in Biomedicine*, **6**, 264–277.

Lee, W.-N.P. 1993. Analysis of tricarboxylic acid cycle using mass isotopomer ratios. *J. Biol. Chem.*, **268**, 25522–25526.

Lee, W.-N.P., Byerley, L.O., Bergner, E.A., & Edmond, J. 1991. Mass isotopomer analysis: Theoretical and practical considerations. *Biological Mass Spectrometry*, **20**, 451–458.

Lee, W.-N.P., Bergner, E.A., & Guo, Z.K. 1992. Mass isotopomer pattern and precursor-product relationship. *Biological Mass Spectrometry*, **21**, 114–122.

Marx, A., de Graaf, A.A., Wiechert, W., Eggeling, L., & Sahm., H. 1996. Determination of the Fluxes in Central Metabolism of *Corynebacterium glutamicum* by NMR spectroscopy combined with Metabolite Balancing. *Biotechnol.Bioeng.*, **49**, 111–129.

Munack, A. 1989. Optimal Feeding Strategy for Identification of Monod-Type Models by Fed-Batch Experiments. *In:* Fish, N.M., & Fox, R.I. (eds), *Modelling and Control of Biotechnical Processes, Proceedings of the 4th international Congress on Computer Applications in Fermentation Technology, Cambridge, September 25-29, 1988*. Pergamon Press.

Pázman, A. 1986. *Foundations of Optimum Experimental Design*. Mathematics and its Applications. Kluwer Academic Publishing.

Petersen, S., de Graaf, A.A., Möllney, M., Kownatzki, D., Wiechert, W., & Sahm, H. 1998. Flux analysis of anaplerotic pathways in *Corynebacterium glutamicum* using ^{13}C labeling techniques. *In: Metabolic Engineering II, Elmau, Germany, October 25-30, 1998*. United Engineering Foundation, New York.

Sauer, U., Hatzimanitakis, V., Bailey, J.E., Hochuli, M., Szyperski, T., & Wüthrich, K. 1997. Metabolic fluxes in riboflavin-producing *bacillus subtilis*. *Nature Biotechnology*, **15**, 448–452.

Schmidt, K., Carlsen, M., Nielsen, J., & Villadsen, J. 1997. Modelling Isotopomer Distribution in Biochemical Networks Using Isotopomer Mapping Matrices. *Biotechnology and Bioengineering*, **55**(6), 831–840.

Schmidt, K., Norregard, L.C., Pedersen, B., Meissner, A., Duus, J.O., Nielsen, J., & Villadsen, J. 1998a. Determination of intracellular metabolic fluxes from fractional enrichment measurements and isotopomer analysis of ^{13}C labeled components. *Metabolic Engineering Journal*. In press.

Schmidt, K., Nielsen, J., & Villadsen, J. 1998b. Quantitative analysis of metabolic fluxes in *E. coli*, using 2 dimensional NMR spectroscopy and complete isotopomer models. *Journal of Biotechnology*. In press.

Sonntag, K., Eggeling, L., de Graaf, A.A., & Sahm, H. 1993. Flux partitioning in the split pathway of lysine synthesis in *Corynebacterium glutamicum* — quantification by ^{13}C- and ^{1}H-NMR spectroscopy. *Eur.J.Biochem.*, **213**, 1325–1331.

Szyperski, T. 1995. Biosynthetically directed fractional ^{13}C-labeling of proteinogenic amino acids — An efficient analytical tool to investigate intermediary metabolism. *Eur.J.Biochem.*, **232**, 433–448.

Takors, R., Wiechert, W., & Weuster-Botz, D. 1997. Experimental Design for the Identification of Macrokinetic Models and Model Discrimination. *Biotechnology and Bioengineering*, **56**, 564–567.

Wiechert, W. 1995. Algebraic Methods for the Analysis of Redundancy and Identifiability in Metabolic ^{13}C Labelling Systems. *Pages 169–184 of:* Schomburg, D., & Lessel, U. (eds), *Bioinformatics; From Nucleic Acids and Proteins to Cell Metabolism*. Verlag Chemie.

Wiechert, W. 1996. *Metabolische Kohlenstoff-Markierungssysteme – Modellierung, Simulation, Analyse, Datenauswertung*. *Habilitation* Thesis, University of Bonn, Jülicher Forschungsbericht 3301, ISSN 0944-2952.

Wiechert, W., & de Graaf, A.A. 1996. In Vivo Stationary Flux Analysis by ^{13}C Labelling Experiments. *Adv.Biochem.Eng.Biotechnol.*, **54**, 109–154.

Wiechert, W., & de Graaf, A.A. 1997. Bidirectional Reaction Steps in Metabolic Networks. Part I: Modelling and Simulation of Carbon Isotope Labelling Experiments. *Biotechnol.Bioeng.*, **55**, 101–117.

Wiechert, W., Siefke, C., de Graaf, A.A., & Marx, A. 1997a. Bidirectional Reaction Steps in Metabolic Networks. Part II: Flux Estimation and Statistical Analysis. *Biotechnol.Bioeng.*, **55**, 118–135.

Wiechert, W., Möllney, M., & Wurzel, M. 1997b. Modelling, Analysis and Simulation of Metabolic Isotopomer Labelling Systems. *In: 15th IMACS World Congress, Berlin, August 1997*.

Wiechert, W., Möllney, M., Isermann, N., Wurzel, M., & de Graaf, A.A. 1999. Bidirectional Reaction Steps in Metabolic Networks. Part III: Explicit Solution and Analysis of Isotopomer Labeling Systems. *Biotechnol.Bioeng.* This volume.

Wurzel, M. 1997. *Stabilität und eindeutige Lösbarkeit von Isotopomeren-Bilanzgleichungssystemen*. Diploma dissertation, University of Bonn.

C Figures and Tables

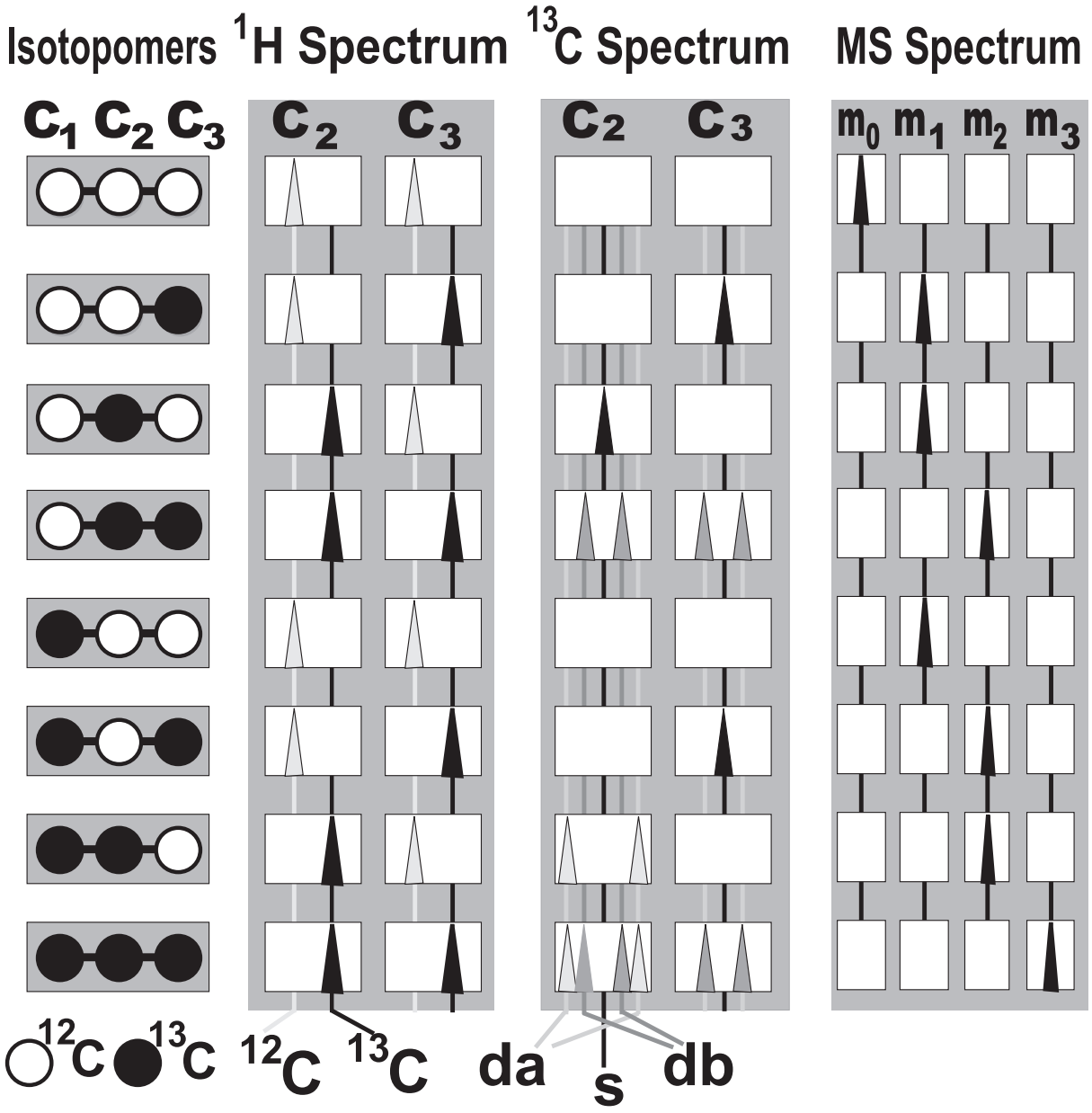

Figure 1: Schematic ^{1}H NMR, ^{13}C NMR, and MS spectra of different alanine isotopomers. The first carbon atom is not considered for NMR because no proton is bonded to it (no ^{1}H NMR peak) and its ^{13}C-NMR-signal is generally very weak. For ^{13}C NMR the scalar coupling constant $J_{C_2C_1}$ is larger than $J_{C_2C_3}$ so that a singlet s, an A-doublet da, a B-doublet db and a double doublet dd=da+db arise. Isotope effects of non-carbon atoms have been neglected in the MS spectra.

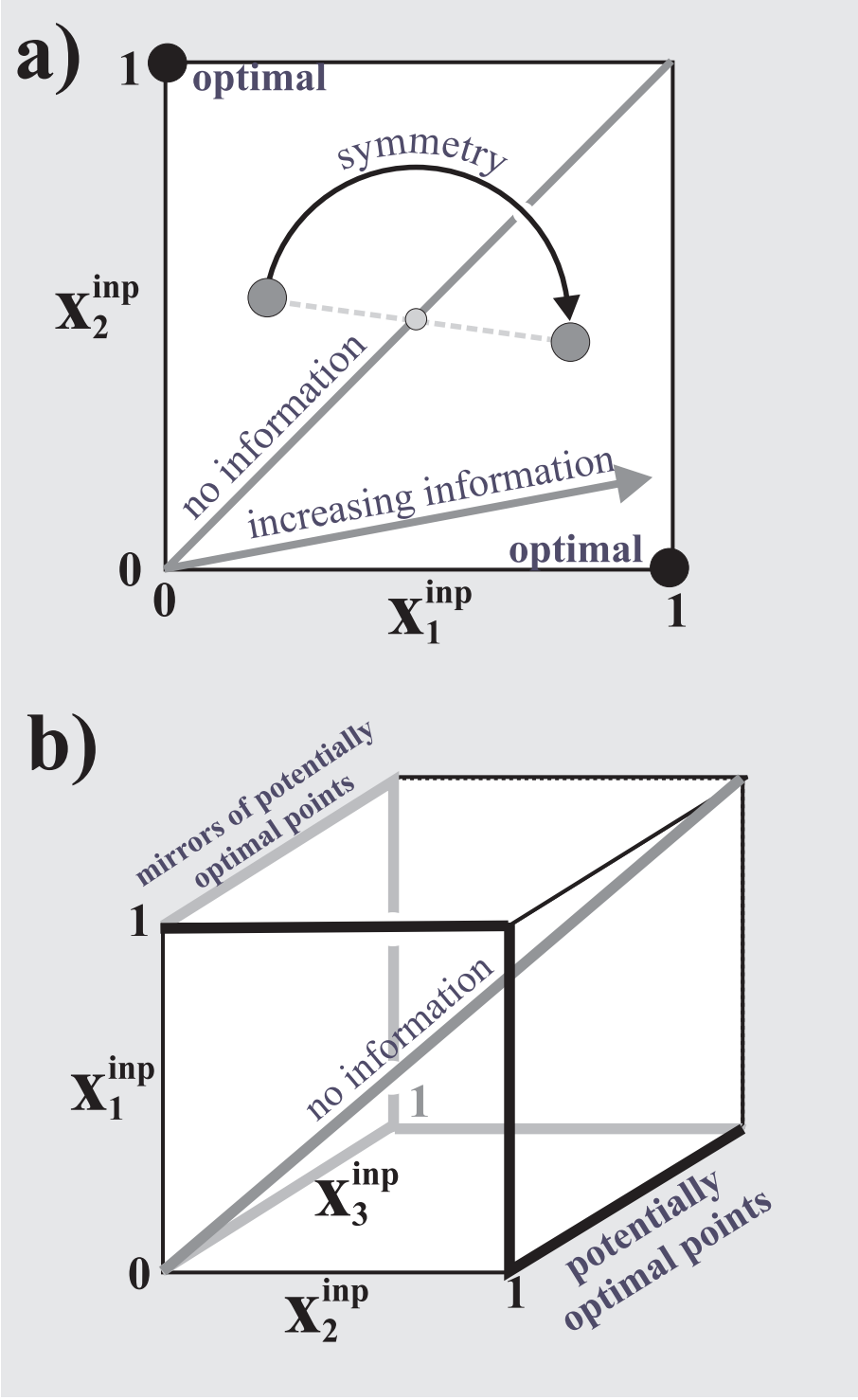

Figure 2: Illustration of the optimal design theorem for PE measurements in a) two and b) three dimensions. Each point in the design square or cube represents the positional enrichments in the input substrate with two or three carbon atoms. There is a D-value which corresponds to each point and which measures how informative the corresponding experiment is.

Figure 3: Relative information yield (compared to the PE reference experiment) for all edges of the design cube in the example system with PE measurements. Additionally the number of labeled carbon atoms of each input isotopomer and the improvement of the pentose phosphate pathway influx estimate is displayed. Mirror images and uniformly labeled substrates are ommited.

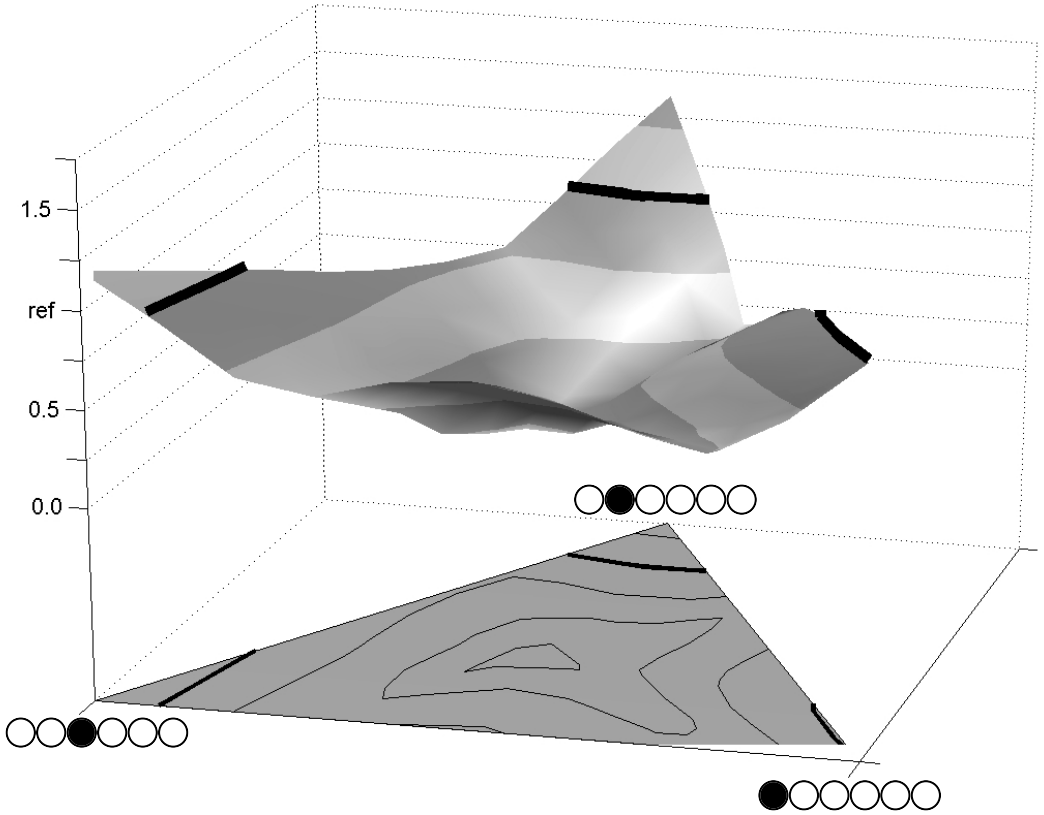

Figure 4: Relative information yield for all mixtures of 1-, 2- and 3-^{13}C labeled glucose when only PE measurements are given. Each corner of the mixture triangle corresponds to a mixture with 100% of a component. The reference value 1 is indicated by thick lines. The labels indicate those experiments which are shown in more detail in Figure 5.

Figure 5: Comparison of all estimated single parameter standard deviations for the PE data set with glucose labeled at positions a) 1,3,6, b) 1,2,4 , c) 2, d) 3, e) 1 as input substrate. The meaning of the flux names is given in the text.

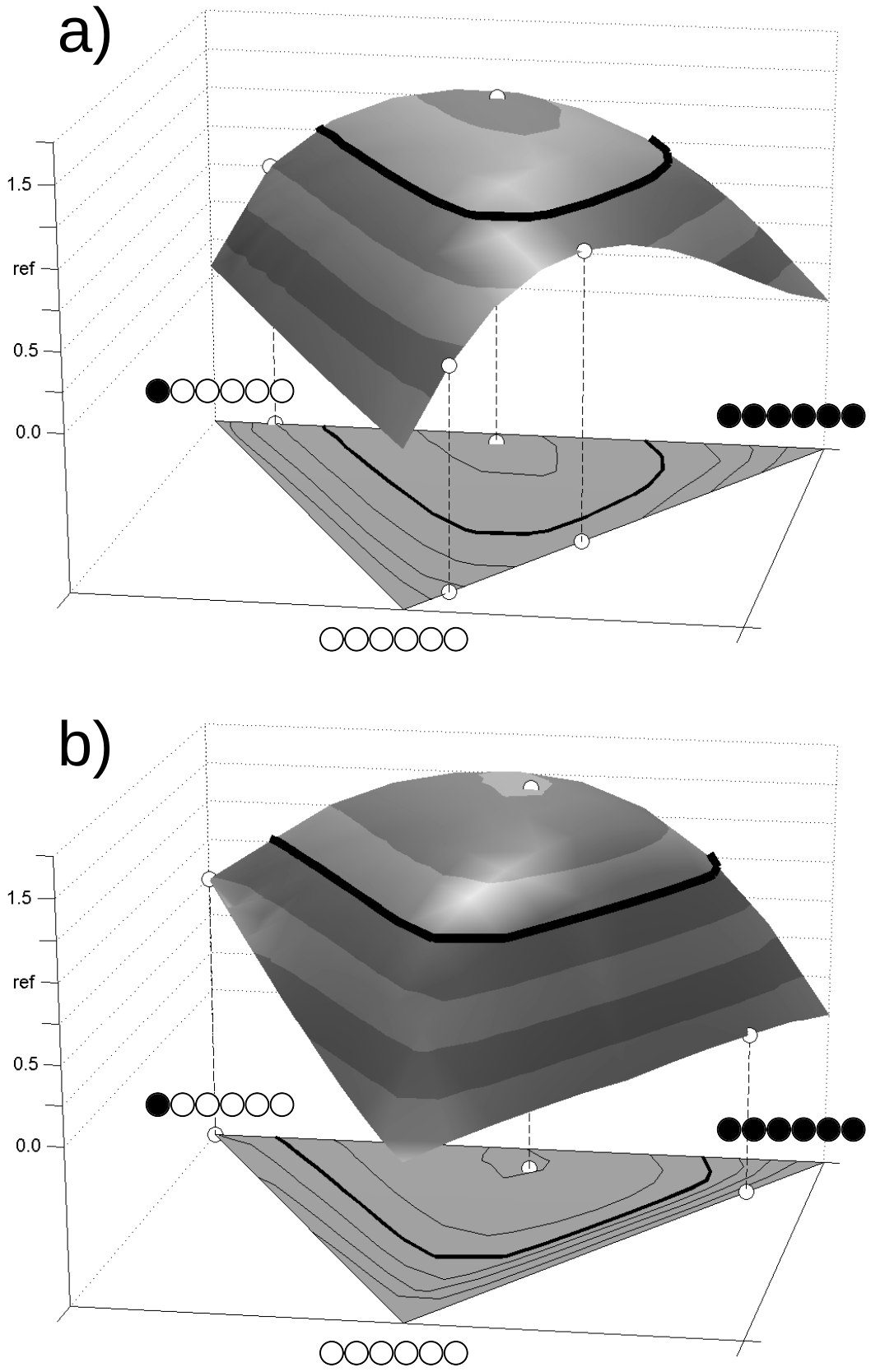

Figure 6: Relative information yield for all mixtures of 1-labeled, unlabeled and uniformly labeled glucose as input substrate with a) MP measurements, b) MS measurements. The labels indicate those experiments which are shown in more detail in Figure 7 and 8.

Figure 7: Comparison of all estimated single parameter standard deviations for the MP data set with a) 0:90:10, b) 0:60:40, c) 90:0:10 and d) 50:5:45 mixtures of 1-labeled, unlabeled and uniformly labeled glucose as input substrate.

Figure 8: Comparison of all estimated single parameter standard deviations for the MI data set with a) 0:20:80, b) 100:0:0 and c) 40:12:48 mixtures of 1-labeled, unlabeled and uniformly labeled glucose as input substrate.

Figure 9: Comparison of all estimated single parameter standard deviations for the best feasible PE, MP and MI data sets found in this investigation. a) PE data with 2-labeled glucose, b) MP data with a 50:5:45 mixture and c) MI data 40:12:48 mixture of 1-labeled, unlabeled and uniformly labeled glucose.

Measured Metabolite	Metabolic Precursor	Carbon atom	PE data Std.Dev.	MP data Std.Dev. s	da	db	dd
RNA/DNA	Ri5P	1	0.005	0.010	0.010		
RNA/DNA	Ri5P	2	0.010	0.010	0.010	0.010	0.010
RNA/DNA	Ri5P	3	0.010	0.010	0.010	0.010	0.010
RNA/DNA	Ri5P	4	0.010	0.010	0.010	0.010	0.010
RNA/DNA	Ri5P	5	0.010	0.010	0.010		
Phe	E4P	5+9	0.010	0.008	0.008		0.008
Phe	E4P	6+7+8	0.020	0.008	0.008		0.008
Ser	GAP	1	0.004				
Ser	GAP	2	0.004	0.005	0.005	0.005	0.005
Ser	GAP	3	0.004	0.005	0.005		
Ala	Pyr	1	0.010				
Ala	Pyr	2	0.010	0.005	0.005	0.005	0.005
Ala	Pyr	3	0.010	0.005	0.005		
Glu	AKG	1	0.005				
Glu	AKG	2	0.005	0.005	0.005	0.005	0.005
Glu	AKG	3	0.005	0.005	0.005		0.005
Glu	AKG	4	0.005	0.005	0.005	0.005	0.005
Glu	AKG	5	0.010				
Asp	OAA	1	0.005				
Asp	OAA	2	0.010	0.005	0.005	0.005	0.005
Asp	OAA	3	0.010	0.005	0.005	0.005	0.005
Asp	OAA	4	0.010				
Lys	Pyr, OAA	2	0.005	0.005	0.005	0.005	0.005
Lys	Pyr, OAA	3	0.010	0.005	0.005		0.005
Lys	Pyr, OAA	4	0.010	0.005	0.005		0.005
Lys	Pyr, OAA	5	0.010	0.005	0.005		0.005
Lys	Pyr, OAA	6	0.005	0.005	0.005		
CO2		1	0.005				
Number of Measurements			56	68			
Number of Groups			28	21			

Table 1: Assumed measurements and the associated standard deviations for positional enrichment (derived from (Wiechert *et al.*, 1997a)) and multiplet data (derived from (Szyperski, 1995) and (Schmidt *et al.*, 1998b)). The scaling factor for the groups is set to 1, i.e. all values can be directly interpreted on a fractional enrichment scale. The multiplet abbreviations are: s-singlet, da-doublet a, db-doublet b, dd-double doublet. If db is missing dd means a triplet peak.

Measured Metabolite	Metabolic Precursor	Measured Fragment	MS data Std.Dev.
RNA/DNA	Ri5P	1-5	0.004
Phe	E4P, PEP	2-9	0.004
Ser	GAP	1-2	0.004
	GAP	2-3	0.004
Gly	GAP	2	0.004
Ala	Pyr	2-3	0.004
Ala	Pyr	1-3	0.004
Leu	Pyr	2-6	0.004
Val	Pyr	2-5	0.004
	Pyr	1-2	0.004
	Pyr	1-5	0.004
Glu	AKG	1-5	0.004
	AKG	1-2	0.004
Pro	AKG	2-5	0.004
Asp	OAA	1-4	0.004
	OAA	2-4	0.004
	OAA	2	0.004
Ile	OAA	2-6	0.004
Thr	OAA	1-2	0.004
Lys	Pyr, Asp	2-6	0.004
Number of Measurements			90
Number of Groups			20

Table 2: Metabolite fragments assumed to be measured by GC-MS in the example and the associated standard deviations (derived from (Christensen & Nielsen, 1999)). The scaling factor for the groups is set to 1, i.e. all values can be directly interpreted on a fractional enrichment scale.