

Methodology article

Open Access

## An analytic and systematic framework for estimating metabolic flux ratios from $^{13}\text{C}$ tracer experiments

Ari Rantanen<sup>\*1,2</sup>, Juho Rousu<sup>1</sup>, Paula Jouhten<sup>3</sup>, Nicola Zamboni<sup>2</sup>, Hannu Maaheimo<sup>3</sup> and Esko Ukkonen<sup>1</sup>

Address: <sup>1</sup>Department of Computer Science, University of Helsinki, Finland, <sup>2</sup>Institute of Molecular Systems Biology, ETH Zurich, Switzerland and <sup>3</sup>VTT Technical Research Centre of Finland, Finland

Email: Ari Rantanen\* - [ari.rantanen@iki.fi](mailto:ari.rantanen@iki.fi); Juho Rousu - [juho.rousu@cs.helsinki.fi](mailto:juho.rousu@cs.helsinki.fi); Paula Jouhten - [pjouhten@mappi.helsinki.fi](mailto:pjouhten@mappi.helsinki.fi); Nicola Zamboni - [zamboni@imsb.biol.ethz.ch](mailto:zamboni@imsb.biol.ethz.ch); Hannu Maaheimo - [hannu.maaheimo@helsinki.fi](mailto:hannu.maaheimo@helsinki.fi); Esko Ukkonen - [ukkonen@cs.helsinki.fi](mailto:ukkonen@cs.helsinki.fi)

\* Corresponding author

Published: 6 June 2008

Received: 14 January 2008

*BMC Bioinformatics* 2008, **9**:266 doi:10.1186/1471-2105-9-266

Accepted: 6 June 2008

This article is available from: <http://www.biomedcentral.com/1471-2105/9/266>

© 2008 Rantanen et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Metabolic fluxes provide invaluable insight on the integrated response of a cell to environmental stimuli or genetic modifications. Current computational methods for estimating the metabolic fluxes from  $^{13}\text{C}$  isotopomer measurement data rely either on manual derivation of analytic equations constraining the fluxes or on the numerical solution of a highly nonlinear system of isotopomer balance equations. In the first approach, analytic equations have to be tediously derived for each organism, substrate or labelling pattern, while in the second approach, the global nature of an optimum solution is difficult to prove and comprehensive measurements of external fluxes to augment the  $^{13}\text{C}$  isotopomer data are typically needed.

**Results:** We present a novel analytic framework for estimating metabolic flux ratios in the cell from  $^{13}\text{C}$  isotopomer measurement data. In the presented framework, equation systems constraining the fluxes are derived automatically from the model of the metabolism of an organism. The framework is designed to be applicable with all metabolic network topologies,  $^{13}\text{C}$  isotopomer measurement techniques, substrates and substrate labelling patterns.

By analyzing nuclear magnetic resonance (NMR) and mass spectrometry (MS) measurement data obtained from the experiments on glucose with the model micro-organisms *Bacillus subtilis* and *Saccharomyces cerevisiae* we show that our framework is able to automatically produce the flux ratios discovered so far by the domain experts with tedious manual analysis. Furthermore, we show by *in silico* calculability analysis that our framework can rapidly produce flux ratio equations – as well as predict when the flux ratios are unobtainable by linear means – also for substrates not related to glucose.

**Conclusion:** The core of  $^{13}\text{C}$  metabolic flux analysis framework introduced in this article constitutes of flow and independence analysis of metabolic fragments and techniques for manipulating isotopomer measurements with vector space techniques. These methods facilitate efficient, analytic computation of the ratios between the fluxes of pathways that converge to a common junction metabolite. The framework can be seen as a generalization and formalization of existing tradition for computing metabolic flux ratios where equations constraining flux ratios are manually derived, usually without explicitly showing the formal proofs of the validity of the equations.

## Background

From microorganisms to animals and plants, cells adjust their metabolic operations to fulfill the demand of energy and biosynthetic precursors. In nature this is a challenging task because substrate availability is typically limited and often changing in its composition. To ensure viability on a broad palette of chemically heterogeneous substrates, cells have developed intertwined enzymatic networks that also confer robustness against genetic mutations. Optimum redistribution of molecular fluxes in metabolism is achieved by multilevel regulation circuits. In recent years, experimental measurement of *in vivo* metabolic fluxes has attracted much attention. For example, in biotechnology metabolic fluxes are utilized to lead rational strain engineering, whereas systems biologists assess fluxes to unravel targets and mechanisms of metabolic regulation.

Metabolic fluxes are often estimated using flux balance analysis (FBA) [1,2]. In FBA, fluxes are solved by fixing some objective for the metabolism of an organism, such as maximal growth. Then, a corresponding optimization problem is solved by using the stoichiometry of the metabolic network as a constraint to the optimization. FBA is a viable method for studying the metabolic capabilities of an organism, but as a method for estimating metabolic fluxes it has some weaknesses. First, selecting the correct objective for the metabolism is far from trivial [3], especially when mutant strains or behaviour in exceptional environmental conditions is analyzed. Second, there can be many biologically interesting flux distributions that give an optimal solution to the optimization problem of FBA.

A more direct method for experimental determination of the metabolic fluxes is to feed an organism with  $^{13}\text{C}$  labelled substrate, observe the fate of  $^{13}\text{C}$  atoms in the cell at isotopomeric steady state with mass spectrometry (MS) or nuclear magnetic resonance (NMR) measurements, and then infer the metabolic fluxes from the measurements. The rationale behind these  $^{13}\text{C}$  tracer experiment is that, often alternative pathways between metabolites in the network manipulate the carbon backbones of the metabolites differently, thus inducing different  $^{13}\text{C}$  labelling patterns to metabolites. Then, constraints to fluxes complementary to the basic stoichiometric constraints can be derived by measuring the relative abundances of different labelling patterns in the metabolites.

The main difficulty in applying the procedure in practice is that current measurement techniques only can produce incomplete information about relative abundances of different  $^{13}\text{C}$  labelling patterns, the *isotopomer distributions*, of some metabolites, usually protein bound amino acids in the network, and no isotopomer information at all for many intermediate metabolites of interest [4-6]. This

imposes a highly non-linear dependency between the measured isotopomer distributions of the metabolites and the metabolic fluxes, which is very challenging to solve both computationally and statistically.

Currently, two main approaches for  $^{13}\text{C}$  metabolic flux analysis exist. In the *global isotopomer balancing* approach, the problem of estimating metabolic fluxes from the isotopomer measurements is formulated as a nonlinear optimization problem, where candidate flux distributions are iteratively generated until they fit well enough to the experimental  $^{13}\text{C}$  labelling patterns [7-11]. Global isotopomer balancing is a versatile approach that can be applied with all network topologies, substrate labelling distributions and with all measurement techniques – also in short time scales where isotopomeric steady state is not reached [12-14]. However, due to the nonlinearity of the problem, it is hard to make sure that the optimization has converged to a global optimum and that this optimum is unique [15]. Also, to apply the global isotopomer balancing approach successfully, one usually needs comprehensive information on the uptake and production rates of external metabolites, as well as about biomass composition of the cell. This information can be hard to obtain, especially in large-scale experiments where dozens to hundreds of mutants or different organisms are comparatively analyzed [16,17].

A metabolic flux ratio approach (METAFor) [4,18,19] for  $^{13}\text{C}$  metabolic flux analysis has traditionally relied more on the expertise of a domain specialist than advanced computational techniques. In metabolic flux ratio analysis, the aim is to write linear equations constraining the *ratios* of fluxes producing the same metabolite. The equations are manually derived by domain experts, by careful (and tedious) inspection of metabolic networks. The motivation for the approach is that, in many cases, the knowledge about the flux ratios already offers enough information about the response of an organism to its environment.

The ratio of competing fluxes or pathways producing the same metabolite is easy to understand, and in many cases easier to estimate reliably than all the fluxes in the network – some interesting flux ratios might be obtainable from scarce measurement data or from the incomplete model of metabolic network that would not allow a reliable estimation of a complete flux distribution using global isotopomer balancing. Flux ratios can also be obtained without knowing the uptake and production rates of external metabolites of the biomass composition. And, if enough non-redundant flux ratios are identified, it is possible to use this information to construct and solve an equation system for the full flux distribution of the metabolic network [20-22].

As a downside, manually derived flux ratio equations depend heavily on the topology of a metabolic network, measurement capabilities and substrate labelling distributions. Thus, each time a new organism or new mixture of substrates are introduced, flux ratio equations have to be verified and new ones possibly derived. To date, flux ratio equations are manually derived for central carbon metabolisms of three model organisms on glucose, *S. cerevisiae* [17,23], *B. subtilis* [24] and for *Escherichia coli* [18,19]. Recently, flux ratio equations of *S. cerevisiae* were modified for *Pichia pastoris* grown on glycerol and on glycerol/methanol mixtures [25,26]. Facilitating the process of deriving flux ratio equations for other organisms and other substrates clearly calls for automatic tools. Also, many times the (simplifying) assumptions made by the expert in the derivation and solution of flux ratio equations, are not reported in detail. Thus, it is often nontrivial to verify the correctness of given flux ratio equations.

In this article we present a novel automatic framework for deriving flux ratios when the measurement data and the model of metabolic network are given as input. The framework is based on the graph algorithmic flow analysis of metabolite fragments in the metabolic network [27] and on the interpretation and manipulation of both NMR and MS data with vector space techniques [21]. The goal of our work is to combine the good aspects of global isotopomer balancing and manual flux ratio analysis: like global isotopomer balancing, our framework is systematic and can be applied with all network topologies, substrates and substrate labelling distributions and with all current isotopomer measurement techniques. Thus, laborious and error-prone manual inspection of metabolic network models and the tailoring of the equation systems constraining the fluxes separately for each experimental setting required in manual flux ratio techniques can be avoided. On the other hand, during the automated construction of flux ratios we resort to linear optimization techniques only, combined with graph algorithms of polynomial worst case time complexity. Thus, our framework is computationally efficient and avoids problems with local and multiple optima frequently met in global isotopomer balancing. The trade-off of this philosophy is, however, the requirement of measuring isotopomer distributions of metabolites more rigorously to obtain full flux distribution. Given insufficient measurements, our framework can solve the flux ratios only for some, but not necessary for all metabolites in the network. We expect that, especially as measuring isotopomers of intermediate metabolites becomes more routine, our framework will be an attractive method for  $^{13}\text{C}$  flux analysis.

## Results and Discussion

In this section we demonstrate the applicability of the presented framework by empirical results. We show that our

automatic and systematic framework is able to reproduce flux ratios previously determined by a manual analysis from NMR and GC-MS isotopomer measurements of protein bound amino acids of *S. cerevisiae* and *B. subtilis* on glucose. Thus, we can conclude that the presented framework is powerful enough to provide interesting flux ratio information in the well studied experimental settings. Furthermore, we show that the framework can be applied to study less known experimental conditions without any further effort by discovering the flux ratios that can be estimated when *B. subtilis* is grown on malate instead of glucose. The results of this analysis show that our framework can detect profound effects the change of external substrate can have to the flux ratio computations. The results indicate that our framework is a good tool to study flux ratios of microbes in different experimental conditions – a claim that will try to validate with more experiments in our further work.

We obtained NMR and GC-MS labelling data, where isotopomer distributions of protein bound amino acids of *S. cerevisiae* and *B. subtilis* grown on different conditions were measured. Then, available flux ratios were computed with the presented framework. Models of metabolic networks applied in the analysis can be found from the supplementary material of this article: additional files 1 and 2 contain the SBML model file and a visualization of the model of *S. cerevisiae*, while additional files 3 and 4 contain the same information for *B. subtilis*. In the models, some simplifications common to  $^{13}\text{C}$  metabolic flux analysis were applied by pooling metabolites whose isotopomer pools can be assumed to be fully mixed (cf. [28]). Pooling of metabolites was carried for (i) the three pentose-phosphates in PPP, (ii) phosphotrioses between GA3P and PEP in glycolysis, and (iii) oxaloacetate and malate in the TCA cycle. In these cases, pooling is justified by the existence of fast equilibrating, bidirectional reactions between the listed intermediates and the empirical evidence that their isotopic labeling is not distinguishable with the current analytical tools. Cofactor metabolites were excluded from the model as cofactor specificities and activities are not accurately known for many reactions.

The bulk of the carbon mappings of reactions in the metabolic network were provided by ARM project [29]. Carbon mappings from amino acids to their precursors were conform to [4] and [23]. Before the analysis of the real measurement data, the correctness of the implementation of the framework was empirically verified by estimating flux ratios for junction metabolites in the metabolic network of *S. cerevisiae* from the artificial data generated by the 13C-FLUX software [8].

### NMR measurements from *S. cerevisiae* on glucose

In the first experiment we analyzed NMR isotopomer measurement data from protein bound amino acids of *S. cerevisiae* that was grown on uniformly labelled glucose (see Section *Experimental NMR and GC-MS methods* for more details on experimental settings).

From the 15 measured amino acids we were able to estimate flux ratios for seven junction metabolites: oxaloacetate, PEP, glycine and serine on cytosol and for oxaloacetate, acetyl-CoA and pyruvate in mitochondria. Furthermore, an upper bound for a ratio of GA3P molecules that have visited the transketolase reaction was obtained by manually simplifying the model to imitate the previously reported ways to manually compute the corresponding upper bound (cf. [4]). (The structural analysis of the metabolic network model described in Section *Structural analysis of isotopomer systems* can help in discovering such simplifications, but they also need some expert insight. As the simplifications are currently not done automatically, the systematical framework is unable to validate them.)

The computed flux ratios were compared with the manually derived ratios [23], when the assumptions made in the manual derivation of flux ratios were consistent with the model used. In all cases, automatically computed flux ratios agreed well with the manually derived ratios (Table 1). Differences between the estimations can be explained by numerical instabilities and by differences in computational procedures: in manually derived ratios the estimations are based on the breakage of a single bond in different routes leading to a metabolite while in our framework more isotopomer information is possible utilized in the estimation.

### GC-MS measurements from *B. subtilis* on glucose

In the second experiment we analyzed GC-MS isotopomer measurement data from protein bound amino acids of *Bacillus subtilis* that was grown on uniformly labelled glucose (see Section *Experimental NMR and GC-MS methods* for more details on experimental settings).

In comparison to eukaryotic *S. cerevisiae*, the metabolic network of prokaryotic *B. subtilis* lacks cellular compartments. Thus, from the point of view of  $^{13}\text{C}$  metabolic flux analysis, there are fewer interesting junction metabolites in the central carbon metabolism of *B. subtilis* where the flux ratios can be estimated. From the GC-MS measurements of 14 amino acids we were able to compute flux ratios for four junction metabolites – oxaloacetate, pyruvate, PEP and glycine – when [U- $^{13}\text{C}$ ]-glucose was used as a carbon source. Furthermore, an upper bound for a ratio of GA3P molecules that have visited transketolase reaction was obtained by manually simplifying the model of the metabolic network. Excluding pyruvate, we were able to compute the same ratios with [1- $^{13}\text{C}$ ]-glucose as a carbon source.

We compared the computed flux ratios to ones obtained with the software FiatFlux [30] that is based on the manually derived analytic equations for computing flux ratios. Currently, manually derived flux ratio equations for [1- $^{13}\text{C}$ ]-glucose as a carbon source exist only for PEP and for the upper bound to the flux through oxidative pentose phosphate pathway. In general, the flux ratios computed with different methods from the same data and with the same assumptions about the topology of metabolic network were in good agreement (Table 2). (As a data cleaning procedure, we removed from [1- $^{13}\text{C}$ ]-glucose data the mass distributions of fragments whose fractional enrichment deviated more than 5% from the assumed fractional enrichment of 20% in [U- $^{13}\text{C}$ ]-data. This was done because differences in fractional enrichments can be tracked in uniformly labelled data where the fractional enrichment of each fragment is known *a priori*, but not in positionally labelled data, where the fractional enrichment of a fragment depends on the network topology and the fluxes. This kind of irregularities are in general caused by noise in fragments with low intensity or by coeluting analytes with overlapping fragment masses.) Again, differences between the flux ratios estimated by different methods can be explained by numerical instabilities and by differences in isotopomer information applied during the estimation. Variation in the estimated flux ratios between repeated experiments (six repetitions for [1- $^{13}\text{C}$ ]-glucose

**Table 1: Estimated flux ratios from NMR measurements of *S. cerevisiae*.**

flux ratios	our framework	METAFor
PYR(mit) from MAL(mit) : PYR(mit) from PYR(cyt)	0.05 : 0.95	0.03 : 0.97
OAA(mit) from TCA-cycle : OAA(mit) from OAA(cyt)	0.50 : 0.50	0.50 : 0.50
PEP from OAA(cyt) : PEP from GA3P	0 : 1	0.04 : 0.96
OAA(cyt) from PYR(cyt) : OAA(cyt) from OAA(mit)	0.40 : 0.60	0.43 : 0.57
GLY from SER : GLY from CI + CO <sub>2</sub>	0.96 : 0.04	0.96 : 0.04

Directly comparable flux ratios computed from the NMR data by the framework presented in this paper and by manual flux ratio analysis (METAFor).

**Table 2: Estimated flux ratios from GC-MS measurements of *B. subtilis*.**

flux ratios	our fw (UL)	our fw (ICL)	FiatFlux (UL)	FiatFlux (ICL)
PYR from MAL : PYR from PEP	0.01 : 0.99		0.04 : 0.96	
OAA from TCA-cycle : OAA from PYR	0.42 : 0.58	0.37 : 0.63	0.41 : 0.59	
PEP from OAA : PEP from GA3P	0 : 1	0 : 1	0.04 : 0.96	0 : 1
GLY from SER : GLY from CI + CO2	1 : 0	1 : 0	1 : 0	
SER from GLY : SER from GA3P	0.09 : 0.91		0.14 : 0.86	

Directly comparable flux ratios computed from the GC-MS data described by the framework presented in this paper (*our fw*) and by FiatFlux software [30]. *UL* denotes uniform labelling of external glucose, *ICL* external glucose labelled to the first carbon position.

experiment, four repetitions for uniformly labelled glucose experiment) was negligible.

#### ***In silico* calculability analysis of *B. subtilis* on malate**

One of the strengths of the presented framework is that it is able to automatically produce metabolic flux ratios also when other external labelled substrates than commonly used glucose are fed to organisms. To demonstrate this ability, we applied our framework to predict what flux information would be available, if we feed *B. subtilis* with malate.

We applied the *in silico* calculability analysis (see Section *Calculability analysis*) to examine which flux ratios are calculable in the best case from GC-MS measurements of amino acids, when *B. subtilis* is grown on [U-13C]-labelled malate. Interestingly, our fragment flow analysis revealed that – with the reaction reversibilities in the applied model – the isotopomer distributions of GA3P, PEP and pyruvate depend only on the isotopomer distribution of the fragment containing the first three carbons of oxaloacetate, but not on the relative fluxes producing these metabolites. Thus, isotopomer balances for GA3P, PEP and pyruvate reduce to mass balances and the ratios of fluxes producing these metabolites cannot be estimated. This somewhat surprising phenomenon is due the fact that the rearrangements of carbon chains occurring in PPP pathway will affect only to the carbon fragments that will be recycled in the upper metabolism but not the carbon fragments that can flow back GA3P, PEP and pyruvate from PPP (we modelled a reaction from GA3P to F6P as unidirectional one).

Preliminary experiments with GC-MS data from *B. subtilis* grown on [U-13C]-labelled malate agreed with the results of fragment flow analysis: constraints to the isotopomer distributions of fluxes entering to these metabolites were identical within the limits of measurement accuracy. On the other hand, our framework was able to estimate for example the TCA-cycle activity also when *B. subtilis* is grown on malate, just as predicted by the calculability analysis.

#### **Conclusion**

In this article we introduce a systematic and analytic framework for <sup>13</sup>C metabolic flux ratio analysis. At the heart of the framework lie the techniques for flow analysis of a metabolic network and for manipulating isotopomer measurements as linear subspaces. These techniques facilitate the efficient and analytic computation of the ratios between the fluxes producing the same junction metabolite. The framework can be seen as a generalization and formalization of existing analytic methods for computing metabolic flux ratios [23,30,31] where equations constraining flux ratios are manually derived. Like the recent methods to improve the speed of the simulation of isotopomer distributions in the global isotopomer balancing framework [10,32], our framework relies on graph algorithms. However, both our goals and applied techniques are quite different from these approaches. In [10] and [32] connected components of isotopomer graphs are discovered to divide the simulation of isotopomer distributions to smaller subtasks. In our framework, flow analysis techniques are applied to discovered metabolite fragments with equivalent isotopomer distributions in every isotopomeric steady state.

Our experiments with NMR and MS data show that the framework is able to provide relevant information about metabolic fluxes, even when only constraints to the isotopomer distributions of protein-bound amino acids are measured.

Thanks to recent advancements in measurement technology improving the feasibility of mass isotopomer measurements of intermediate metabolites [13,33], we expect that the full power of the framework will be harnessed in near future. Measurements from intermediates will make it possible to use larger models of metabolic networks and estimate flux ratios more accurately, without simplifying assumptions about the topology of the metabolic network or directionality of the fluxes. However, these measurements will not be easy to conduct, because of the low abundances of intermediates in the cell. Thus, systematic methods for experimental planning and data quality control are required. The presented framework provides pow-

erful tools for these tasks. First, the framework facilitates time saving *in silico* calculability analysis.

Second, the manipulation of isotopomer measurements as linear subspaces offers a natural way for comparing measurements from different metabolites that contain overlapping information to detect inconsistencies in the measurements: it is enough to compare propagated isotopomer information in the fragments that belong to the same equivalence class. Third, as MS isotopomer measurement techniques have to be developed separately for different intermediate metabolites or metabolite classes, it will be very useful to select a small subset of intermediates that gives enough information about the interesting metabolic fluxes with least experimental effort. In future research, we want to tackle this problem by generalizing our earlier experimental planning method [34], to all measurement data and to realistic measurement error models.

As the presented framework for  $^{13}\text{C}$  metabolic flux analysis only resorts to linear optimization techniques, it is not always able to provide as much information about the metabolic fluxes as the global isotopomer balancing frameworks utilizing more powerful, nonlinear optimization techniques [8,35], that do not necessarily converge to the global optimum. On the other hand, some flux ratios might be computable from the scarce data or incomplete model of metabolic network that does not allow global isotopomer balancing. The differences in the practical performance of different approaches require further research. We see these alternative approaches as complementary ones. A very nice goal would be an integration of our work with global isotopomer balancing: our analytic flux ratios could speed up and direct the optimization process of global isotopomer balancing, that would then fill in the flux ratios possibly missed by our framework.

## Methods

In this section we describe a systematic and analytic computational framework for  $^{13}\text{C}$  metabolic flux analysis. At the end of the section we also shortly describe the experimental method that were applied to produce the isotopomer measurement data that was analyzed in Section *Results and Discussion*.

The overall goal of our computational framework is to automatically infer from the available isotopomer measurement data produced by MS, MS-MS or NMR techniques, an equation system constraining the fluxes. The crucial question is to derive as many non-redundant equations as possible, ideally constraining the flux distribution to a point solution, or in general, as low-dimensional convex set as possible.

In short, the framework consists of the following steps:

1. The model of the metabolic network of an organism is constructed by selecting a set of biochemical reactions operating in the organism and by designating them to correct cellular compartments;

2. Structural analysis of the isotopomer system is conducted, consisting of the following steps:

- (a) Flow analysis of the metabolic network is conducted in order to discover *equivalent fragments*, fragments of carbon backbones of metabolites that will have the same theoretical isotopomer distribution, regardless of the fluxes.

- (b) Independence analysis of fragments is conducted to find statistically independent carbon subsets from metabolites, that is, subsets that have been at some point separated along every pathway able to producing them and that have flux invariant isotopomer distributions. This guarantees that the isotopomer distribution of their union assumes the form of a product distribution.

- (c) *In silico* calculability analysis is performed to test if the available measurement techniques and substrate labelings make it *in principle* possible to obtain the required flux information.

3. Wet-lab isotopomer tracer experiments are conducted and constraints to isotopomer distributions are measured;

4. The fluxes of the network are estimated. The process consists of the following steps:

- (a) Isotopomer measurement data is *propagated* in the metabolic network model from the measured metabolites to unmeasured ones according to the equivalences discovered in step 2.

- (b) An equation system tying the isotopomer data and the fluxes together is constructed and solved, either to obtain a flux distribution for the metabolic network as a whole, or for a single junction metabolite to obtain the ratios of fluxes producing it.

- (c) The statistical analysis of obtained fluxes or flux ratios is carried out.

In the following we first formalize the  $^{13}\text{C}$  flux analysis problem and then detail the computational steps above.

## Preliminaries

In  $^{13}\text{C}$  metabolic flux analysis the carbon atoms of metabolites are of special interest. We denote with  $M$  the set of carbon locations  $M = \{c_1, \dots, c_k\}$  of a  $k$ -carbon metabolite.

By  $|M| = k$  we denote the number of carbons in  $M$ . *Fragments* of metabolites are subsets  $F = \{f_1, \dots, f_h\} \subseteq M$  of the carbons of the metabolite. A fragment  $F$  of  $M$  is denoted as  $M|F$ . A *metabolic network*  $G = (C, \mathcal{R})$  is composed of a set  $C = \{M_1, \dots, M_m\}$  of *metabolites* and a set  $\mathcal{R} = \{\rho_1, \dots, \rho_n\}$  of *reactions* that perform the interconversions of metabolites.

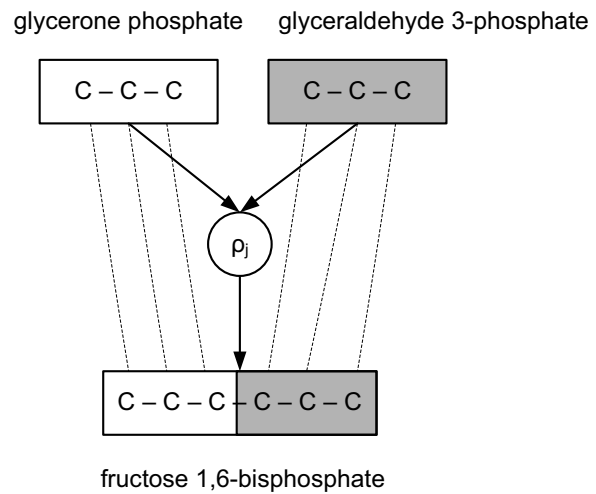
With *isotopomers* we mean molecules with similar element structure but different combinations of  $^{13}\text{C}$  labels. Isotopomers of the molecule  $M = \{c_1, \dots, c_k\}$  are represented by binary sequences  $b = \{b_1, \dots, b_k\} \in \{0, 1\}^k$  where  $b_i = 0$  denotes a  $^{12}\text{C}$  and  $b_i = 1$  denotes a  $^{13}\text{C}$  in location  $c_i$ . Molecules that belong to the  $b$ -*isotopomer* of  $M$  are denoted by  $M(b)$ . Isotopomers of metabolite fragments  $M|F$  are defined in an analogous manner: a molecule belongs to the  $F(b)$ -*isotopomer* of  $M$ , denoted  $M|F(b_1, \dots, b_h)$ , if it has a  $^{13}\text{C}$  atom in all locations  $f_j$  that have  $b_j = 1$ , and  $^{12}\text{C}$  in other locations of  $F$ . Isotopomers with equal numbers of labels belong to the same *mass isotopomer*. We denote *mass isotopomers* of  $M$  by  $M(+p)$ , where  $p \in \{0, \dots, |M|\}$  denotes the number of labels in isotopomers belonging to  $M(+p)$ .

The *isotopomer distribution*  $D_M$  of metabolite  $M$  gives the relative abundances  $0 \leq P_M(b) \leq 1$  of each isotopomer  $M(b)$  in the pool of  $M$  such that

$$\sum_{b \in \{0,1\}^{|M|}} P_M(b) = 1.$$

The isotopomer distribution  $D_{M|F}$  of fragment  $M|F$  and the *mass isotopomer* distribution  $D_M^m$  of mass isotopomers  $M(+p)$  are defined analogously:  $D_{M|F}$  of metabolite  $M$  gives the relative abundances  $0 \leq P_{M|F}(b) \leq 1$  of each isotopomer  $M|F(b)$  and  $(\rho_1^p, \rho_2^p, \dots, \rho_n^p)$  gives the relative abundances  $0 \leq P_M(+p) \leq 1$  of each mass isotopomer  $M(+p)$ .

Reactions are pairs  $\rho_j = (\alpha_j, \lambda_j)$  where  $\alpha_j = (\alpha_{1j}, \dots, \alpha_{mj}) \in m$  is a vector of *stoichiometric coefficients*-denoting how many molecules of each kind are consumed and produced in a single reaction event-and  $\lambda_j$  is a carbon mapping describing the transition of carbon atoms in  $\rho_j$  (see Figure 1). Metabolites  $M_i$  with  $\alpha_{ij} < 0$  are called *substrates* and with  $\alpha_{ij} > 0$  are called *products* of  $\rho_j$ . If a metabolite is a product of at least two reactions, it is called a *junction*. If  $\alpha_{ij} < 0$ , a reaction event of  $\rho_j$  consumes  $|\alpha_{ij}|$  molecules of  $M_i$ , and if  $\alpha_{ij} > 0$ , it produces  $|\alpha_{ij}|$  molecules of  $M_i$ . Bidirectional reactions are modelled as a pair of reactions.



**Figure 1**  
**An example of a metabolic reaction.** In the reaction  $\rho_j$ , a fructose 1,6-bisphosphate ( $\text{C}_6\text{H}_{14}\text{O}_{12}\text{P}_2$ ) molecule is produced from glycerone phosphate ( $\text{C}_3\text{H}_7\text{O}_6\text{P}$ ) and glyceraldehyde 3-phosphate ( $\text{C}_3\text{H}_7\text{O}_6\text{P}$ ) molecules. Carbon maps are shown with dashed lines. Glyceraldehyde 3-phosphate is equivalent to the gray fragment of fructose 1,6-bisphosphate while glycerone phosphate is equivalent to the white fragment.

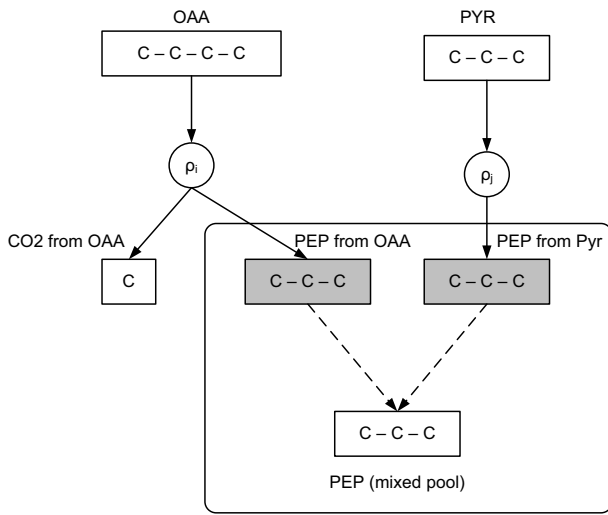
A *pathway*  $p$  in network  $G$  from metabolite fragments  $\{F_1, \dots, F_k\}$  to fragment  $F'$  is a sequence  $(\rho_1^p, \rho_2^p, \dots, \rho_n^p)$  of reactions such that a *composite carbon mapping*  $(\lambda_n^p \circ \dots \lambda_2^p \circ \lambda_1^p)(\cup_{i \in [1,k]} F_i) = F'$ , defined by  $p$  maps the carbons of  $\{F_1, \dots, F_p\}$  to the carbons of  $F'$ .

For the rest of the article, it is important to distinguish between the *subpools* of a metabolite pool produced by different reactions. Therefore, we denote by  $M_{ij}$  the subpool of the pool of  $M_i$  produced ( $\alpha_{ij} > 0$ ) or consumed ( $\alpha_{ij} < 0$ ) by reaction  $\rho_j$ . The concept of the subpools of a metabolite pool is illustrated in Figure 2.

By  $M_{i0}$  we denote the subpool of  $M_i$  that is related to the external in flow or external out flow of  $M_i$ . We call the sources of external in flows *external substrates*. Subpools of fragments are defined analogously. In  $^{13}\text{C}$  metabolic flux analysis, the quantities of interest are the rates or the fluxes  $v_j \geq 0$  of the reactions  $\rho_j$ , giving the number of reaction events of  $\rho_j$  per time unit. We denote by  $\mathbf{v}$  the vector  $[v_1, \dots, v_n]$  of fluxes, or a *flux distribution*.

**Generalized isotopomer balances**

The framework for  $^{13}\text{C}$  metabolic flux analysis presented in this article rests on the assumption that the metabolic network is in metabolic and isotopomeric steady state. In



**Figure 2**  
**An example of subpools of a metabolite pool.** Phosphoenolpyruvate (PEP) is produced by two different reactions ( $\rho_i$  and  $\rho_j$ ), either from Oxaloacetate (OAA) or from glyceraldehyde 3-phosphate (GA3P). Thus, PEP has two in flow subpools, *PEP from OAA* and *PEP from GA3P* (grey boxes) that are mixed in the common PEP pool (at the bottom).

the metabolic steady state, *metabolite balance*, or *mass balance*

$$\sum_{j=1}^n \alpha_{ij} v_j = \beta_i \tag{1}$$

holds for each metabolite  $M_i$ . Here,  $\beta_i$  is the measured external in flow ( $\beta_i < 0$ ) or external out flow ( $\beta_i > 0$ ) of metabolite  $M_i$ . From balance equations (1) defined for every metabolite  $M_i$  we will obtain a metabolite balancing, or stoichiometric equation system.

In isotopomer steady state, for each isotopomer  $b$  of each metabolite  $M_i$  in the metabolic network the following *isotopomer balance* holds:

$$\sum_{j=1}^n \alpha_{ij} v_j P_{M_{ij}}(b) = \beta_i P_{M_{i0}}(b). \tag{2}$$

For metabolic flux analysis (1) and (2) bear a fundamental difference: the former cannot be used to estimate fluxes of alternative pathways producing  $M_i$  while the latter can. However, using (2) is not in general admissible in practice: typically abundances  $P_M(b)$  of isotopomers are not fully determined from the measurements, and we need to settle for some constraints to the distribution  $D_M$ . A crucial building block of our framework is the representation

of isotopomer measurements as systems of linear equations (c.f. [21])

$$\sum_b s_{bh} P_M(b) = d_h, h = 1, \dots, r \tag{3}$$

where  $s_{bh}$  is the weight of isotopomer  $b$  in the  $h$ 'th constraint,  $d_h$  is a value derived from isotopomer measurements, and  $r$  is the total number of constraints. We call (3) *isotopomer constraints*. For a  $k$  carbon metabolite,  $2^k$  linearly independent isotopomer constraints – one for each isotopomer – are necessary and sufficient to constrain the isotopomer distribution  $D_M$  to a point solution. A set of isotopomer constraints has a natural matrix representation  $SD_M = \mathbf{d}$ , where  $S = (s_{bh})_{b,h}$  is a  $2^k$  by  $r$  matrix, where  $1 \leq r \leq 2^k$  is the number of isotopomer constraints (the trivial constraint  $\sum_b P_M(b) = 1$  by definition always holds).

The use of (3) follows from the simple observation that isotopomer balance (2) implies that each linear combination of isotopomers is balanced. Thus, we can write a new balance equation that constrains the fluxes producing  $M_i$  as soon as we know the value of the *same* linear combination of isotopomer abundances for each subpool of  $M_i$ . We have

$$\sum_{j=1}^n \alpha_{ij} v_j d_{ij} = \beta_i d_{i0}, \tag{4}$$

where each  $d_{ij} = \sum_b s_b P_{M_{ij}}(b)$  is a linear combination of the form (3), with coefficients  $s_b$  that do not depend on  $j$ , i.e. they are the same for each subpool  $M_{ij}$ . We call (4) a *generalized isotopomer balance*.

**Representing MS and NMR isotopomer measurements as linear constraints**

In the following, we will show by examples how MS and NMR data can be represented as isotopomer constraints. Let us first consider mass isotopomer distributions obtained from MS. Here we omit discussion on practicalities such as corrections for natural abundances of  $^{13}\text{C}$  isotopes (c.f., [6,36]) and concentrate on the conceptual level. For example, the +2 mass isotopomer of a three carbon metabolite  $M$  satisfies

$$P_M(+2) = P_M(011) + P_M(110) + P_M(101)$$

which conforms to (3) by taking  $s_{011,2} = s_{110,2} = s_{101,2} = 1$ , and  $s_{b,2} = 0$  otherwise. Similarly, the coefficients  $s_{bk}$  can be derived for all mass isotopomers  $+k, k = 0, \dots, 3$ .

Isotopomer data originating from Tandem MS, or MS-MS, falls into the same representation. Consider, for example,



a fragment  $M|F$  of a three-carbon metabolite, containing the first and second carbon of  $M$ . The following equation holds for the mass isotopomer  $M|F(+1)$ :

$$P_{M|F(+1)} = P_{M|F(01)} + P_{M|F(10)}.$$

The equation can be written in terms of the precursor  $M$ , but the exact form of the equation depends on the mode of tandem MS. If the full scan mode is used, we obtain

$$P_{M|F(+1)} = P_M(010) + P_M(011) + P_M(100) + P_M(101),$$

as all precursor molecules  $M$  that have exactly one carbon among the first and second location contribute to the fragment mass isotopomer  $M|F(+1)$ . On the other hand, in the daughter-ion scanning mode a single mass isotopomer, for example  $M(+2)$ , is selected as the precursor. Then we obtain

$$P_{M|F(+1)} = P_M(011) + P_M(101),$$

as the precursor must always have two  $^{13}\text{C}$  atoms in total. We refer the reader to [6,36] for further details. Also NMR  $^{13}\text{C}$  isotopomer measurements, where relative intensities of different combinations of  $^{13}\text{C}$  and  $^{12}\text{C}$  atoms that are coupled to an observed  $^{13}\text{C}$  atom are measured, can be expressed as linear combinations of isotopomer abundances. For example, for a three-carbon metabolite  $M$ , the following constraints to  $D_{M_i}$  can be inferred for the labeling pattern 010:

$$\frac{P_M(010)}{\sum_{b_1, b_3 \in \{0,1\}} P_M(b_1 1 b_3)} = d(010)$$

where  $d(010)$  is the measured intensity. Rewriting the above as

$$d(010) \cdot \sum_{b \in \{110, 011, 111\}} P_M(b) = (1 - d(010)) \cdot P_M(010)$$

and denoting  $s_b = d(010)$ , for  $b \in \{110, 011, 111\}$ ,  $s_{010} = d(010) - 1$  and  $s_b = 0$  for  $b \in \{000, 100, 001, 101\}$ , the above can be seen to conform to (3). Similar derivation can be used for other isotopomer signals present in the NMR spectrum to obtain the corresponding isotopomer constraints.

*Projection of isotopomer measurements to fragments*

In our computational framework, it will be necessary to project the measurement data obtained for a metabolite  $M$  to its fragments  $M|F$  and vice versa. In this projection, we want to avoid any unnecessary loss of measurement information, that is, we want to obtain as many linearly

independent constraints to the isotopomer distribution of  $F$  as possible. For example, if we have measured that a two-carbon metabolite  $M$  has the isotopomer distribution

$$P_M(00) = 0.2, P_M(01) = 0.3, P_M(10) = 0.4, P_M(11) = 0.1$$

and we need to know the isotopomer distribution of the fragment  $M|F$  consisting of the first carbon of  $M$ , we should obtain

$$P_F(0) = P_M(0*) = P_M(00) + P_M(01) = 0.5, \\ P_F(1) = P_M(1*) = P_M(10) + P_M(11) = 0.5.$$

For the general model of isotopomer measurements (4) the projection of measurement information from a metabolite to its fragments can be done by the techniques of linear algebra introduced in [21]. We recapitulate the techniques in the following.

Recall the general form of isotopomer measurement  $SD_M = \mathbf{d}$ , where  $S$  denotes a matrix with  $2^k$  columns, one column for each isotopomer  $b$  of  $k$ -carbon metabolite  $M$ , and each row  $h$  of  $S$  corresponds to a measured isotopomer constraint (3). The rows of  $S$  span a subspace  $S \subseteq \mathcal{I}_M$  in a  $2^k$  dimensional vector space  $\mathcal{I}_M$  spanned by all possible isotopomer distributions  $D_M$ .

Also the metabolite fragments are naturally represented as vector subspaces. Let  $U_F$  denote a matrix with also a column for each isotopomer  $M(b)$  and a row for each isotopomer  $F(b')$  of  $M|F$ , that is,

$$U_F(b', b) = \begin{cases} 1 & \text{if } b_j = b'_j \text{ for all carbon positions } j \in F_k \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The rows of  $U_F$  span another subspace  $\mathcal{U}_F \subseteq \mathcal{I}_M$ . Any isotopomer distribution  $D_{M|F}$  lies in this subspace, and hence also any isotopomer constraint  $S_F D_{M|F} = \mathbf{d}$  for fragment  $M|F$  necessarily lies in the same subspace.

In conclusion, the available information about  $D_M$  is given as its linear projection onto  $S$ , and anything we can express about  $D_{M|F}$  in terms of isotopomer constraints is contained within  $\mathcal{U}_F$ . Thus, any isotopomer constraint for  $D_{M|F}$  that we can derive from the measurements can be expressed in terms of the *vector space intersection*  $S \cap \mathcal{U}_F$ .

Thus, to obtain isotopomer constraints for fragment  $M|F$  from a measurement  $SD_M = \mathbf{d}$ , we need to compute the

vector space intersection  $\mathcal{Y}_{M|F} = \mathcal{S} \cap \mathcal{U}_F$  and project the measurement to  $\mathcal{Y}_{i,F}$ . This can be done by standard linear algebra (c.f. [21]). This process gives us as output isotopomer constraints of the required form

$$Y_F D_{M|F} = \mathbf{d}_F.$$

Finally, transforming a fragment constraint  $Y_F D_{M|F} = \mathbf{d}_F$  into an isotopomer constraint  $S D_M = \mathbf{d}$  is easy: we post-multiply the fragment constraint with the matrix  $U_{M|F} : S = Y_F U_{M|F}$  and  $\mathbf{d} = \mathbf{d}_F U_{M|F}$ .

### Structural analysis of isotopomer systems

The incomplete nature of  $^{13}\text{C}$  measurement data requires us to find ways to use the available data the best way possible. The central concept is to find invariants of isotopomer distributions that remain through the pathways, and allow us to trade or *propagate* measurement information from one metabolite to another. This allows us to write or augment generalized isotopomer balances for metabolites for which the isotopomer distributions are not completely determined by measurements. Thus, the fluxes are potentially more tightly pinpointed as well.

In particular, we use two techniques: First, *flow analysis* is used to uncover sets of metabolite fragments that have the same isotopomer distribution regardless of the fluxes. Second, *independence analysis* of fragments is used to uncover situations where two fragments of the same metabolite induce the product distribution for the isotopomer distribution of their union.

#### Flow analysis of metabolic networks

The goal of the flow analysis [27] is to partition the fragments of the metabolites in the network to *equivalence classes* such that fragments in the same equivalence class have identical isotopomer distributions in every steady state. This can be guaranteed if a fragment is produced from another a such a manner that the carbons within the fragment never depart from each other regardless of the pathway that is being used.

Formally, we say that fragment  $F'$  *dominates* fragment  $F$  if the following conditions are met

1.  $F$  and  $F'$  have the equal number of carbons;
2. all carbons of  $F$  originate always from the carbons of  $F'$ ;
3. carbon of  $F'$  stay connected to each other via all pathways from  $F'$  to  $F$ ;
4. composite carbon mappings are the same in all pathways from  $F'$  to  $F$ .

Intuitively, a *dominated fragment* ( $F$ ) is always produced from its *dominator* ( $F'$ ) without manipulating the carbon chain of the fragment. Thus, isotopomer distribution of the dominated fragment does not contain any information about the metabolic fluxes. For a fragment  $F$  that has no dominators, the transitive closure of the domination relation corresponds to the class of equivalent fragments in the network.

The simplest example of fragment equivalence is the one between a substrate  $M_k$  and product  $M_i$  in a single reaction  $\rho$ . If the atoms in  $M_i|F$  originate from  $M_k|F'$ , then the fragments  $M_k|F'$  in the subpool  $M_{ij}$  produced by reaction  $\rho_j$  are equivalent with the fragment  $M_i|F$  (Figure 1). Furthermore, if metabolite  $M_i$  has only one producing reaction  $\rho_j$ , isotopomer distributions of subpool  $M_{ij}$  and  $M_i$  coincide. Thus, if fragment  $M_i|F$  is produced from a single fragment  $M_k|F'$  of some substrate  $M_k$  of  $\rho_j$ ,  $F$  and  $F'$  are equivalent. By transitivity, all fragments in the linear pathway are equivalent.

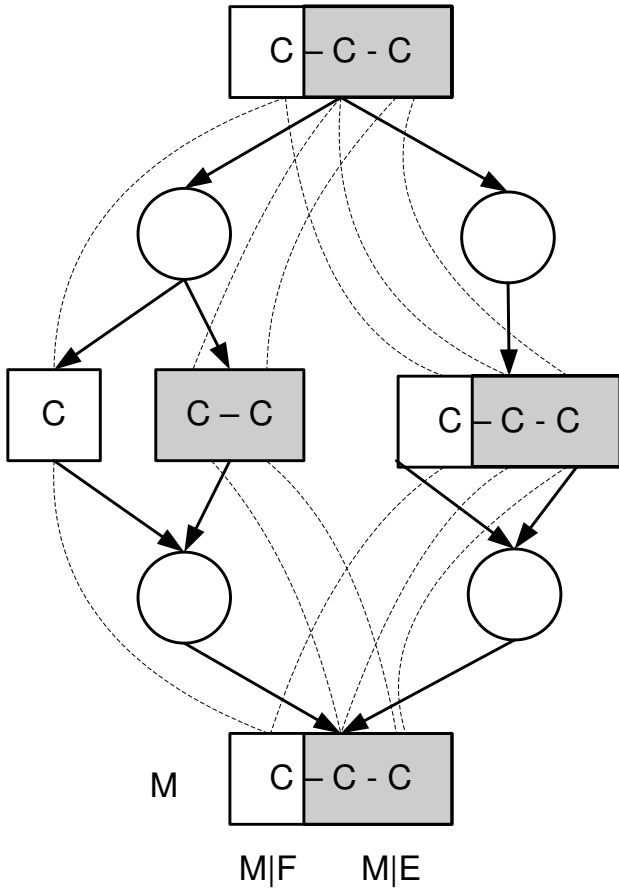
More complicated case of fragment equivalence is found when a fragment of a junction metabolite is dominated by an upstream fragment (Figure 3). In [27] we show that the the classes of equivalent fragments corresponding the conditions (1–4) can be efficiently computed. Very briefly, first the metabolic network is transformed to a *fragment flow graph* that connects substrate metabolite fragments to their product fragments for each reaction in the network. Then, the dominator tree [37,38] of the fragments in the fragment flow graph is constructed. It turns out that the subtrees of this dominator tree correspond to the required fragment equivalence classes (see Figure 4).

Fragment equivalence classes have many uses [27]. Most importantly, measured isotopomer constraints to fragment  $F$  can be directly propagated to another fragment  $F'$  in the same equivalence class, by applying the joint carbon mappings between  $F$  and  $F'$ . This helps in the construction of generalized balance equations (4) where the same isotopomer information is required for each subpool of junction metabolites.

#### Independence analysis of fragments

A complementary property to fragment equivalence  $^{13}\text{C}$  is the statistical independence of fragment isotopomer distributions. Intuitively, if two fragments of the same metabolite are statistically independent, new isotopomer constraints to the union of them can be obtained by taking a product of the isotopomer distributions of the independent fragments.

More formally, the basic question is on what conditions the distribution  $D_{M|E \cup F}$  of union of two fragments will necessarily have the form of a *product distribution*



**Figure 3**  
**An example of fragment equivalence classes in a branched pathway.** An example of equivalence classes of fragments in the metabolic network that contains dominated junction fragments  $M|E$  and  $M|F$ . Grey and white fragments constitute two equivalence classes. Dashed lines depict carbon mappings.

$$D_{M|E \cup F} = D_{M|E} \otimes D_{M|F} \quad (6)$$

where  $\otimes$  denotes the tensor product consisting all terms of the form  $P_{E \cup F}(b) = P_E(b') \cdot P_F(b'')$ , where  $b'$  (resp.  $b''$ ) ranges over all isotopomers of  $E$  (resp.  $F$ ), and  $b$  is the isotopomer of  $M|E \cup F$  formed by joining  $E$  and  $F$ .

The utility of fragment independence is in that it gives us constraints to the isotopomer distributions that are complementary to the isotopomer constraints (3) obtained from the measurements.

In general, two criteria need to be satisfied for statistical independence of two fragments  $M|E$  and  $M|F$ . First, the fragments need to be *structurally independent*, meaning that along all pathways producing the metabolite, at some point all carbons of the fragments have resided in different metabolite molecules. This property can be defined in

recursive manner. Fragments  $M|E$  and  $M|F$  are structurally independent if for all carbon pairs  $(a, b)$ ,  $a \in E$  and  $b \in F$ , for all reactions  $\rho$  producing  $M$ , it holds that

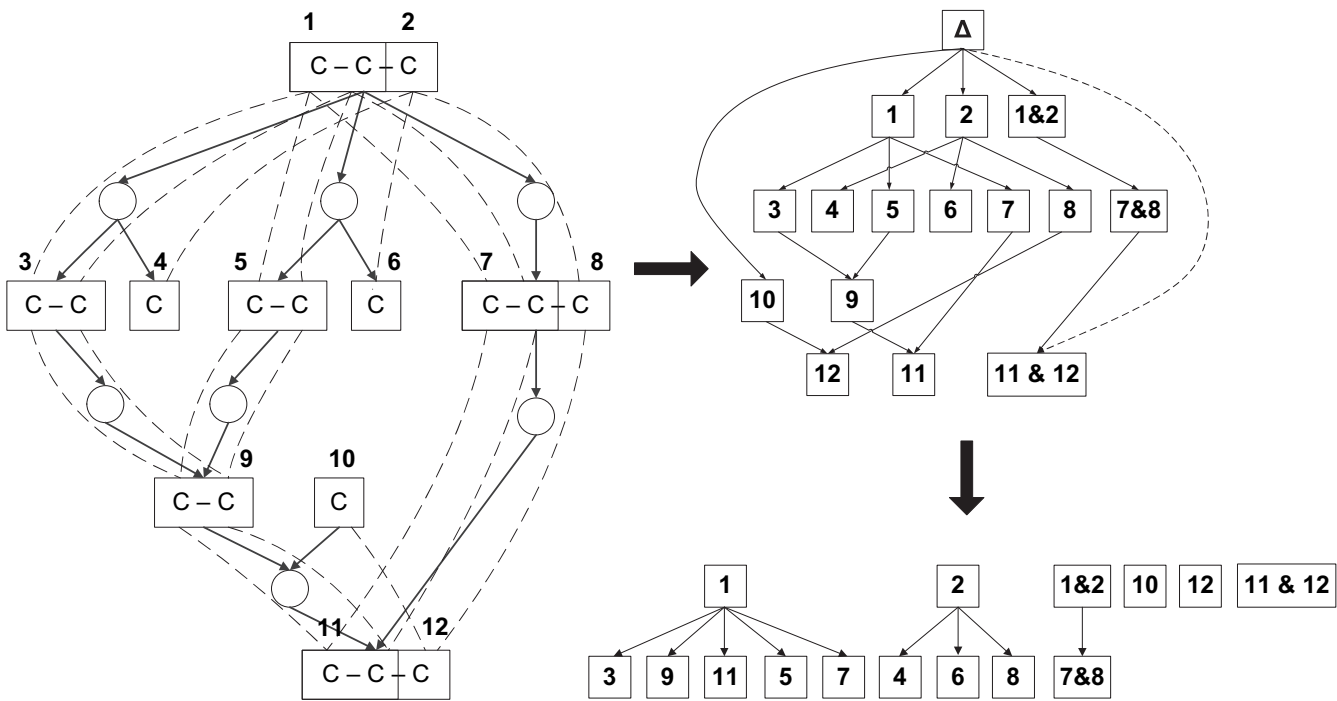
- $a$  and  $b$  originate from different reactants of  $\rho$ , or
- $a$  and  $b$  originate from the same reactant  $M'$ , and the reactant fragments  $M'|F_a$  and  $M'|F_b$ , where  $F_a = \lambda_\rho^{-1}(a)$ ,  $F_b = \lambda_\rho^{-1}(b)$ , are structurally independent.

The second necessary condition is that the two fragments need to be dominated by some other metabolite fragments in the network. This will make the fragment distributions flux invariant. Together, the two criteria guarantee (6) to hold.

A simple case of statistical independence of fragments is a (subpool) product metabolite  $M_i$  of a single reaction  $\rho_j$ , where the fragments  $M_i|E$  and  $M_i|F$  are disjoint and originate from different reactants. The fragments are structurally independent (by originating from different reactants) and are dominated (by reactant fragments of  $\rho_j$ ). Hence (6) holds. The underlying assumption here is that enzymes pick their reactants independently and randomly from the available pools. This case of statistical independence of fragments is depicted Figure 1, where white and grey fragments of D-fructose 1,6-bisphosphate are statistically independent (in the subpool of reaction  $\rho_j$ ).

Another simple example is a junction metabolite  $M_i$  that has two or more producers with associated subpools  $M_{ij}$ . If  $M_{ij}|E$  and  $M_{ij}|F$  are structurally independent in all subpools,  $M_i|E$  and  $M_i|F$  are structurally independent as well. If  $M_i|E$  and  $M_i|F$  are dominated by some fragments in the network, all subpools have the same distribution which takes the form of (6). Without dominance the distribution  $D_{M_i|E \cup F}$  will in general be a *flux-dependent mixture* of product distributions  $D_{M_i|E \cup F} = \sum_j v_j D_{M_{ij}|E} \otimes D_{M_{ij}|F}$ . This case of statistical independence is depicted in Figure 5.

A generalized form of (6), useful for propagation of isotopomer constraints, is derived as follows. Assume independent fragments  $M|E$  and  $M|F$  of metabolite  $M$  and isotopomer constraints  $S_{E \cup F} D_{M|E \cup F} = \mathbf{d}_{E \cup F}$ ,  $S_E D_{M|E} = \mathbf{d}_E$  and  $S_F D_{M|F} = \mathbf{d}_F$ , where  $S = S_E \otimes S_F$ . Now, the  $h$ 'th constraint for fragment  $E$  and  $g$ 'th constraint for fragment  $F$ , written as  $\sum_a s_{ah,E} P_{M|E}(a) = d_{h,E}$  and  $\sum_c s_{cg,F} P_{M|F}(c) = d_{g,F}$ .



**Figure 4**  
**An example of a fragment flow graph and a dominator tree.** A metabolic network (left), the corresponding fragment flow graph (up right) and the subtrees of the dominator tree (down right).

Multiplying the constraints, and denoting  $s_b = s_{ah,E} \cdot s_{cg,F}$  where  $b$  is the isotopomer consistent with fragment isotopomers  $a$  and  $c$ , we get the following equation

$$d_{h,E} \cdot d_{g,F} = \left( \sum_a s_{ah,E} P_{M|E}(a) \right) \cdot \left( \sum_c s_{cg,F} P_{M|F}(c) \right) = \sum_a s_b P_{M|E \cup F}(b) = d_{l,E \cup F}$$

for the  $l$ 'th constraint for  $E \cup F$ . The equations of the above kind can be concisely written in terms of tensors:

$$\mathbf{d} = S_{D_{M|E \cup F}} = (S_E \otimes S_F) D_{M|E} \otimes D_{M|F} = \mathbf{d}_E \otimes \mathbf{d}_F \tag{7}$$

From above, if two of the three vectors  $\mathbf{d}$ ,  $\mathbf{d}_E$ ,  $\mathbf{d}_F$  are known, the remaining unknown one can be solved.

We note in passing that computing constraints to the metabolite given constraints to the fragments is straightforward application of (7).

*Applying fragment independence analysis to flux ratio computation*  
 In our framework, statistical independence of fragments has two uses. We apply it

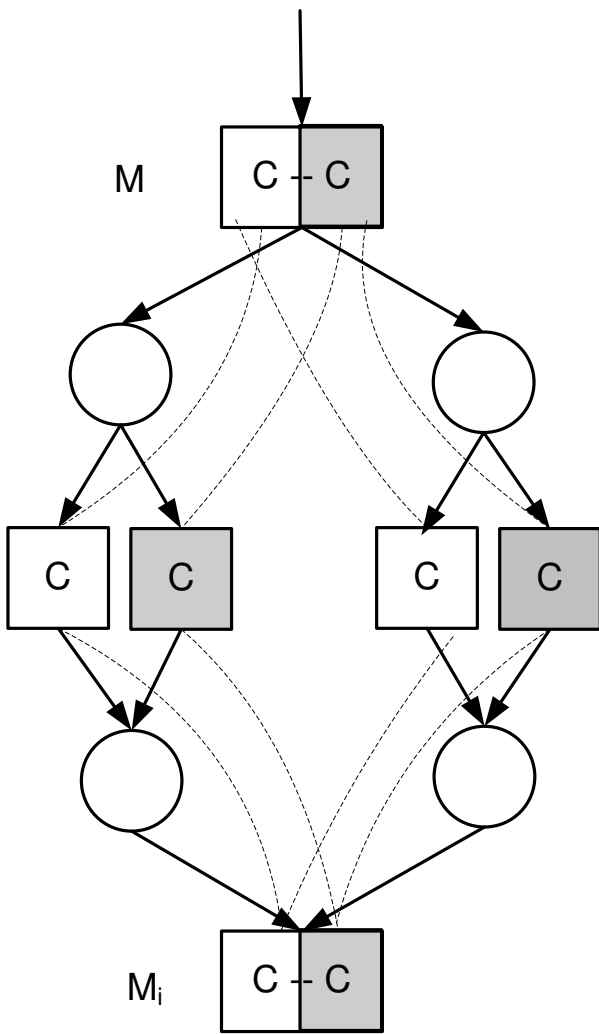
1. to compute isotopomer constraints for the union of independent fragments, given isotopomer constraints to its independent fragments, and

2. to compute isotopomer constraints for an independent fragment given isotopomer constraints to the other fragments and the metabolite as a whole.

In both cases making use of (6) gives us a larger set of constraints than the vector space and flow analysis approach alone.

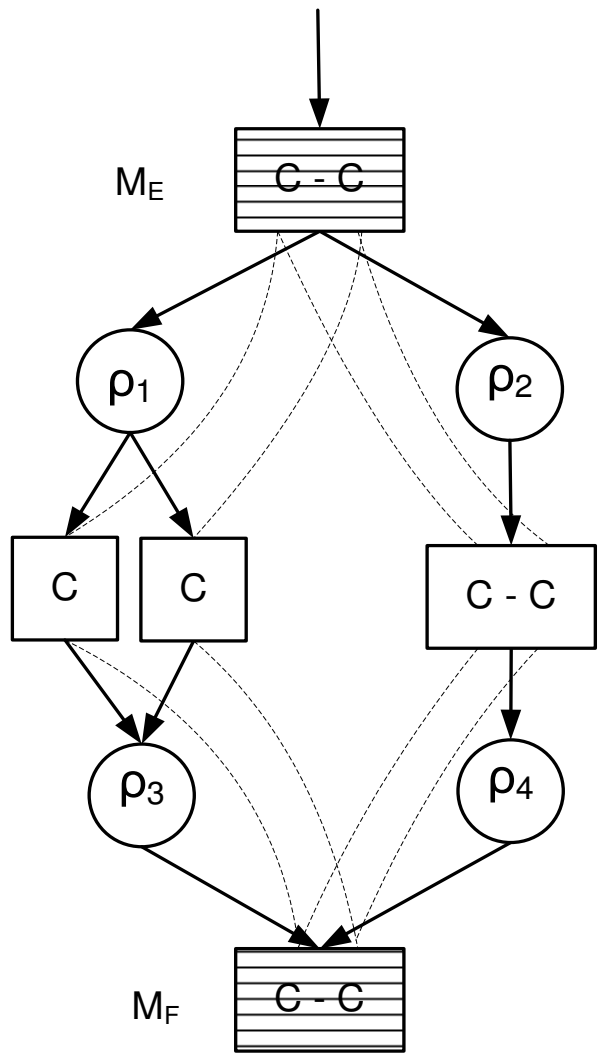
Next we describe how (7) generalizes the basic measurement propagation step of the traditional metabolic flux ratio analysis [31]. In the basic case, the flux ratios are solved for two competing pathways  $p$  and  $q$ , which  $p$  cleaves a certain carbon-carbon bond  $b$  of junction  $M$  while the  $q$  preserves  $b$  intact from the external substrate. (See Figure 6 for an example). This serves also as an example of applying (7) to compute isotopomer constraints for the union of independent fragments.

When a uniformly labelled substrate is used, the labelling degree of every carbon in the network is the same (and known *a priori*) when the system reaches isotopomeric steady state. Thus, the isotopomer distribution  $D_{F_p}$  of a two-carbon fragment  $F$  (metabolite  $M_F$  in Figure 6) containing bond  $b$  can be computed by (7) for pathway  $p$  cleaving  $b$ , while for pathway  $q$ ,  $D_{F_q}$  can be propagated



**Figure 5**  
**An example of statistical independence of fragments.**  
 White and grey one-carbon-fragments of  $M_i$  are statistically independent: both fragments are dominated by one-carbon-fragments of  $M$ , and the fragments are disjoint in every pathway that produce  $M_i$  from  $M$ .

from the external substrate (metabolite  $M_E$  in Figure 6) using the fragment equivalence classes of the previous section. If we are able to measure (constraints to) the isotopomer distribution of the mixed pool  $F$ , we can then automatically derive a generalized isotopomer balance corresponding the manually derived ratio. To use (7) to compute isotopomer constraints for an independent fragment from the known isotopomer constraints to the other independent fragment and to the whole metabolite is complicated by the incompleteness of the measurement data: an arbitrary measurement  $SD_{M|E \cup F} = \mathbf{d}$  might not be directly representable via a tensor product  $S = S_E \otimes S_F$ . Instead, we need to first compute isotopomer subspaces

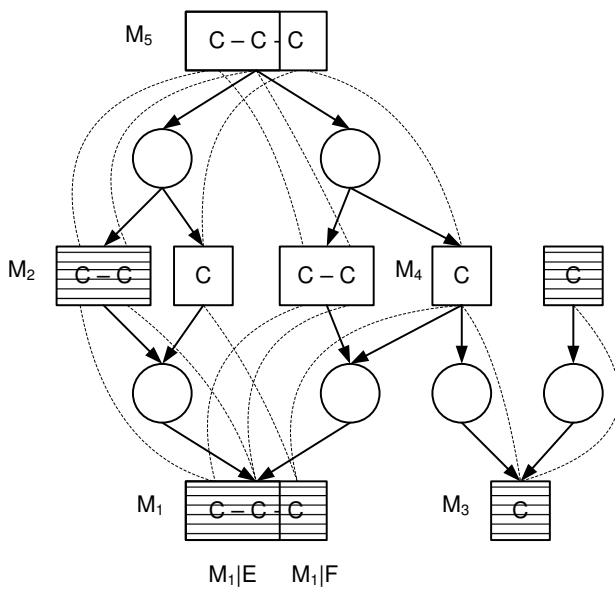


**Figure 6**  
**An example of using fragment independence to obtain new isotopomer constraints under uniform substrate labelling.** Constraints to the isotopomer distributions of striped metabolites are assumed to be known, either by direct measurement of measurement propagation. In pathway  $q = (\rho_2, \rho_4)$ , the isotopomer distribution of  $M_F$  molecules will be the same as in  $M_E$ . In pathway  $p = (\rho_1, \rho_3)$ , the isotopomer distribution of  $M_F$  can be derived by applying fragment independence: the isotopomer distributions of single carbon metabolites produced by  $\rho_1$  are known *a priori* to be equal to the labelling degree of uniformly labelled substrate. As the two carbons of  $M_{F_3}$  are produced from two different metabolites, these carbons are statistically independent to each other in the subpool and the isotopomer distribution  $D(M_{F_p})$  of  $M_F$  molecules produced by  $p$  can be computed by applying Equation 7.

for known isotopomer constraints where (7) can be applied.

The detailed description of this technique is rather technical and omitted from this article. Here we give an example of the technique (See Figure 7). We assume that we know the mass isotopomer distributions  $D_{M_i}^m$  of metabolite  $M_i$  (metabolite  $M_1$  in Figure 7) and  $D_{M_i|E}^m$  of fragment  $M_i|E$ . We furthermore assume that  $M_i|E$  and  $M_i|F$  are independent. From this information the mass isotopomer distribution of  $D_{M_i|E}^m$  can be solved. To be exact,  $D_{M_i|E}^m$  can be solved from the system containing an equation

$$P_{M_i(+p)} = \sum_{\substack{b+l=p \\ M_5|E' \quad M_5|F'}} P_E(+k) \cdot P_F(+l) \quad (8)$$



**Figure 7**  
**An example of using fragment independence to obtain new isotopomer constraints for a reactant.**  
 The mass isotopomer distributions of striped metabolites are assumed to be measured. Fragments  $M_1|E$  and  $M_2$  belong to the same fragment equivalence class. Thus,  $D_m(M_1|E)$  can be derived from  $D_m(M_2)$  by the measurement propagation inside equivalence classes. Furthermore, fragments  $M_5|E'$  and  $M_5|F'$  dominate fragments  $M_1|E$  and  $M_1|F$ , and the bond between  $M_1|E$  and  $M_1|F$  is broken in all pathways producing  $M_1$  from  $M_5$ . Thus,  $M_1|E$  and  $M_1|F$  are statistically independent, and  $D_m(M_1|F)$  can be deduced from  $D_m(M_1)$  and  $D_m(M_1|E)$  by utilizing Equation 7. Computed  $D_m(M_1|F)$  can then be propagated to  $M_4$ , as  $M_1$  and  $M_4$  belong to the same fragment equivalence class. Finally,  $D_m(M_4)$  helps to solve the ratios of fluxes entering to  $M_3$ .

for each mass isotopomer  $p$  of  $M_i$ . To see that (8) conforms to (7), we denote the measured mass isotopomer distribution  $D_{M_i}^m = S \cdot D(M_i)$  of  $M_i$  by  $\mathbf{d}_M$  (i.e. rows of coefficient matrix  $S$  correspond to different mass isotopomers of  $M_i$ ) and the measured mass isotopomer distributions of  $E$  and  $F$  by  $\mathbf{d}_E$  and  $\mathbf{d}_F$ . Let  $|E| = 2$ ,  $|F| = 1$  and  $|M_i| = 3$ , thus  $M_i = E \cup F$ . We have

$$S_E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, S_F = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

with the tensor product

$$S_E \otimes S_F = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \text{ and } S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

As the two matrices are not the same (7) is not directly applicable. However, by summing up the second and the third rows and the fourth and the fifth rows of  $S_E \otimes S_F$  we obtain  $S$ . Intuitively, this means that Equation (7) can be applied to compute  $D_{M_i|F}^m$ , when we take into account that the isotopomer constraints corresponding both second and third rows of  $S_E \otimes S_F$  contribute to mass isotopomer  $M_i(+1)$ , while the isotopomer constraints corresponding fourth and fifth rows of  $S_E \otimes S_F$  contribute to mass isotopomer  $M_i(+2)$ . (From the definition of the tensor product we see that, for example, the second row of  $S_E \otimes S_F$  corresponds to isotopomer constraints  $P_E(00) \cdot P_F(1) = P_E(+0) \cdot P_F(+1)$  and the third row corresponds to the constraints  $P_E(01) \cdot P_F(0) + P_E(10) \cdot P_F(0) = P_E(+1) \cdot P_F(+0)$ , thus validating our intuitive observation.) When the similar information for all rows of  $S_E \otimes S_F$  is collected to a linear equation system, we will obtain the following constraints to the mass isotopomer distribution  $D_{M_i|F}^m$  (which in the case of one-carbon-fragment  $M|F$  coincides with the isotopomer distribution  $D_{M|F}$ ):

$$\begin{pmatrix} P_E(+0) & 0 \\ P_E(+1) & P_E(+0) \\ P_E(+2) & P_E(+1) \\ 0 & P_E(+2) \end{pmatrix} D_m(M|F) = \mathbf{d}_M = \begin{pmatrix} P_{M_i(+0)} \\ P_{M_i(+1)} \\ P_{M_i(+2)} \\ P_{M_i(+3)} \end{pmatrix} = S D_{M_i}$$

which is equal to (8).

#### Calculability analysis

Isotopomer tracer experiments using less common carbon sources can be very costly because of the prices of purposefully labelled substrates. Thus, it is very useful to be able to first conduct *in silico* calculability analysis to find out, whether it is even in principle possible to obtain the required flux information from the tracer experiment. By analyzing the fragment equivalence classes, it is relatively easy to perform this kind of "structural identifiability analysis" (cf. [39,40] for global isotopomer balancing), that is, to discover the set of junction metabolites for which the flux ratios can be calculated (in the best case) from the given measurement data: it is enough to check what type of isotopomer constraints

$$\sum_b s_{bij} P_{M_{ij}}(b) = d_{ij}, \quad (9)$$

can be propagated to each subpool  $M_{ij}$  of junction metabolites  $M_i$  from the measured metabolites (we need to know only coefficients  $s_{bij}$ , not the isotopomer abundances  $d_{ij}$ ). Then, by applying the techniques of computing vector subspace intersection described above, we can compute the maximal number of linearly independent constraints obtainable for the flux ratios of each junction. Thus, it is possible to check before costly and time-consuming wet lab experiments, whether the experiments even have potential to answer the biological questions at hand. The results of the calculability analysis tell which flux ratios are in principle determinable, given the labelling of external substrates, topology of the metabolic network and the available measurement data. It then depends on the actual flux distribution and the accuracy of the measurements, whether these ratios can be reliably determined from the experimental data.

#### Estimating the flux distribution of the metabolic network

In the main step of our framework for  $^{13}\text{C}$  metabolic flux analysis, the fluxes of the metabolic network are estimated by forming and solving generalized isotopomer balance equations (4). The generalized isotopomer balance equations are based on the isotopomer measurement data that is first propagated in the network to unmeasured metabolites by utilizing the results of the structural analysis presented above.

#### Measurement propagation

The aim of the propagation of measurement data is to infer from the isotopomer constraints of measured metabolites as many isotopomer constraints as possible to unmeasured metabolites. As a rule of thumb, more constraints the unmeasured metabolites will get more gener-

alized isotopomer balance equations (4) bounding the fluxes can be written.

Fragment equivalence classes can be utilized in the measurement propagation: from isotopomer constraints known for fragment  $M_i|F$  isotopomer constraints for other fragments  $M_l|F_k$  in the equivalence class of  $F$  can be easily computed. The process is the following:

1. Before measurements are propagated from fragment  $M|F$  of measured metabolite  $M$  to other fragments in the equivalence class of  $F$ , isotopomer constraints to  $F$  are computed from the constraints measured to the whole metabolite  $M$  by using the vector space projection techniques (see Section *Projection of isotopomer measurements to fragments*).
2. The fragment constraints are propagated to all fragments  $F'$  that have been found equivalent to  $F$  via the flow analysis technique.

This requires mapping of isotopomers of  $F$  to isotopomers of  $F'$  by applying the carbon mappings of the reactions along any pathway between  $F$  and  $F'$ .

3. After the propagation of measurement data inside the fragment equivalence classes, new isotopomer constraints for independent fragments of the same metabolite can be derived, as described in Section *Independence analysis of fragments*.

Steps 2 and 3 can be iterated until no new isotopomer constraints to the fragments are discovered.

#### Construction of generalized isotopomer balances

After the propagation step, we have some isotopomer constraints  $S_{ij} D_{M_{ij}} = d_{ij}$  for each subpool  $j$  of every junction metabolite  $M_i$ . (For non-junction metabolites, isotopomer balance equations do not contain any additional flux information compared to the mass balances.) In the best case we know complete isotopomer distribution  $D_{M_{ij}}$ , in the worst case we have only trivial constraints stating that the sum of relative abundances of all isotopomers equals one.

Next, a linear equation system containing flux constraints obtained from mass balances (1) and generalized isotopomer balances (4) is constructed.

However, the isotopomer constraints of different subpools do not yet conform to (4) as the matrices  $S_{ij}$  are not necessarily the same.

Thus we still need to compute a common subspace  $\mathcal{Y}_i = \bigcap_j \mathcal{S}_{ij}$  ( $\mathcal{S}_{ij}$  is spanned by the rows of  $S_{ij}$ ) of the isotopomer constraints known for each subpool  $M_{ij}$  and project subpool constraints  $S_{ij}D_{M_{ij}} = d_{ij}$  to  $\mathcal{Y}_i$ .

This can be done with the same techniques that were previously applied to project measured isotopomer information of a metabolite to its fragments. Let  $Y_i$  be a matrix with row space  $\mathcal{Y}_i$ . After the projection we obtain isotopomer constraints  $Y_i D_{M_{ij}} = z_{ij}$  for each subpool  $M_{ij}$  (See Figure 8 for an example).

Now the isotopomer constraints of all the subpools lie in the same subspace of  $\mathcal{I}_{M_i}$  and we are ready to write the system of generalized isotopomer balance equations (4) for every junction  $M_i$ :

$$\sum_{j=1}^n \alpha_{ij} v_j z_{ij} = \beta_i z_{i0}, \quad (10)$$

that is,

$$A_i v = \begin{bmatrix} \alpha_{1i}(z_{1i})_1 & \cdots & \alpha_{ni}(z_{ni})_1 \\ \vdots & \ddots & \vdots \\ \alpha_{1i}(z_{1i})_r & \cdots & \alpha_{ni}(z_{ni})_r \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = g_i, \quad (11)$$

where  $g_i = \beta_i z_{i0}$ .

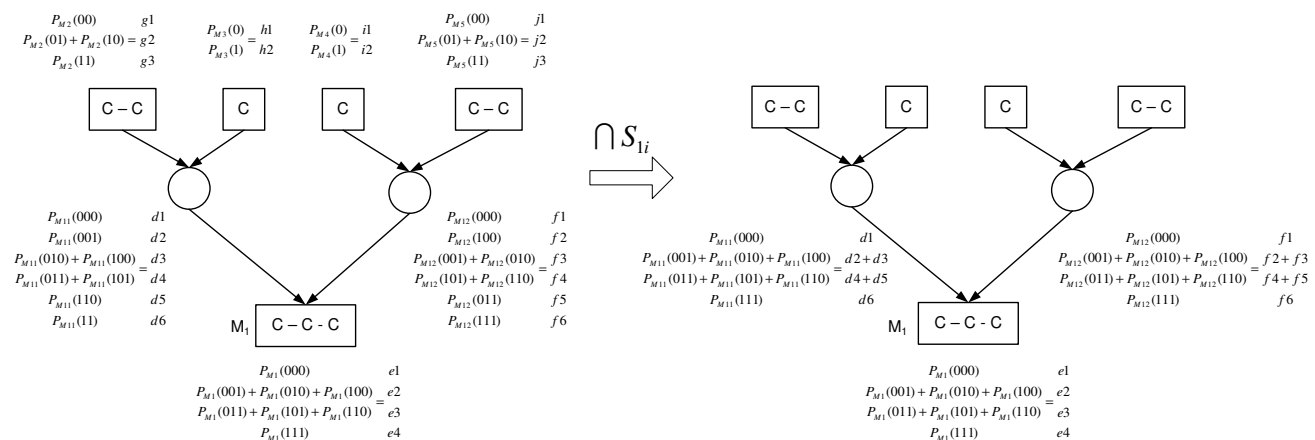
**Estimating the fluxes**

The ratios of (forward) fluxes producing  $M_i$  can be computed by solving the corresponding Equation (11) augmented with a constraint that fixes the out flow from  $M_i$  to equal 1. Thus, we obtain flux ratios of junction metabolites without manual derivation of ratio equations, without nonlinear optimization and without knowing intake and outtake rates of external metabolites or biomass composition.

In addition, when the equations (11) of all junction metabolites are combined with the mass balances (1) of non-junctions, we obtain a linear equation system

$$A v = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix} = g \quad (12)$$

constraining the fluxes  $v$  of the network that contains a block (junctions) or a row (non-junctions)  $A_k$  for each metabolite  $M_k$ . Measured external fluxes and other known constraints, such as the composition of biomass can also



**Figure 8**  
**An example of the computation of the common subspace of isotopomer constraints in different subpools.** The mass isotopomer distribution of junction metabolite  $M_1$  is assumed to be measured. For the in flow subpools  $M_{11}$  and  $M_{12}$  we obtain isotopomer constraints from the above reactant metabolites by measurement propagation. These propagated constraints must be projected to mass isotopomer to the subspace defined by the mass isotopomer distribution of  $M_1$  before generalized isotopomer balances are constructed.



be added to (12) as additional constraints. Additional constraints, like ones derived from gene regulatory information [41] or from thermodynamic analysis of metabolism [42-44] can also easily be included to (12).

If (12) is of full rank, the whole flux distribution can be solved with standard linear algebra [45]. Also, more complex, nonlinear methods can be applied to model the effect of experimental errors to the estimated flux distribution [20]. In a common case where the system is of less than full rank, a single flux distribution can not be pinpointed without additional constraints. Instead, (12) defines the space of feasible flux distributions, that are all equally good solutions. In that case we can apply techniques developed for the analysis of stoichiometric matrices to determine as many fluxes as possible [46] from (12). More generally, by linear programming we can obtain maximum (resp. minimum) values for each flux  $v_i$ :

$$\begin{aligned} &\text{For each } v_i : \\ &\quad \max \quad v_i \\ &\quad \text{s.t. } \quad Av = g \\ &\quad \quad v_i^{\min} \leq v_i \leq v_i^{\max} \forall v_i \in v, \end{aligned} \quad (13)$$

where  $v_i^{\min}$  and  $v_i^{\max}$  are predetermined minimum and maximum allowable values for  $v_i$

Furthermore, it is possible to search for in some sense optimal flux distribution – for example a flux distribution maximizing the production of biomass – from the feasible space defined by (12) by linear programming techniques of flux balance analysis [1,3,47,48]. In that case, isotopomer data constrain the feasible space more than the stoichiometric information would alone do, thus possibly allowing more accurate estimations of the real flux distribution.

### Statistical analysis

For an experimentalist, it is important to know how sensitive the obtained estimation of fluxes is to measurement errors. If enough repeated measurements are not available to assess this sensitivity, it has to be estimated by computational techniques. In the global isotopomer balancing framework for  $^{13}\text{C}$  metabolic flux analysis, many mathematically or computationally involved methods have been developed to analyze the sensitivity of estimated flux distributions to errors in isotopomer measurements and the sensitivity of the objective function to the changes in the generated candidate flux distributions [49-53].

As our direct method for  $^{13}\text{C}$  metabolic flux analysis is computationally efficient, we can afford to a simple, yet powerful Monte Carlo procedure to obtain estimates on

the variability of individual fluxes due to measurement errors:

1. For each measured metabolite  $M_i$ : By studying the variability in the repeated measurements, fix the distribution  $\Omega_i$  from which the measurements of  $M_i$  are sampled.
2. Repeat  $k$  times:
  - (a) For each measured metabolite  $M_i$ : sample a measurement from  $\Omega_i$ .
  - (b) Estimate fluxes  $v_i$  from the sampled measurements.
3. Compute appropriate statistics from the set  $V = \{v_1, \dots, v_k\}$  to describe the sensitivity of fluxes to measurement errors.

Possible statistics that can be applied in the last step of the above algorithm include standard deviation, empirical confidence intervals [53], kurtosis, standard error etc. of each individual flux  $v_j$  and measures of "compactness" of  $V$ , such as (normalized) average distance of items of  $V$  from the sample average.

### Experimental NMR and GC-MS methods

In this section we shortly describe the experimental procedures applied in NMR and GC-MS isotopomer measurements that produced the data for Section.

In the first experiment *S. cerevisiae* was grown in an aerobic glucose-limited chemostat culture at dilution rate  $0.1 \text{ h}^{-1}$ . After reaching a metabolic steady state, as determined by constant physiological parameters 10% of the carbon source in the medium was replaced with fully carbon labelled glucose ([U- $^{13}\text{C}$ ]) for approximately 1.5 residence times, so that about 78% of the biomass was uniformly labelled. 2D [ $^{13}\text{C}$ ,  $^1\text{H}$ ] COSY spectra of harvested and hydrolysed biomass were acquired for both aliphatic and aromatic resonances at  $40^\circ\text{C}$  on a Varian Inova 600 MHz NMR spectrometer. The software FCAL v.2.3.0 [19] was used to compute isotopomer constraints for 15 amino acids from the spectra. Detailed description of the cultivation set up can be found in [54] whereas similar  $^{13}\text{C}$  labeling set up, NMR experiments and spectral data analysis as were applied here have been described in [55].

In the second experiment *B. subtilis* was grown on shake flasks containing 50 ml M9 minimal medium. In the experiment, the medium was supplemented with 50 mg/L tryptophan and 3 g/L glucose labelled to the first carbon position ([1- $^{13}\text{C}$ ]) (99%; Cambridge Isotope Laboratories) or a mixture of 0.6 g/L fully carbon labelled glucose ([U- $^{13}\text{C}$ ]) (99%; Cambridge Isotope Laboratories) and 2.4 g/L unlabeled glucose as the sole carbon source. Four-

teen derivatized amino acids were analyzed for  $^{13}\text{C}$  labeling patterns with a series 8000GC combined with an MD800 mass spectrometer (Fisons instruments). More information about the details of the measurement procedure can be found from [20] where identical measurement techniques were applied.

### Authors' contributions

AR, JR and EU are the main contributors of the computational methods given in the article. AR implemented the methods and ran the experiments. PJ constructed the models of metabolic networks and contributed to the development of the computational methods. NZ and HM contributed the development of the computational methods. AR, JR, NZ and PJ wrote the manuscript. All authors have read and accepted the contents of the manuscript.

### Additional material

#### Additional file 1

*Model of central carbon metabolism of S. cerevisiae in SBML format. Additional file 1 contains the stoichiometric model of central carbon metabolism of S. cerevisiae used in the experiments of the article. The model is provided as a text file containing SBML compliant descriptions of metabolites and reactions in the model. Metabolites are identified by their KEGG LIGAND codes.*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2105-9-266-S1.txt>]

#### Additional file 2

*Illustration of the model of central carbon metabolism of S. cerevisiae. Bidirectional reactions in PPP pathways are depicted as unidirectional for better readability.*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2105-9-266-S2.pdf>]

#### Additional file 3

*Model of central carbon metabolism of B. subtilis in SBML format. Additional file 2 contains the stoichiometric model of central carbon metabolism of B. subtilis used in the experiments of the article. The model is provided as a text file containing SBML compliant descriptions of metabolites and reactions in the model. Metabolites are identified by their KEGG LIGAND codes.*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2105-9-266-S3.txt>]

#### Additional file 4

*Illustration of the model of central carbon metabolism of B. subtilis. Bidirectional reactions in PPP pathways are depicted as unidirectional for better readability.*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2105-9-266-S4.pdf>]

### Acknowledgements

This research has been supported by Academy of Finland (SYSBIO programme, grant 207436, SYSFYS project). We thank Uwe Sauer and Merja Penttilä for providing the data and for their comments to the work presented in the article; Markus Heinonen, Esa Pitkänen and Arto Åkerlund for the implementation of software tools supporting the presented framework; Roelco Kleijn, Sarah-Maria Fendt, Simon Tännler and Martin Rühl for providing the data and for fruitful discussions. We also thank the anonymous reviewers whose comments helped us to improve the manuscript substantially.

### References

1. Varma A, Palsson B: **Metabolic flux balancing: basic concepts, scientific and practical use.** *Nature Biotechnology* 1994, **12(10)**:994-998.
2. Edwards JS, Ibarra RU, Palsson B: **In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data.** *Nature Biotechnology* 2001, **19(2)**:125-130.
3. Schütz R, Kuepfer L, Sauer U: **Systematic evaluation of objective functions for predicting intracellular fluxes in E. coli.** *Molecular Systems Biology* 2007, **3(119)**.
4. Szyperki T: **Biosynthetically directed fractional  $^{13}\text{C}$ -labelling of proteinogenic amino acids.** *European Journal of Biochemistry* 1995, **232(2)**:433-448.
5. Dauner M, Sauer U: **GC-MS analysis of amino acids rapidly provides rich information for isotopomer balancing.** *Biotechnology Progress* 2000, **16(4)**:642-649.
6. Rousu J, Rantanen A, Ketola R, Kokkonen J: **Isotopomer distribution computation from tandem mass spectrometric data with overlapping fragment spectra.** *Spectroscopy* 2005, **19**:53-67.
7. Schmidt K, Carlsen M, Nielsen J, Viladsen J: **Modeling Isotopomer Distributions in Biochemical Networks Using Isotopomer Mapping Matrices.** *Biotechnology and Bioengineering* 1997, **55(6)**:831-840.
8. Wiechert W, Petersen Möllney S, de Graaf A: **A Universal Framework for  $^{13}\text{C}$  Metabolic Flux Analysis.** *Metabolic Engineering* 2001, **3(3)**:265-283.
9. Wiechert W, de Graaf A: **Bidirectional Reaction Steps in Metabolic Networks: I. Modeling and simulation of carbon isotope labeling experiments.** *Biotechnology and Bioengineering* 1997, **55(1)**:101-117.
10. Antoniewicz M, Kelleher J, Stephanopoulos G: **Elementary Metabolite Units (EMU): a novel framework for modeling isotopic distributions.** *Metabolic Engineering* 2007, **9(1)**:68-86.
11. van Winden W, Heijnen J, Verheijen P: **Cumulative bondomers: a new concept in flux analysis from 2D [ $^{13}\text{C}$ ,  $^1\text{H}$ ] COSY NMR data.** *Biotechnology and Bioengineering* 2002, **80(7)**:731-745.
12. Nöh K, Wvahl A, Wiechert W: **Computational tools for isotopically instationary  $^{13}\text{C}$  labeling experiments under metabolic steady state conditions.** *Metabolic Engineering* 2006, **8(6)**:554-577.
13. Nöh K, Grönke K, Luo B, Takors R, Oldiges M, Wiechert W: **Metabolic flux analysis at ultra short time scale: Isotopically non-stationary  $^{13}\text{C}$  labeling experiments.** *Journal of Biotechnology* 2007, **129(2)**:249-267.
14. Young J, Walther J, M A, Yoo H, Stephanopoulos G: **An elementary metabolite unit (EMU) based method of isotopically nonstationary flux analysis.** *Biotechnology and Bioengineering* 2007, **99(3)**:686-699.
15. Ghosh S, Zhu T, Grossmann I, Ataai M, Domach M: **Closing the loop between feasible flux scenario identification for construct evaluation and resolution of realized fluxes via NMR.** *Computers and Chemical Engineering* 2005, **29(3)**:459-466.
16. Blank L, Kuepfer L, Sauer U: **Large-scale  $^{13}\text{C}$ -flux analysis reveals mechanistic principles metabolic network robustness to null mutations in yeast.** *Genome Biology* 2005, **6(6)**:R49.
17. Blank L, Lehmebeck F, Sauer U: **Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts.** *FEMS Yeast Research* 2005, **5(6-7)**:545-558.
18. Fischer E, Sauer U: **Metabolic flux profiling of Escherichia coli mutants in central carbon metabolism using GC-MS.** *European Journal of Biochemistry* 2003, **270(5)**:880-891.

19. Szyperski T, Glaser R, Hochuli M, Fiaux J, Sauer U, Bailey J, Wüthrich K: **Bioreaction Network Topology and Metabolic Flux Ratio Analysis by Biosynthetic Fractional <sup>13</sup>C Labeling and Two-Dimensional NMR Spectrometry.** *Metabolic Engineering* 1999, **1(3)**:189-197.
20. Fischer E, Zamboni N, Sauer U: **High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived <sup>13</sup>C constraints.** *Analytical Biochemistry* 2004, **325(2)**:308-316.
21. Rousu J, Rantanen A, Maaheimo H, Pitkänen E, Saarela K, Ukkonen E: **A method for estimating metabolic fluxes from incomplete isotopomer information.** *Computational Methods in Systems Biology, Proceedings of the First International Workshop, CMSB Volume 2602 of Lecture Notes in Computer Science* 2003:88-103.
22. Sauer U, Hatzimanikatis V, Bailey J, Hochuli M, Szyperski T, Wüthrich K: **Metabolic fluxes in riboflavin-producing *Bacillus subtilis*.** *Nature Biotechnology* 1997, **15(5)**:448-452.
23. Maaheimo H, Fiaux J, Cakar Z, Bailey J, Sauer U, Szyperski T: **Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional <sup>13</sup>C labeling of common amino acids.** *European Journal of Biochemistry* 2001, **268(8)**:2464-2479.
24. Zamboni N, Sauer U: **Knockout of the high-coupling cytochrome aa3 oxidase reduces TCA cycle fluxes in *Bacillus subtilis*.** *FEMS Microbiology Letters* 2003, **226(1)**:121-126.
25. Sola A, Maaheimo H, Ylönen K, Ferrer P, Szyperski T: **Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*.** *European Journal of Biochemistry* 2004, **271(12)**:2462-2470.
26. Sola A, Jouhten P, Maaheimo H, Sanchez-Ferrando F, Szyperski T, Ferrer P: **Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates.** *Microbiology* 2007, **153(1)**:281-290.
27. Rantanen A, Maaheimo H, Pitkänen E, Rousu J, Ukkonen E: **Equivalence of metabolite fragments and flow analysis of isotopomer distributions for flux estimation.** *Transactions on Computational Systems Biology VI, Lecture Notes in Bioinformatics* 2006, **4220**:198-220.
28. Kleijn R, Geertman J, Nfor B, Ras C, Schipper D, Pronk J, Heijnen J, van Maris A, van Winden W: **Metabolic flux analysis of a glycerol-overproducing *Saccharomyces cerevisiae* strain based on GC-MS, LC-MS and NMR derived <sup>13</sup>C-labeling data.** *FEMS Yeast Research* 2006, **7(2)**:216-231.
29. Arita M: **In Silico Atomic Tracing of Substrate-Product Relationships in *Escherichia coli* Intermediary Metabolism.** *Genome Research* 2003, **13(11)**:2455-2466.
30. Zamboni N, Fischer E, Sauer U: **FiatFlux – a software for metabolic flux analysis from <sup>13</sup>C-glucose experiments.** *BMC Bioinformatics* 2005, **6(209)**:
31. Szyperski T: **<sup>13</sup>C-NMR, MS and metabolic flux balancing in biotechnology research.** *Quarterly Reviews of Biophysics* 1998, **31(1)**:41-106.
32. Weitzel M, Wiechert W, Nöh K: **The topology of metabolic isotope labeling networks.** *BMC Bioinformatics* 2007, **8(315)**:
33. van Winden W, van Dam J, Ras C, Kleijn R, Vinke J, Gulik W, Heijnen J: **Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on mass isotopomer measurements of <sup>13</sup>C-labeled primary metabolites.** *FEMS Yeast Research* 2005, **5(6-7)**:559-568.
34. Rantanen A, Mielikäinen T, Rousu J, Maaheimo H, Ukkonen E: **Planning optimal measurements of isotopomer distributions for estimation of metabolic fluxes.** *Bioinformatics* 2006, **22(10)**:1198-1206.
35. Schmidt K, Nielsen J, Villadsen J: **Quantitative analysis of metabolic fluxes in *Escherichia coli*, using two-dimensional NMR spectroscopy and complete isotopomer models.** *Journal of Biotechnology* 1999, **71(1-3)**:175-189.
36. Rantanen A, Rousu J, Ketola R, Kokkonen J, Tarkiainen V: **Computing Positional Isotopomer Distributions from Tandem Mass Spectrometric Data.** *Metabolic Engineering* 2002, **4**:285-294.
37. Aho A, Hopcroft J, Ullman J: *The Design and Analysis of Computer Algorithms* Addison Wesley; 1974.
38. Appel A: *Modern Compiler Implementation in Java* Cambridge University Press; 1998.
39. Isermann N, Wiechert W: **Metabolic isotopomer labeling systems. Part II: structural identifiability analysis.** *Mathematical Biosciences* 2003, **183(2)**:175-214.
40. van Winden W, Heijnen J, Verheijen P, Grievink J: **A priori analysis of metabolic flux identifiability from <sup>13</sup>C-labeling data.** *Biotechnology and Bioengineering* 2001, **74(6)**:505-516.
41. Covert M, Schilling C, Palsson B: **Regulation of Gene Expression in Flux Balance Models of Metabolism.** *Journal of Theoretical Biology* 2001, **213(1)**:73-88.
42. Kümmel A, Panke S, Heinemann M: **Systematic assignment of thermodynamic constraints in metabolic network models.** *BMC Bioinformatics* 2006, **7(512)**:
43. Henry C, Broadbelt L, Hatzimanikatis V: **Thermodynamics-Based Metabolic Flux Analysis.** *Biophysical Journal* 2007, **92(5)**:1792-1805.
44. Hoppe A, Hoffmann S, Holzhütter H: **Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks.** *BMC Systems Biology* 2007, **1(23)**:
45. Schwarz H: *Numerical Analysis: A Comprehensive Introduction* John Wiley & Sons; 1989.
46. Klamt S, Schuster S: **Calculating as many fluxes as possible in underdetermined metabolic networks.** *Molecular Biology Reports* 2002, **29(1-2)**:243-248.
47. Bonarius H, Schmidt G, Tramper J: **Flux analysis of underdetermined metabolic networks: the quest for the missing constraints.** *Trends in Biotechnology* 1997, **15(8)**:308-314.
48. Edwards J, Covert M, Palsson B: **Metabolic modelling of microbes: the flux-balance approach.** *Environmental Microbiology* 2002, **4(3)**:133-140.
49. Möllney M, Wiechert W, Kownatzki D, de Graaf A: **Bidirectional Reaction Steps in Metabolic Networks IV: Optimal Design of Isotopomer Labeling Experiments.** *Biotechnology and Bioengineering* 1999, **66(2)**:86-103.
50. Kleijn R, van Winden W, Ras C, van Gulik W, Schipper D, Heijnen J: **<sup>13</sup>C-Labeled Gluconate Tracing as a Direct and Accurate Method for Determining the Pentose Phosphate Pathway Split Ratio in *Penicillium chrysogenum*.** *Applied and Environmental Microbiology* 2006, **72(7)**:4743-4754.
51. Araújo-Bravo M, Shimizu K: **An improved method for statistical analysis of metabolic flux analysis using isotopomer mapping matrices with analytical expressions.** *Journal of Biotechnology* 2003, **105(1-2)**:117-133.
52. Wiechert W, Siefke C, de Graaf A, Marx A: **Bidirectional Reaction Steps in Metabolic Networks: II. Flux Estimation and Statistical Analysis.** *Biotechnology and Bioengineering* 1997, **55(1)**:118-134.
53. Antoniewicz M, Kelleher J, Stephanopoulos G: **Determination of confidence intervals of metabolic fluxes estimated from stable isotopomer measurements.** *Metabolic Engineering* 2006, **8(4)**:324-337.
54. Wiebe M, Rintala E, Tamminen A, Simolin H, Salusjärvi L, Toivari M, Kokkonen J, Kiuru J, Ketola R, Jouhten P, Huuskonen A, Maaheimo H, Ruohonen L, Penttilä M: **Central carbon metabolism of *Saccharomyces cerevisiae* in anaerobic, oxygen-limited and fully aerobic steady-state conditions and following a shift to anaerobic conditions.** *FEMS Yeast Research* 2008, **8(1)**:140-154.
55. Fiaux J, Cakar P, Sonderegger M, Wüthrich K, Szyperski T, Sauer U: **Metabolic-Flux Profiling of the Yeasts *Saccharomyces cerevisiae* and *Pichia stipitis*.** *Eukaryotic Cell* 2003, **2(1)**:170-180.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
http://www.biomedcentral.com/info/publishing\_adv.asp

