Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Microbiology

# Novel biological insights through metabolomics and $^{13}$C-flux analysis

Nicola Zamboni and Uwe Sauer

Metabolomics and $^{13}$C-flux analysis have become instrumental for analyzing cellular metabolism and its regulation. Driven primarily by technical advances in mass spectrometry-based analytics, they provide unmatched readouts on metabolic state and activity. Functional genomics leverages metabolomics for the discovery of novel enzymes and unexpected secondary activities of annotated enzymes. $^{13}$C-flux analyses are frequently used for empirical elucidation of pathways in poorly characterized species and for network-wide analysis of mechanisms that realize energy and redox balancing. Integration of metabolomics, $^{13}$C-flux analysis and other data enable the condition-dependent characterization of regulatory circuits that ultimately govern the metabolic phenotype.

**Address**
Institute of Molecular Systems Biology, ETH Zurich, 8093 Zurich, Switzerland

Corresponding author: Sauer, Uwe (sauer@ethz.ch)

**Current Opinion in Microbiology** 2009, **12**:553–558

This review comes from a themed issue on
Genomics
Edited by Carmen Buchrieser and Stewart Cole

Available online 8th September 2009

1369-5274/$ – see front matter


**DOI** 10.1016/j.mib.2009.08.003

## Introduction

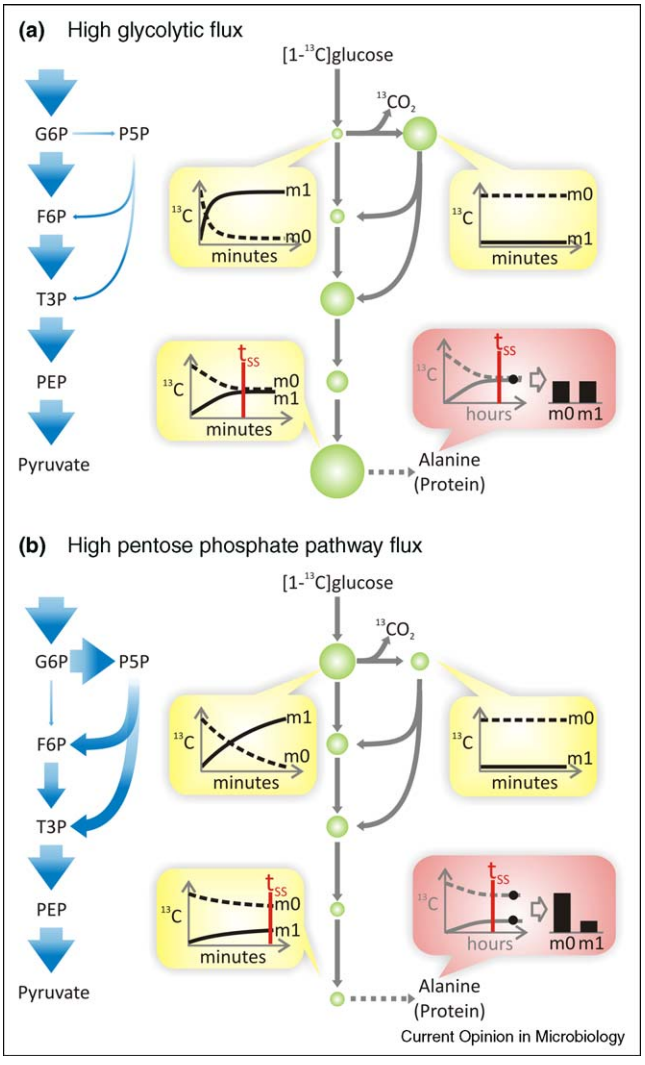
Getting closer to the whole picture of metabolism requires — among other measurements and mechanistic modeling — global analysis of metabolite concentrations and *in vivo* reaction fluxes [1], commonly referred to as metabolomics and $^{13}$C-flux analysis, respectively. While individual metabolite concentrations have been measured for a long time, metabolomics relates to the analysis of the full small-molecule content of cells (or other biological samples), and is largely driven by analytical advances of mass spectrometry (MS) [2]. Conceptually different from the concentration-based 'omics', metabolic fluxes are the time-dependent passage of metabolites through a reaction (i.e. *in vivo* reaction rates). Since intracellular fluxes cannot be determined directly from the concentration of any cellular component, they must be determined from either firstly, local time course data of dynamic stable isotope enrichment in pathway intermediates or secondly, mathematically inferred from measurable quantities such as macroscopic uptake and production rates and steady state $^{13}$C-labeling data (Figure 1). The difference between metabolite concentrations and fluxes is roughly equivalent to the relationship between the number of cars on a street versus their traffic pattern [3]. After an update on technological developments since 2007, we focus here on applications of metabolomics and flux methods to bacteria that address biological key questions in functional genomics, pathway identification, regulation networks, and biotechnology.

## Technological development

Over the past two years, microbial metabolomics experienced a consolidation with a particular focus on increasing metabolite coverage and improving quantification. The currently best quantitative methods are MS-based and detect up to 100–200 cellular metabolites [4], in rigorous applications even with experimental validation of peak identity [5•,6•,7]. Such high coverage is typically achieved by combining multiple chromatographic and/or detection methods that specialize on different compound classes [5•]. In the so far largest bacterial example, 198 hydrophilic metabolites could be quantified in *Escherichia coli* with three capillary electrophoresis time-of-flight (CE-TOF)MS methods [6•]. A yet unsolved problem remains quantitative sample preparation from liquid cultures (i.e. rapid quenching) because metabolites leak from bacteria — but not yeast — when exposed to organic solvents or cold environments [8]. The currently most reliable workaround is differential measurement of concentrations in whole broth and cell-free broth aliquots [9•]. At least when the concentration of metabolites of interest is low in the medium, this approach delivers accurate results. A major challenge is the abundance of salts in broth extracts that can severely affect chromatography, sample derivatization, and ionization in a compound-dependent manner [7]. These detrimental effects are efficiently counteracted individually for each analyte by using fully $^{13}$C-labeled biomass extracts as internal standard [7,10].

The above quantitative and typically targeted approaches rely on chromatographic separation before MS detection, hence analysis of one sample is on the order of 30 min. Without chromatography, direct acquisition of unseparated metabolite extracts with MS emerges as an orthogonal strategy for high throughput at the cost of isomer separation. In combination with either electrospray, desorptive, or extractive ionization, mass spectrometers with

**Figure 1**




Schematic view of metabolome, fluxes and $^{13}$C-data obtained for two hypothetical strains with predominant glycolytic or pentose phosphate pathway flux. Fluxome (blue arrows) and metabolome (green bubbles) are interconnected by poorly known enzyme kinetics, which is insufficient to directly calculate fluxes from concentrations of intermediates or vice versa. Fluxes have to be measured indirectly by $^{13}$C-experiments with stable tracer isotopes. For example, the complementary catabolic fluxes through glycolysis and pentose phosphate pathway are well resolved with [1 – $^{13}$C]glucose because the label is lost as $^{13}CO_2$ in the pentose phosphate pathway. The most detailed flux information is obtained from dense and local time course data of dynamic stable isotope enrichment (yellow callouts), whose trajectories depend on flux and metabolome. Stationary $^{13}$C-flux methods estimate fluxes from $^{13}$C-pattern at isotopic steady state ($t_{ss}$). This is technically simpler because abundant end-products such as protein-bound amino acids (red callouts) can be analyzed instead of pathway intermediates with short lifetime, but comes with the caveat that long labeling times are necessary to attain isotopic stationarity.

high-resolving power such as modern TOF instruments or Fourier transform ion traps enable relative quantification of potentially hundreds of intracellular compounds with distinct molecular weight [11,12]. This strategy promises unique opportunities for discovery and screening of metabolic activities, although it is currently limited by the low fraction of detectable MS signals that can be assigned to structures or metabolites. Identification of ions is still a challenging and time-demanding task [13], but considering pre-existing knowledge on microbial enzymes can substantially help [14].

Increased sensitivity and coverage of metabolomics methods is also a key advance for $^{13}$C-flux analysis because it enables direct detection of labeling patterns in pathway intermediates [15]. The higher turnover rates of intermediates and their better distribution in metabolism, relative to the eight carbon precursors for the traditionally used protein-bound amino acids, greatly improves temporal and pathway resolution. In particular the initial dynamics of label propagation through pathway intermediates allow to assess flux responses that cannot be resolved by traditional stationary flux analysis [16] (Figure 1). Such dynamic experiments have recently enabled major biological discoveries in pathway discovery [17••] and biotechnology [18••] that are reviewed later. These analytical developments are accompanied by algorithmic improvements to facilitate the integration of dynamic data, with the Wiechert lab method of the highest potential for general applicability [19].

## Functional genomics

A key expectation for just about any omics technology is to assign functions for the still about 30% unassigned genes within most genomes. In contrast to the common notion of a well understood metabolism, about 30–40% of the metabolic activities are still without known enzymes or genes in any organism [20•]. Slowly but continuously, gene functions are empirically unraveled, and the value of $^{13}$C-flux and metabolomics methods is elucidation of such activities within the network context [21,22].

For at least the known part of metabolism, however, more systematic, model-driven approaches have been developed that start from known or suspected enzymatic reactions that are not yet linked to a gene; that is about 20% of such orphan activities exist in typical genome-scale models of microbial metabolism [23]. To rank their likelihood and suggest candidate genes for experimental testing, an algorithm has been developed that combines the known network structure with multiple types of functional association evidence, including genome position, similarity of phylogeny, and co-expression [20•]. This gap-filling approach was successfully demonstrated for the identification of the broadly distributed *yneI* gene to encode one of the two succinate semialdehyde dehydrogenase activities in *E. coli* [24]. A related, optimization-based algorithm predicts missing reactions that are required to reconcile genome-scale model prediction and high-throughput phenotypic data, thereby assigning