Eur. J. Biochem. **270**, 579–588 (2003) © FEBS 2003 doi:10.1046/j.1432-1033.2003.03427.x

MINIREVIEW

Deciphering metabolic networks

Oliver Fiehn and Wolfram Weckwerth

Max-Planck Institute of Molecular Plant Physiology, 14424 Potsdam/Golm, Germany

All higher organisms divide major biochemical steps into different cellular compartments and often use tissue-specific division of metabolism for the same purpose. Such spatial resolution is accompanied with temporal changes of metabolite synthesis in response to environmental stimuli or developmental needs. Although analyses of primary and secondary gene products, i.e. transcripts, proteins, and metabolites, regularly do not cope with this spatial and temporal resolution, these gene products are often observed to be highly coregulated forming complex networks. Methods to study such networks are reviewed with respect to data acquisition, network statistics, and biochemical interpretation.

Keywords: metabolomics; proteomics; protein networks; metabolite networks; metabolite profiling.

Introduction

Now that a variety of unicellular and multicellular genomes have successfully been sequenced and (partially) annotated (see The Arabidopsis Genome Initiative, 2000 [1,2]), functional genomics has become a focal point for many research efforts. For most genomes a significant number of genes cannot be annotated by homology to genes in other organisms, and (with the exception for yeast) for the majority of the annotated genes, no experimental proof is supporting these annotations. Moreover, if focusing on metabolism, many levels of regulation occur after genes have been transcribed, such as post-transcriptional, translational, post-translational, and all forms of biochemical control such as allosteric or feedback regulation. Taking this view into account, it is hard to believe that functional genomics can stop at the mRNA level. Instead, systems biology approaches need to be undertaken that comprehensively analyse the structure of cellular organization and try to model all acquired data with the ultimate aim to be able to predict the effect of any perturbation of the system (e.g. by gene deletion or as response to physiological alterations). This view necessarily demands a new paradigm: the analysis of multiple interactions of gene products in biological networks rather than assuming simple hierarchical control of genes over metabolic pathways. In a recent case study, control of glycolysis was shown to be shared between metabolic, proteomic, or genomic levels [3]. It might well be that substrate availability and protein clusters turn out to challenge the classical view of 'fixed' metabolic pathways. Instead, there might be preferred routes [4] through biochemical networks that usually utilize shortest paths between two nodes, but that might change in response to altered conditions or needs. In the following, methods and applications towards comprehensive analysis of biochemical networks will be summarized from the level of proteomic and metabolomic data acquisition to network computation.

Correspondence to O. Fiehn, Max-Planck Institute of Molecular Plant Physiology, 14424 Potsdam/Golm, Germany.
E-mail: fiehn@mpimp-golm.mpg.de

Abbreviations: 2-DE, two-dimensional gel electrophoresis; CE, capillary electrophoresis; FT-MS, Fourier-transform ion cyclotron mass spectrometry.

(Received 18 September 2002, accepted 12 December 2002)

Quantitative proteomics

Besides a qualitative description of proteins expressed in an organism, the need for quantification of protein abundance in response to specific perturbations is a prerequisite in systems biology. Protein quantification has even more importance due to the fact that correlation between RNA levels and protein levels has been shown to be remarkable low in all recent studies so far [5,6]. In that respect, post-translational regulation is discussed to be a major regulatory event in metabolism, but techniques to identify and quantify post-translational modification of proteins on a systems level are still at a preliminary state of development [7–12].

2D Gels combined with mass spectrometry, usually MALDI-TOF, allow detection and identification of a large number of proteins from a tissue, and the comparison of protein profiles in different tissues, different genotypes or after different treatments [13,14]. At present, a systematic programme is sequencing proteins running at different positions in 2D gels to allow more rapid interpretation of the results. In addition to protein identification, 2D gel technology can be combined with the use of radiolabelling of the tissue before extraction, and subsequent autoradiography. Incubation with ^{32}P will label proteins that are subject post-translational regulation by phosphorylation [15].

Nevertheless, current approaches by two-dimensional gel electrophoresis (2-DE) cannot fulfill the utmost requirements for proteomic research [16], i.e. a rapid, robust, sensitive and unbiased identification and quantification of complex proteomes. Mostly, researchers focus either on specific questions, for example soluble proteins from

isolated organelles [17,18] (see also the review by Dreger in this series) or they limit their interest to proteins that appear to be heavily up- or down regulated by looking for novel or disappearing spots which can then be cut and analysed. Quantifiying by measuring the intensity of stained spots suffer from limited loading material in combination with the large range of natural protein abundances. In an excellent overview approaches are summarized for using differential labelling in 2-DE for quantifying proteins [19].

Non-gel approaches exploit specific physical properties of proteins, such as isoelectric points (by liquid or capillary based isoelectric focusing), hydrophobicity (by novel column types such as monolithic or polystyrene microcolumns), charge state (by strong cation exchange microcolumns), affinity (by immunoprecipitation or affinity columns), or protein residues (by chemical linkers specific for protein moieties such as Cys containing peptide, or phosporylated Ser/Thr residues). Wall *et al.* have shown that protein mixtures can be fractionated using isoelectric focusing in the liquid phase [20,21]. If coupled to monolithic columns and protein digestion prior to mass spectrometry, they obtained identification of hundreds of proteins from cell cultures. In a study on relative protein quantification in yeast [22], Gygi, Aebersold and coworkers were able to show that method repeatability is better than 12% relative standard deviation when protein Cys-residues were linked to stable isotopically labelled chemicals that included biotin moieties for sample purification and preconcentration. However, the protocol became again quite tedious when propagated to the analysis of phosphoproteins, with only three phosphoproteins found in crude yeast cell lysates [23] although in general, labeling is also helpful for *de novo* peptide sequencing [24] and might pose an elegant alternative to classical Western blots for protein targeting [25].

Various groups used the β-elimination of organophosphate groups prior to stable isotope chemical linkers and LC/MS/MS [9,10,26], for the identification of phosphoproteins. Quite a number of phosphoproteins could be detected using classical labeling approaches by feeding cell cultures with ^{15}N [8]. One of the problems associated with chemical linkers is the high excess of reagents that cause high background noise in mass spectrometry. Another is the sheer number of peptides to be analysed. All commercially available mass spectrometers first perform a survey scan before ions are chosen in subsequent MS/MS fragmentation experiments. Regularly, only the most abundant ions in each scan get fragmented, so that minor components easily remain undetected. One way around was found by Yates who increased the run times of liquid chromatography in order to ensure that only a few peptides per time interval reach the mass spectrometer [27]. This was achieved by stepwise elution of the peptide mixture from strong cation exchangers onto nanoscale reverse phase columns prior to ion trap MS/MS experiments. By this strategy, he could reach a genome coverage of $\approx 25\%$ for yeast. In total, 1500 proteins were detected, ranging from low abundant transcription factors to proteins with up to 14 transmembrane domains, and high and low abundant enzymes. More importantly, Yates could also prove that no bias against codon usage was found in the detected proteins. In a subsequent study Washburn *et al.* extended this approach to quantify differential labelled peptides [28].

Besides these impressive results of comprehensive 2D chromatography coupled to mass spectrometry it is of crucial importance to provide direct comparisons of this novel technology with 2D gels to reveal advantages and disadvantages. It may be expected that both approaches will have complementary input for proteomic technology.

An alternative way around using MS and subsequent MS/MS experiments is to use mass spectrometers with ultimate sensitivity and mass resolution. In a series of related papers, Smith and coworkers observed up to 50 000 peptides in a single run by a combination of nanoLC columns [29–31], or capillary isoelectric focusing [32] prior to fourier-transform ion cyclotron mass spectrometry (FT-MS). Smith *et al.* integrated the labeling approach into proteomic experiments [33,34], and showed that up to a thousand proteins could be identified and quantified in 4 h runs. Using infrared laser photodissociation, peptides were fragmented within the cyclotron in order to confirm peptide identification that was solely based on accurate masses [35].

But how would all these proteins act together, and how are they assembled *in vivo*? In model systems such as yeast, pioneering studies have been able to characterize protein networks on the basis of intracellular associations using two hybrid systems [36]. In two recent studies on proteomic clusters, different approaches were chosen to detect protein–protein interactions. Ho *et al.* captured interacting proteins by immunoprecipitation of yeast bait proteins and subsequently on 2D SDS-polyacrylamide gel electrophoresis and mass spectrometry [37]. In a breakthrough study, Gavin *et al.* used tandem affinity purification and mass spectrometry observing 233 distinct multiprotein complexes in yeast [38], some of them consisting of enzymes that were previously believed to solely act in isolation. From these networks the potential to unravel connectivities via biochemical pathways have arisen [39,40], and it would be exciting to see combinations of quantitative proteomics, proteomic clustering, and results on metabolite levels in the future.

Quantitative metabolomics

Despite extensive knowledge of fundamental metabolic processes such as enzyme kinetics and substrate specificities, the actual output of large and branched biosynthetic networks is largely unpredictable to date. At first glance it might further be surprising, that even metabolome sizes are unknown for a particular organism. Not even for *Arabidopsis* or related fully sequenced and well-annotated organisms have metabolic reconstructions reached a level of comprehensiveness that would allow counting all possibly synthesized metabolites (http://ergo.integratedgenomics.com/ERGO-Plant/). Furthermore, using the reconstructions available, not even the presence of metabolites are explainable that already have been detected and structurally identified by metabolomic techniques such as citramalic acid in plants [41]. Apart from lack of enzyme specificities, a possible explanation could be lack of substrate availability under normal situations due to spatial compartmentation of biochemical pathways. Such compartmentation exists in all higher organisms, and additionally, many enzymes exist in several isoforms that are most regularly expressed in different compartments, cell types or tissues at different

time points. Spatial and temporal control over protein synthesis as well as enzyme isoforms would enable organisms to rapidly adapt to varying needs of metabolite synthesis. Considering the variability of protein networks, metabolome variations could be explainable even for related species by subtle alterations in the ability to form of protein complexes [37]. Protein complexes might work as supplementary tool to allow adjustments of metabolome complexity by mechanism similar to metabolite channelling. Even if channelling in its exact terms is not reality, the likelihood that a given substrate is available to a certain enzyme is greatly enhanced if all 'pathway enzymes' are connected in close neighbourhood via protein complexes.

Taking this lack of theoretical information about the nature and size of a given organism's metabolome into account, tremendous tasks are posed to the analytical methods for comprehensive and unbiased metabolite identification and quantitation. To date, no rigid comparison of extraction techniques has been published that highlights advantages and drawbacks comparing different classes of compounds. Mostly, simple alcohol or water/alcohol mixtures are used to extract multiple components from homogenized tissues [42,43], but systematic and rigorous validation is lacking. A general purpose method is also hard to imagine as no general characteristics account equally for all metabolites: they are different in size, number and nature of functional groups, volatility, charge states or electromobility, polarity and other physicochemical parameters. Moreover, each analytical detection method itself has a certain bias. For example, using mass spectrometry requires that metabolites are ionisable, coulometry needs analyte responses to varying redox potentials [44], ultraviolet absorption or fluorescence emission presumes that biochemical compounds bear moieties with excitable electrons (such as found in aromatic rings), and most other techniques are either two special (such as radioactivity detection), too insensitive (such as light scattering) or too difficult to be coupled to on-line separations (such as infrared spectroscopy). Therefore, no single metabolomic technique exists but a combination of aforementioned methods needs to be used. The largest scope with respect to universality, sensitivity and selectivity is clearly achieved using mass spectrometry (MS). Applying different ionization techniques has proven very appropriate to detect a large variety of metabolites. For example, simple terpenes, carotenoids, or aliphatics are hardly chargeable by electrospray ionization (ESI), the standard technique used in conjunction with liquid chromatography (LC). Such hydrocarbons, however, are often volatile and can therefore easily be detected by a combination of gas chromatography (GC) and MS, for example using classical electron impact ionization. With this regards, a combination of GC/MS and LC/MS methods is already capable to analyse a wide range of metabolites. However, especially for LC/MS effects of ion suppression due to matrix effects must not be ignored [45,46]. Ion suppression can only partly be circumvented by reducing the size of liquid droplets [47] and it might invalidate any approach to large-scale pathway elucidation or metabolic reconstruction that fails to properly preseparate metabolites prior to mass spectrometric detection. In order to down weight problems with ion suppression, LC/MS might be coupled to further detectors by splitting LC flows, e.g. to UV diode-array or electrochemical detectors, hence increasing the range of detectable metabolites. For the same reason, also use of capillary electrophoresis (CE) should be taken into consideration [48]. CE separates molecules with respect to their apparent charge radius, and it is therefore best applicable to analysis of easily ionisable or ionic compounds. In this sense, CE would be perfectly complementary to LC preseparation of metabolic mixtures prior to MS detection as the majority of existing LC methods are focusing on the analysis of lipophilic compounds that would be retained by reverse-phase LC columns. Only a few examples exist for analysis of highly polar compounds by LC/MS, such as by hydrophilic interaction chromatography [49]. In GC/MS, recent advances with respect to fast acquisitions as well as accurate mass determinations have been achieved by applying time-of-flight technology (TOF) [50]. This technology can be used to discriminate different genotypes based on the variance of individual compounds in contrast to metabolic fingerprinting technologies lacking the identification of individual metabolites (see Fig. 1) Additionally, powerful deconvolution algorithms [51] have been developed to find peaks without prior knowledge of their abundance, mass spectral characteristics, or retention time. Such deconvolution algorithms still do not work in LC/MS but would be urgently needed with respect of unbiased analysis of arbitrary complex mixtures.

All one dimensional detection methods lack resolution to separate the full suite of individual metabolites. However, such 1D approaches alone are in most cases sufficient to distinguish genotypes by 'metabolic fingerprinting' and multivariate data analysis. Metabolic fingerprinting does not aim to separate all individual components but instead has usually the focus of rapid screening and characterizing a

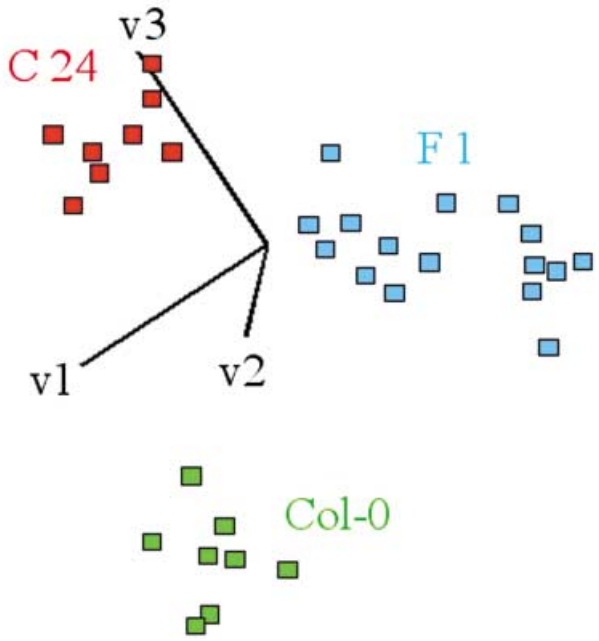

Fig. 1. Principal component analysis based on metabolomic profiling by GC/TOF. A clear discrimination between the *Arabidopsis* ecotypes C24, Col-0 and their F1 hybrid generation is observed.

variety of different samples according to predefined classes [52–57].

Typically, metabolic fingerprinting is carried out with little or no sample preparation. For example, nuclear magnetic resonance (NMR) could be used as metabolite detection technique. NMR is even more specific and universal than MS, but it lacks sensitivity. Moreover, most metabolites will result in several to many signals in NMR, causing a large problem in resolving individual metabolites if no chromatographic separation is coupled prior to NMR detection and identification. Therefore, only sparse information can be derived from NMR fingerprints with respect to individual metabolites. In a prominent example, silent yeast mutants were readily discriminated from wild type genotypes by NMR fingerprinting [58], but the data were needed to be complemented by classical enzyme assays to derive any biochemical meaningful conclusion. Similarly to NMR, mass spectrometry may be used for fingerprinting purposes. Direct-infusion mass spectrometry has shown to be ideally suited for high throughput classifications of sample origins [59]. In such applications, partial ion suppression in electrospray ionization may eventually be helpful as it can be regarded as amplifier of slight matrix differences, in addition to metabolites that are more abundant or unique in one of the sample populations under study.

According to the technical limits and capabilities, analytical methods of metabolites can roughly be grouped into four classes of metabolic characterizations [60] which would all have their own and special requirements for rigidity of method validation [61]. First, there is rapid genotype discrimination by 'metabolic fingerprinting' as given above. Next, there is classical target-driven hypothesis testing looking for one or a few single compounds. This may be called 'metabolite target analysis'. Sophisticated methods for extractions, sample preparation, sample clean-ups and internal references may be used in this area, rendering target analysis much more precise than other methods. Applications are found widely in literature of both biology and analytical chemistry, and it would be out of scope to try to cover them appropriately. Thirdly, one can aim at quantitation of a number of analytes belonging to a certain biochemical pathway or a set of chemically related compounds. This approach could be called 'metabolite profiling' although in pharmaceutical research, this term is also used for tracking down the catabolic fate of administered drugs. In metabolite profiling, *de novo* identification is mostly not a problem because all analytes have been predefined based on assumed biological relevance. Many applications can be found exemplifying this approach, including analysis of isoprenoids [62], phenylpropanoids [63], redox-active compounds [64,65], polar lipds [66], oxylipins [67], flavonoids [68,69], saponins [70], hydroxy and amino acids [71], carotenoids [72], or flavour ingredients like esters and aldehydes [73]. It is clear that sample preparation and quantification must undergo compromises if many different individual compounds are analysed simultaneously. For example, in metabolite profiling normally no isotopes are used and often, only crude sample fractionation and clean-up steps are carried out. Correspondingly, metabolite profiling is usually accompanied by rather focused biological questions and deep background knowledge, trying to cover all relevant compounds with authentic standards.

This level of initial understanding is not reached if researchers broaden their view to cover many or all metabolites simultaneously. In analogy to proteomics, this approach could be termed 'metabolomics'. It is more defining an aim than reality, as no truly metabolomic techniques been developed that were able to cover all metabolites irrespective of size and molecular nature (e.g. including cell walls, starch, etc.). However it can be clearly distinguished from metabolite profiling by the view which is underlying a certain experiment, for example if it is tried to take all detectable peaks into account, including unknowns. It may therefore be regarded as characteristic of metabolomic approaches [74] to cover as many individual metabolites as possible, disregarding if their exact chemical structures are known *a priori*. Metabolomics has the ultimate goal of unbiased identification and quantitation of all the metabolites present in a certain biological sample which was garnered from accurately defined experimental conditions. In this view, there are a few examples of metabolomic experiments. Different compound classes have been investigated using refined fractionation techniques of rice grain extracts prior to GC/MS analysis [75], and a similar approach was taken for GC/MS analyses of leaf extracts from plant mutants that combined all known and unknown peaks that were detectable from total ion chromatograms [76]. An interesting approach to the analysis of metabolites detectable in rat serums was carried out by Vigneau-Callahan *et al.* [77]. In this paper, the analytical variability in rat blood serums was carefully investigated to derive metabolites that were later assigned as targets in terms of metabolite profiling. It is interesting to note that from over 1200 detectable metabolites only 250 were regarded as significant to determine metabolic phenotypes, and it might be questionable if biological variability needs to be regarded as unwanted rather than as intrinsic property of a certain tissue.

There is another clear bottleneck for turning metabolomic data acquisition into biological significant information: metabolite identification. Disregarding post-translational modifications for the moment, the identification task in transcriptomics and proteomics is comparatively easy: the basic sequences of gene products can be linearly read from the 4-letter code of nucleotides or the 20-letter-code of amino acids, respectively. However in metabolomics, each metabolite needs to be identified *de novo* including the sequence and stereochemistry of all its atoms. Just like metabolite detection, there is no single technique fulfilling all requirements for full metabolite structural elucidation. With either NMR or MS, tentative structures can be suggested if reference compounds are available for direct comparison. However, only with both methods together or with a high level of prior background information, definite molecular structures can be proposed if no such standards are accessible. This process is tedious and cumbersome, although some level of automation has been achieved during the last 10 years. Unfortunately, many unknown metabolites can be expected from the theoretical considerations of metabolome complexity given above, and in fact, this is what is observed in reality. In a variety of metabolomic studies, each organ and each plant species were clearly different [76,78], indicating the important role of metabolism to define phenotypes. However, this situation

makes it difficult for the researcher to do deliver publishable results: there are too many unknown peaks to regard them unimportant, especially because in many cases unknowns are also revealed to be statistically different in the biological situations under study.

Therefore structure elucidation is an essential part of metabolomic approaches achievable by combinations of chromatographic metabolite separation, off-line or online NMR and MS (Fig. 2).

A faster road to structural elucidation could be the use of the unrivalled power of high-resolution Fourier-transform ion cyclotron mass spectrometry (FT-MS). In principle, FT-MS has the capability to resolve many coeluting compounds, fragment these by electron-capture dissociation [79] or other means, and then characterize both parent ion and fragments with inaccuracies < 0.001 Da. This accuracy would be enough to calculate elemental compositions, especially if the search space is limited using isotope ratios. Using large chemical and biochemical libraries such as Beilstein and CAS, a single best hit would result from the combined information, which could be confirmed in some cases by authentic standards or NMR. However, no such study has been presented so far. Most work in functional genomics using FT-MS has been carried out in the field of proteomics with the aid of capillary isoelectric focusing or nano-LC for preseparation of complex peptide mixtures (see above). However, duty cycles of FT-MS instruments are long, charge repulsion effects occur if too many ion species are simultaneously in the cyclotron chamber [80], and the instrument itself is far from being a routine benchtop mass spectrometer but requires high technical skills for maintenance. Some efforts in resolving complex mixtures have been performed by direct infusion. Several thousand nonaliphatic components have been detected in (crude) oil samples [81], but besides detection, no structural investigations were carried out. It is interesting to note that further advances have been performed to increase the resolution of FT-MS using ion mobility spectrometry [82]. For biology-directed research of complex metabolite mixtures, FT-MS has seldomly been used. Among the few examples was the characterization of lipooligosaccharides [83], discovery of central nervous system agents [84] or high throughput screening of combinatorial libraries [85]. Regarding the efforts to describe the inventory of natural products, for example in medicinal plants, FT-MS might turn out to be very helpful in characterization of novel compounds [86].

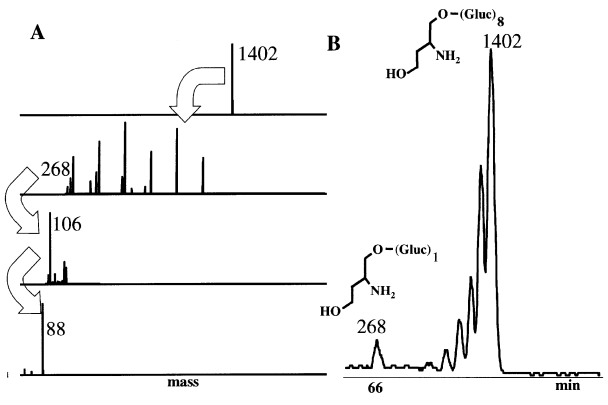

Fig. 2. Structure elucidation using an IONTRAP CID. Fragmentation on specified fragments. (A) A novel N-glycan structure with the mass 1402 was selected and fragmented. The resulting fragmentation pathway revealed a repetitive structure based on a glycoside-chain linked to an amino-hydroxy-aliphate. (B) All intermediates were found in HILIC chromatography [50].

Network computation

How can we estimate the structure of metabolism derived from measurements of cellular metabolite levels? In the past, only small networks with high internal fluxes have been studied in detail. For example glycolysis was investigated by isotopomer or flux balancing methods [87] [88]. However, such isotopomer methods can necessarily only work on very small pathways. Although very detailed knowledge is gained on the internal structure of such networks, general applicability may be limited by flux constraints and network topology. Moreover, the effects of isotope spiking will rapidly dilute through larger networks, and it is hard to imagine that the results of a variety of experiments on different small networks could be combined to get a generalized view of apparently unrelated paths of metabolism.

Among the four nonisotopic approaches towards metabolic analysis described above, metabolomics seems best applicable for the study of large metabolic networks, as it focuses on quantifying individual metabolites without having a bias concerning the metabolite's chemical nature.

We have recently reported on use of metabolite-metabolite correlation analysis [78] and generation and graphical visualization of large metabolic networks [89]. Such correlations can be understood by the theory of 'metabolic control analysis' [90] and its extension to coresponse analysis [91]. However, direct biological interpretation of the experimental observation of metabolic coregulation is restricted by some basic considerations. Obviously, any subcellular compartmentalization is lost in the process of sample preparation of nearly all experimental data sets published so far. Whereas protein expression can partly be denoted to plant compartments based on their target sequences, the actual compartmentation of most metabolites is quite unclear. Accordingly, more detailed studies revealed astonishing facts [92]. It is surprising that metabolite correlations can be seen despite this fact. It might be more understandable if defining metabolic roles for organs (and tissues) or part of organs rather than thinking too strictly in terms of compartmentation of biochemical pathways. It is clear from the data that deviations from steady state average levels are far more frequent than previously assumed, supporting the proper use of metabolic snapshots rather than comparing average levels [93]. The strength as well as the reasons for these deviations might well shade light on biochemical differences between, for example, mutants and wild type genotypes. Using pair-wise comparisons of all metabolite snapshot levels and combining these correlations to protein regulation, it should in principle be possible to reconstruct biochemical pathways. In a test case study, Arkin *et al.* have demonstrated the suitability of time series measurements of metabolite concentrations for deducing small networks [94]. This group

was also able to deduce reaction orders and rate constants for a part of glycolysis from kinetics of pulsed changes in metabolic reaction networks [95]. This idea of deducing chemical mechanism from the linear response around steady state levels was also tried by Diaz-Sierra *et al.* [96] who used a mass action law approach in order to calculate stoichiometries and rate constants from metabolic kinetics.

If all these local small networks were connectable, large metabolic networks would result. How could we compare and analyse such large networks? There is a large pile of experience in the field of mathematics and statistics on handling clustered and complex data sets, and some physicists now apply these experiences to biological data. A nice overview about the statistical mechanics of complex networks is given by Albert and Barabási [97] who have reviewed existing methods with a focus on the relation between robustness and network topology. Earlier, metabolic networks have been described as small-world networks [98,99], which have apparent topologies that render metabolism robust against random mutations, but vulnerable against attack on specific highly connected nodes [100,101]. Such analysis of inherent characteristics and topology can also be deduced from analysis of stoichiometric connections in biochemical nets that allow calculating the outcome of knock-out mutations on growth of *Escherichia coli* [102], or, in another example, the prediction of novel metabolic pathways [103]. A related analysis on the lethality of knock-out mutations has also been carried out investigating protein networks by yeast-two-hybrid screens consisting of 1870 nodes and 2240 edges [40]. Again, experimental findings and predictions from network calculations were found to be highly correlated. In a different approach, persistence in metabolic nets [104] has been investigated by Hurst's 'rescaled analysis' in order to analyse memory effects after local perturbations. The degree of network connectivity has also been related to the global consequences of such perturbations on other branches of the network [105], far away from the initial event. Although it is yet to early to predict which combinations of tools will best describe the dynamics, structure, flexibility and robustness of large metabolic, it seems timely to approach these questions by carrying out adequate experiments.

Conclusions

Integrating different levels of metabolomic and proteomic data to understand complex metabolic networks can be reached using existing analytical and mathematical methods [106] (Fig. 3). Such approach targets to a comprehensive analysis of biological systems, the so-called 'Systems Biology' approach. The paradigm of Systems Biology is challenging classical descriptive biology in many ways. First, data on the quantitative response of gene products have to be acquired. Although even the step of multiparallel and quantitative data acquisition is extremely difficult, it is not sufficient at all for understanding details of biological networks: the nature and strength of the interactions between them needs to be accurately described in mathematically and statistically sound ways. If enough biologically meaningful experiments have been carried out, data analysis of such network descriptions may eventually result in a multitude of differential equations, with which modelling and predictions could be possible. Demands on such approaches are so high that even large institutions or research networks will have problems pursuing the Systems Biology idea. Furthermore, relevant new information gathered by such an approach will always rely on comparisons to background knowledge which are backed by detailed studies of parts of the system that have been carried out in the classical step-by-step and protein-by-protein way. Therefore, network assessments will probably not be more appropriate than classical studies. Instead,

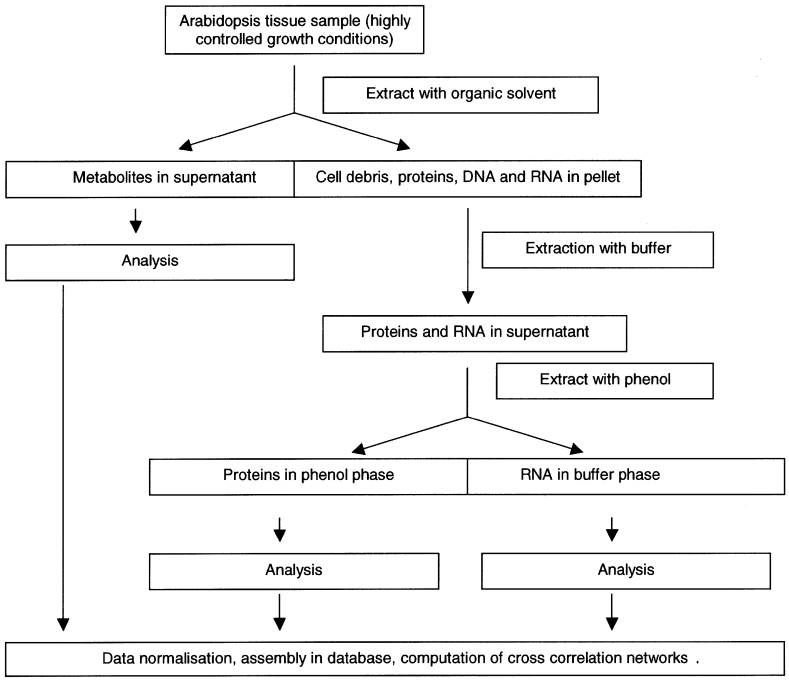

Fig. 3. Proposed flux-scheme to investigate cross-correlations of metabolites, proteins and mRNA by integrative extraction of metabolites, proteins and mRNA from one sample [106].

Systems Biology approaches may be complementary to detailed and very focused work, and it could turn out that both approaches rely on each other with respect to functional genomics.

References

1. The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815.
2. Matsumoto, T., Wu, J.Z., Baba, T., Katayose, Y., Yamamoto, K., Sakata, K., Yano, M. & Sasaki, T. (2001) Rice genomics: current status of genome sequencing. *Novart. Found. Symp.* **236**, 28–41.
3. ter Kuile, B.H. & Westerhoff, H.V. (2001) Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Lett.* **500**, 169–171.
4. Marcotte, E.M. (2001) The path not taken. *Nat. Biotechnol.* **19**, 626–627.
5. Gygi, S.P., Rochon, Y., Franza, B.R. & Aebersold, R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720–1730.
6. Smith, R.D., Pasa-Tolic, L., Lipton, M.S., Jensen, P.K., Anderson, G.A., Shen, Y.F., Conrads, T.P., Udseth, H.R., Harkewicz, R., Belov, M.E., Masselon, C. & Veenstra, T.D. (2001) Rapid quantitative measurements of proteomes by Fourier transform ion cyclotron resonance mass spectrometry. *Electrophoresis* **22**, 1652–1668.
7. Mann, M., Ong, S.E., Gronborg, M., Steen, H., Jensen, O.N. & Pandey, A. (2002) Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends Biotechnol.* **20**, 261–268.
8. Oda, Y., Huang, K., Cross, F.R., Cowburn, D. & Chait, B.T. (1999) Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl Acad. Sci. USA* **96**, 6591–6596.
9. Weckwerth, W., Willmitzer, L. & Fiehn, O. (2000) Comparative quantification and identification of phosphoproteins using stable isotope labeling and liquid chromatography/mass spectrometry. *Rapid Comm. Mass Spectrom.* **14**, 1677–1681.
10. Oda, Y., Nagasu, T. & Chait, B.T. (2001) Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.* **19**, 379–382.
11. Ficarro, S.B., McCleland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., Hunt, D.F. & White, F.M. (2002) Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **20**, 301–305.
12. Zhu, H., Klemic, J.F., Chang, S., Bertone, P., Casamayor, A., Klemic, K.G., Smith, D., Gerstein, M., Reed, M.A. & Snyder, M. (2000) Analysis of yeast protein kinases using protein chips. *Nat. Genet.* **26**, 283–289.
13. Klose, J., Nock, C., Herrmann, M., Stuhler, K., Marcus, K., Bluggel, M., Krause, E., Schalkwyk, L.C., Rastan, S., Brown, S.D.M., Bussow, K., Himmelbauer, H. & Lehrach, H. (2002) Genetic analysis of the mouse brain proteome. *Nat. Genet.* **30**, 385–393.
14. Nordhoff, E., Egelhofer, V., Giavalisco, P., Eickhoff, H., Horn, M., Przewieslik, T., Theiss, D., Schneider, U., Lehrach, H. & Gobom, J. (2001) Large-gel two-dimensional electrophoresis-matrix assisted laser desorption/ionization-time of flight-mass spectrometry: An analytical challenge for studying complex protein mixtures. *Electrophoresis* **22**, 2844–2855.
15. Peck, S.C., Nuhse, T.S., Hess, D., Iglesias, A., Meins, F. & Boller, T. (2001) Directed proteomics identifies a plant-specific protein rapidly phosphorylated in response to bacterial and fungal elicitors. *Plant Cell.* **13**, 1467–1475.
16. Gygi, S.P., Corthals, G.L., Zhang, Y., Rochon, Y. & Aebersold, R. (2000) Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl Acad. Sci. USA* **97**, 9390–9395.
17. Schubert, M., Petersson, U.A., Haas, B.J., Funk, C., Schroder, W.P. & Kieselbach, T. (2002) Proteome map of the chloroplast lumen of *Arabidopsis thaliana*. *J. Biol. Chem.* **277**, 8354–8365.
18. Prime, T.A., Sherrier, D.J., Mahon, P., Packman, L.C. & Dupree, P. (2000) A proteomic analysis of organelles from *Arabidopsis thaliana*. *Electrophoresis* **21**, 3488–3499.
19. Lilley, K.S., Razzaq, A. & Dupree, P. (2002) Two-dimensional gel electrophoresis: recent advances in sample preparation, detection and quantitation. *Curr. Opin. Chem. Biol.* **6**, 46–50.
20. Wall, D.B., Kachman, M.T., Gong, S.Y.S., Parus, S.J., Long, M.W. & Lubman, D.M. (2001) Isoelectric focusing nonporous silica reversed-phase high- performance liquid chromatography/electrospray ionization time- of-flight mass spectrometry: a three-dimensional liquid-phase protein separation method as applied to the human erythroleukemia cell-line. *Rapid Comm. Mass Spectrom.* **15**, 1649–1661.
21. Wall, D.B., Parus, S.J. & Lubman, D.M. (2002) Three-dimensional protein map according to pI, hydrophobicity and molecular mass. *J. Chromatogr. B.* **774**, 53–58.
22. Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H. & Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999.
23. Zhou, H.L., Watts, J.D. & Aebersold, R. (2001) A systematic approach to the analysis of protein phosphorylation. *Nat. Biotechnol.* **19**, 375–378.
24. Goodlett, D.R., Keller, A., Watts, J.D., Newitt, R., Yi, E.C., Purvine, S., Eng, J.K., von Haller, P., Aebersold, R. & Kolker, E. (2001) Differential stable isotope labeling of peptides for quantitation and de novo sequence derivation. *Rapid Comm. Mass Spectrom.* **15**, 1214–1221.
25. Arnott, D. (2000) Bridging the genome and the proteome: a mass spectrometric equivalent to the Western. blot. In *Proceedings of the 48th Meeting of the American Society for Mass Spectrometry*. www.asms.org, Long Beach, CA, USA.
26. Goshe, M.B., Conrads, T.P., Panisko, E.A., Angell, N.H., Veenstra, T.D. & Smith, R.D. (2001) Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. *Anal. Chem.* **73**, 2578–2586.
27. Washburn, M.P., Wolters, D. & Yates, J.R. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19**, 242–247.
28. Washburn, M.P., Ulaszek, R., Deciu, C., Schieltz, D.M. & Yates, J.R. (2002) Analysis of quantitative proteomic data generated via multidimensional protein identification technology. *Anal. Chem.* **74**, 1650–1657.
29. Li, L.J., Masselon, C.D., Anderson, G.A., Pasa-Tolic, L., Lee, S.W., Shen, Y.F., Zhao, R., Lipton, M.S., Conrads, T.P., Tolic, N. & Smith, R.D. (2001) High-throughput peptide identification from protein digests using data-dependent multiplexed tandem FTICR mass spectrometry coupled with capillary liquid chromatography. *Anal. Chem.* **73**, 3312–3322.
30. Shen, Y.F., Tolic, N., Zhao, R., Pasa-Tolic, L., Li, L.J., Berger, S.J., Harkewicz, R., Anderson, G.A., Belov, M.E. & Smith, R.D. (2001) High-throughput proteomics using high efficiency multiple- capillary liquid chromatography with on-line high-performance ESI FTICR mass spectrometry. *Anal. Chem.* **73**, 3011–3021.
31. Shen, Y.F., Zhao, R., Belov, M.E., Conrads, T.P., Anderson, G.A., Tang, K.Q., Pasa-Tolic, L., Veenstra, T.D., Lipton, M.S., Udseth, H.R. & Smith, R.D. (2001) Packed capillary

reversed-phase liquid chromatography with high-performance electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for proteomics. *Anal. Chem.* **73**, 1766–1775.
32. Jensen, P.K., Pasa-Tolic, L., Peden, K.K., Martinovic, S., Lipton, M.S., Anderson, G.A., Tolic, N., Wong, K.K. & Smith, R.D. (2000) Mass spectrometic detection for capillary isoelectric focusing separations of complex protein mixtures. *Electrophoresis* **21**, 1372–1380.
33. Smith, R.D., Pasa-Tolic, L., Lipton, M.S., Jensen, P.K., Anderson, G.A., Shen, Y.F., Conrads, T.P., Udseth, H.R., Harkewicz, R., Belov, M.E., Masselon, C. & Veenstra, T.D. (2001) Rapid quantitative measurements of proteomes by Fourier transform ion cyclotron resonance mass spectrometry. *Electrophoresis* **22**, 1652–1668.
34. Conrads, T.P., Alving, K., Veenstra, T.D., Belov, M.E., Anderson, G.A., Anderson, D.J., Lipton, M.S., Pasa-Tolic, L., Udseth, H.R., Chrisler, W.B., Thrall, B.D. & Smith, R.D. (2001) Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and N-15-metabolic labeling. *Anal. Chem.* **73**, 2132–2139.
35. Conrads, T.P., Anderson, G.A., Veenstra, T.D., Pasa-Tolic, L. & Smith, R.D. (2000) Utility of accurate mass tags for proteome-wide protein identification. *Anal. Chem.* **72**, 3349–3354.
36. Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M.J., Johnston, M., Fields, S. & Rothberg, J.M. (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623–627.
37. Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L.Y., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A.R., Sassi, H., Nielsen, P.A., Rasmussen, K.J., Andersen, J.R., Johansen, L.E., Hansen, L.H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B.D., Matthiesen, J., Hendrickson, R.C., Gleeson, F., Pawson, T., Moran, M.F., Durocher, D., Mann, M., Hogue, C.W.V., Figeys, D. & Tyers, M. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**, 180–183.
38. Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G. & Superti-Furga, G. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147.
39. Schwikowski, B., Uetz, P. & Fields, S. (2000) A network of protein–protein interactions in yeast. *Nat. Biotechnol.* **18**, 1257–1261.
40. Jeong, H., Mason, S.P., Barabasi, A.L. & Oltvai, Z.N. (2001) Lethality and centrality in protein networks. *Nature* **411**, 41–42.
41. Fiehn, O., Kopka, J., Trethewey, R.N. & Willmitzer, L. (2000) Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Anal. Chem.* **72**, 3573–3580.
42. Streeter, J.G. & Strimbu, C.E. (1998) Simultaneous extraction and derivatization of carbohydrates from green plant tissues for analysis by gas-liquid chromatography. *Anal. Biochem.* **259**, 253–257.
43. Johansen, H.N., Glitso, V. & Knudsen, K.E.B. (1996) Influence of extraction solvent and temperature on the quantitative determination of oligosaccharides from plant materials by high performance liquid chromatography. *J. Agric. Food Chem.* **44**, 1470–1474,.
44. Shi, H.L., Vigneau-Callahan, K.E., Matson, W.R. & Kristal, B.S. (2002) Attention to relative response across sequential electrodes improves quantitation of coulometric array. *Anal. Biochem.* **302**, 239–245.
45. Choi, B.K., Hercules, D.M. & Gusev, A.I. (2001) Effect of liquid chromatography separation of complex matrices on liquid chromatography-tandem mass spectrometry signal suppression. *J. Chromatogr. A* **907**, 337–342.
46. Sterner, J.L., Johnston, M.V., Nicol, G.R. & Ridge, D.P. (2000) Signal suppression in electrospray ionization Fourier transform mass spectrometry of multi-component samples. *J. Mass Spectrom.* **35**, 385–391.
47. Bahr, U., Pfenniger, A., Karas, M. & Stahl, B. (1997) High-sensitivity analysis of neutral underivatized oligosaccharides by nanoelectrospray mass spectrometry. *Anal. Chem.* **69**, 4530–4535.
48. Terabe, S., Markuszewski, M.J., Inoue, N., Otsuka, K. & Nishioka, T. (2001) Capillary electrophoretic techniques toward the metabolome analysis. *Pure Appl. Chem.* **73**, 1563–1572.
49. Tolstikov, V.V. & Fiehn, O. (2002) Analysis of highly polar compounds from plant origin: combination of hydrophilic interaction chromatography and electrospray ion trap mass spectrometry. *Anal. Biochem.* **301**, 298–307.
50. Veriotti, T. & Sacks, R. (2001) High-speed GC and GC/time-of-flight MS of lemon and lime oil samples. *Anal. Chem.* **73**, 4395–4402.
51. Stein, S.E. (1999) An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. *J. Am. Soc. Mass Spectrom.* **10**, 770–781.
52. Bundy, J.G., Spurgeon, D.J., Svendsen, C., Hankard, P.K., Osborn, D., Lindon, J.C. & Nicholson, J.K. (2002) Earthworm species of the genus Eisenia can be phenotypically differentiated by metabolic profiling. *FEBS Lett.* **521**, 115–120.
53. Griffin, J.L., Mann, C.J., Scott, J., Shoulders, C.C. & Nicholson, J.K. (2001) Choline containing metabolites during cell transfection: an insight into magnetic resonance spectroscopy detectable changes. *FEBS Lett.* **509**, 263–266.
54. Griffin, J.L., Nicholls, A.W., Keun, H.C., Mortishire-Smith, R.J., Nicholson, J.K. & Kuehn, T. (2002) Metabolic profiling of rodent biological fluids via H-1 NMR spectroscopy using a 1 mm microlitre probe. *Analyst* **127**, 582–584.
55. Gavaghan, C.L., Holmes, E., Lenz, E., Wilson, I.D. & Nicholson, J.K. (2000) An NMR-based metabonomic approach to investigate the biochemical consequences of genetic strain differences: application to the C57BL10J and Alpk: ApfCD mouse. *FEBS Lett.* **484**, 169–174.
56. Aranibar, N., Singh, B.K., Stockton, G.W. & Ott, K.H. (2001) Automated mode-of-action detection by metabolic profiling. *Biochem. Biophys. Res. Comm.* **286**, 150–155.
57. Griffin, J.L., Williams, H.J., Sang, E., Clarke, K., Rae, C. & Nicholson, J.K. (2001) Metabolic profiling of genetic disorders: a multitissue H-1 nuclear magnetic resonance spectroscopic and pattern recognition study into dystrophic tissue. *Anal. Biochem.* **293**, 16–21.
58. Raamsdonk, L.M., Teusink, B., Broadhurst, D., Zhang, N.S., Hayes, A., Walsh, M.C., Berden, J.A., Brindle, K.M., Kell, D.B., Rowland, J.J., Westerhoff, H.V., van Dam, K. & Oliver, S.G.

(2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat. Biotechnol.* **19**, 45–50.
59. Vaidyanathan, S., Rowland, J.J., Kell, D.B. & Goodacre, R. (2001) Discrimination of aerobic endospore-forming bacteria via electrospray-ionization mass spectrometry of whole cell suspensions. *Anal. Chem.* **73**, 4134–4144.
60. Fiehn, O. (2001) Combining genomics, metabolome analysis, and biochemical modeling to understand metabolic networks. *Comp. Funct. Genom.* **2**, 155–168.
61. Krull, I.S. & Swartz, M. (1999) Analytical method development and validation for the academic researcher. *Anal. Lett.* **32**, 1067–1080.
62. Lange, B.M., Ketchum, R.E.B. & Croteau, R.B. (2001) Isoprenoid biosynthesis. Metabolite profiling of peppermint oil gland secretory cells and application to herbicide target analysis. *Plant Phys.* **127**, 305–314.
63. Anterola, A.M., Jeon, J.H., Davin, L.B. & Lewis, N.G. (2002) Transcriptional control of monolignol biosynthesis in *Pinus taeda*–Factors affecting monolignol ratios and carbon allocation in phenylpropanoid metabolism. *J. Biol. Chem.* **277**, 18272–18280.
64. Shi, H.L., Vigneau-Callahan, K.E., Shestopalov, A.I., Millbury, P.E., Matson, W.R. & Kristal, B.S. (2002) Characterization of diet-dependent metabolic serotypes: Proof of principle in female and male rats. *J. Nutr.* **132**, 1031–1038.
65. Shi, H.L., Vigneau-Callahan, K.E., Shestopalov, A.I., Milbury, P.E., Matson, W.R. & Kristal, B.S. (2002) Characterization of diet-dependent metabolic serotypes: Primary validation of male and female serotypes in independent cohorts of rats. *J. Nutr.* **132**, 1039–1046.
66. Welti, R., Li, W., Sang, Y., Biesiada, H., Zhou, H.-E., Rajashekar, C.B., Williams, T.D. & Wang, X. (2002) Profiling membrane lipids in plant stress responses: role of phosholipase Dα in freezing-induced lipid changes in *Arabidopsis*. *J. Biol. Chem.* **35**, 31994–32002.
67. Weichert, H., Kohlmann, M., Wasternack, C. & Feussner, I. (2000) Metabolic profiling of oxylipins upon sorbitol treatment in barley leaves. *Biochem. Soc. Trans.* **28**, 861–862.
68. Pillai, B.V.S. & Swarup, S. (2002) Elucidation of the flavonoid catabolism pathway in *Pseudomonas putida* PML2 by comparative metabolic profiling. *Appl. Env. Microbiol.* **68**, 143–151.
69. Wang, S.Y., Kuo, Y.H., Chang, H.N., Kang, P.L., Tsay, H.S., Lin, K.F., Yang, N.S. & Shyur, L.F. (2002) Profiling and characterization antioxidant activities in *Anoectochilus formosanus*. *J. Agr. Food Chem.* **50**, 1859–1865.
70. Huhman, D.V. & Sumner, L.W. (2002) Metabolic profiling of saponins in *Medicago sativa* and *Medicago truncatula* using HPLC coupled to an electrospray ion-trap mass spectrometer. *Phytochem.* **59**, 347–360.
71. Gibson, K.M., Schor, D.S.M., Gupta, M., Guerand, W.S., Senephansiri, H., Burlingame, T.G., Bartels, H., Hogema, B.M., Bottiglieri, T., Froestl, W., Snead, O.C., Grompe, M. & Jakobs, C. (2002) Focal neurometabolic alterations in mice deficient for succinate semialdehyde dehydrogenase. *J. Neurochem.* **81**, 71–79.
72. Ferruzzi, M.G., Sander, L.C., Rock, C.L. & Schwartz, S.J. (1998) Carotenoid determination in biological microsamples using liquid chromatography with a coulometric electrochemical array detector. *Anal. Biochem.* **256**, 74–81.
73. Eglinton, J.M., Heinrich, A.J., Pollnitz, A.P., Langridge, P., Henschke, P.A. & Lopes, M.D. (2002) Decreasing acetic acid accumulation by a glycerol overproducing strain of *Saccharomyces cerevisiae* by deleting the ALD6 aldehyde dehydrogenase gene. *Yeast* **19**, 295–301.
74. Fiehn, O. (2002) Metabolomics – the link between genotypes and phenotypes. *Plant Mol. Biol.* **48**, 155–171.
75. Frenzel, T., Miller, A. & Engel, K.H. (2002) Metabolite profiling – a fractionation method for analysis of major and minor compounds in rice grains. *Cereal Chem.* **79**, 215–221.
76. Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R.N. & Willmitzer, L. (2000) Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* **18**, 1157–1161.
77. Vigneau-Callahan, K.E., Shestopalov, A.I., Milbury, P.E., Matson, W.R. & Kristal, B.S. (2001) Characterization of diet-dependent metabolic serotypes: Analytical and biological variability issues in rats. *J. Nutr.* **131**, 924S–932S.
78. Roessner, U., Luedemann, A., Brust, D., Fiehn, O., Linke, T., Willmitzer, L. & Fernie, A.R. (2001) Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *Plant Cell* **13**, 11–29.
79. Palmblad, M., Tsybin, Y.O., Ramstrom, M., Bergquist, J. & Hakansson, P. (2002) Liquid chromatography and electron-capture dissociation in Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Comm. Mass Spectrom.* **16**, 988–992.
80. Masselon, C., Tolmachev, A.V., Anderson, G.A., Harkewicz, R. & Smith, R.D. (2002) Mass measurement errors caused by 'local' frequency perturbations in FTICR mass spectrometry. *J. Am. Soc. Mass Spectrom.* **13**, 99–106.
81. Hughey. C.A., Rodgers, R.P. & Marshall, A.G. (2002) Resolution of 11000 compositionally distinct components in a single electrospray ionization fourier transform ion cyclotron resonance mass spectrum of crude oil. *Anal. Chem.* **74**, 4145–4149.
82. Collins, D.C. & Lee, M.L. (2002) Developments in ion mobility spectrometry-mass spectrometry. *Anal. Bioanal. Chem.* **372**, 66–73.
83. Leavell, M.D., Leary, J.A. & Yamasaki, R. (2002) Mass spectrometric strategy for the characterization of lipo-oligosaccharides from *Neisseria gonorrhoeae* 302 using FTICR. *J. Am. Soc. Mass Spectrom.* **13**, 571–576.
84. Prokai, L., Zharikova, A., Janaky, T., Li, X.X., Braddy, A.C., Perjesi, P., Matveeva, L., Powell, D.H. & Prokai-Tatrai, K. (2001) Integration of mass spectrometry into early-phase discovery and development of central nervous system agents. *J. Mass Spectrom.* **36**, 1211–1219.
85. Schmid, D.G., Grosche, P., Bandel, H. & Jung, G. (2000) FTICR-Mass spectrometry for high-resolution analysis in combinatorial chemistry. *Biotechnol. Bioeng.* **71**, 149–161.
86. Burja, A.M., Abou-Mansour, E., Banaigs, B., Payri, C., Burgess, J.G. & Wright, P.C. (2002) Culture of the marine cyanobacterium, *Lyngbya majuscula* (Oscillatoriaceae), for bioprocess intensified production of cyclic and linear lipopeptides. *J. Microbiol. Meth.* **48**, 207–219.
87. Szyperski, T. (1998) 13C-NMR, MS and metabolic flux balancing in biotechnology research. *Quat. Rev. Biophys.* **31**, 41–106.
88. Wiechert, W. & de Graaf, A.A. (1997) Bidirectional reaction steps in metabolic networks. I. Modeling and simulation of carbon isotope labelling experiments. *Biotechnol. Bioeng.* **55**, 112–117.
89. Kose, F., Weckwerth, W., Linke, T. & Fiehn, O. (2001) Visualising plant metabolomic correlation networks using clique-metabolite matrices. *Bioinformatics* **17**, 1198–1208.
90. de Vienne, D., Bost, B., Fievet, J., Zivy, M. & Dillmann, C. (2001) Genetic variability of proteome expression and metabolic control. *Plant Physiol. Biochem.* **39**, 271–283.
91. Hofmeyr, J.-H.S. (1995) Metabolic regulation: a control analytic perspective. *J. Bioenerget. Biomembr.* **27**, 479–490.
92. Farre, E.M., Tiessen, A., Roessner, U., Geigenberger, P., Trethewey, R.N. & Willmitzer, L. (2001) Analysis of the

compartmentation of glycolytic intermediates, nucleotides, sugars, organic acids, amino acids, and sugar alcohols in potato tubers using a nonaqueous fractionation method. *Plant Phys.* **127**, 685–700.
93. Kell, D.B. & Mendes, P. (2000) Snapshots of systems: Metabolic control analysis and biotechnology in the postgenomic era. In *Technical and medical implications of metabolic control analysis* (Cornish-Bowden, A.J. & Cárdenas, M.L., eds), pp. 3–25. Kluwer Academic Publishers, Dordrecht, the Netherlands.
94. Arkin, A., Shen, P. & Ross, J. (1997) A test case of correlation metric construction of a reaction pathway from measurements. *Science* **277**, 1275–1279.
95. Vance, W., Arkin, A. & Ross, J. (2002) Determination of causal connectivities of species in reaction networks. *Proc. Natl Acad. Sci. USA* **99**, 5816–5821.
96. Díaz-Sierra, R., Lozano, J.B. & Fairén, V. (1999) Deduction of chemical mechanisms from the linear response around steady state. *J. Phys. Chem. A* **103**, 337–343.
97. Albert, R. & Barabási, A.-L. (2002) Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97.
98. Almaas, E., Kulkarni, R.V. & Stroud, D. (2002) Characterizing the structure of small-world networks. *Phys. Rev. Lett.* **88**, article 098101. http://prl.aps.org/
99. Fell, D.A. & Wagner, A. (2000) The small world of metabolism. *Nat. Biotechnol.* **18**, 1121–1122.
100. Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N. & Barabasi, A.-L. (2000) The large-scale organization of metabolic networks. *Nature* **407**, 651–654.
101. Amaral, L.A.N., Scala, A., Barthelemy, M. & Stanley, H.E. (2000) Classes of small-world networks. *Proc. Natl Acad. Sci. USA* **7**, 11149–11152.
102. Edwards, J.S., Ibarra, R.U. & Palsson, B.O. (2001) *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* **19**, 125–130.
103. Schuster, S., Fell, D.A. & Dandekar, T. (2000) A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat. Biotechnol.* **18**, 326–332.
104. De la Fuente, I.M., Benitez, N., Santamaria, A., Aguirregabiria, J.M. & Veguillas, J. (1999) Persistence in metabolic nets. *Bull. Math. Biol.* **61**, 573–595.
105. Watts, D.J. (2002) A simple model of global cascades on random networks. *Proc. Natl Acad. Sci. USA* **99**, 5766–5771.
106. Weckwerth, W. & Fiehn, O. (2002) PCT patent application. European Patent Office (submitted).