Anal. Chem. **2006,** *78,* 7954−7958

Perspectives

Metabolomics: The Greatest Omics of Them All?

Danielle Ryan* and Kevin Robards

School of Science and Technology, Charles Sturt University, Locked Bag 588, Wagga Wagga, New South Wales, Australia 2678

Is metabolomics the greatest "omics" of all? Certainly, it has been suggested that metabolomics may in fact provide the most "functional" information of the omics technologies.[1] This reflects the limitations associated with transcriptomics and proteomics; for example, changes in the transcriptome and proteome do not always result in altered biochemical phenotypes (the metabolome).[1,2] Furthermore, the metabolome represents the final "omic" level in a biological system, and metabolites represent functional entities, unlike messenger RNA molecules, which constitute the transcriptome.[3] Metabolites thus have a clear function in the life of the biological system and are also contextual,[3] reflecting the surrounding environment. The metabolome can thus be thought of as a looking glass, which if looked through can show information concerning the physiological, developmental, and pathological status of a biological system.

Terminology relating to metabolomics has been controversial.[4] The term "metabolome" was first used by Olivier et al. in 1998[5] to describe the set of metabolites synthesized by an organism, in a fashion analogous to that of the genome and proteome. This definition has been limited[6] to "the quantitative complement of all of the low molecular weight molecules present in cells in a particular physiological or developmental state". Metabolomics was coined by Fiehn[7] and defined as a comprehensive analysis in which all metabolites of a biological system were identified and quantified.

Much of the confusion related to terminology arises from the similar "metabonomics", since the two terms are often used interchangeably. Furthermore, metabonomics has been described as a subset of metabolomics,[8] yet metabolomics can also be thought of as a subset of systems covered by metabonomics[3] cited in ref 9. The two fields employ similar methodologies and have a common goal of analyzing the metabolome. Hence, technologies will often be translatable across both fields. However, both precedence and the bulk of the literature support use of the term metabolomics to describe a comprehensive analytical approach that is *nonselective* and universally applicable to *identify* and *quantify all* metabolites of a biological system without size restriction on the analyte.

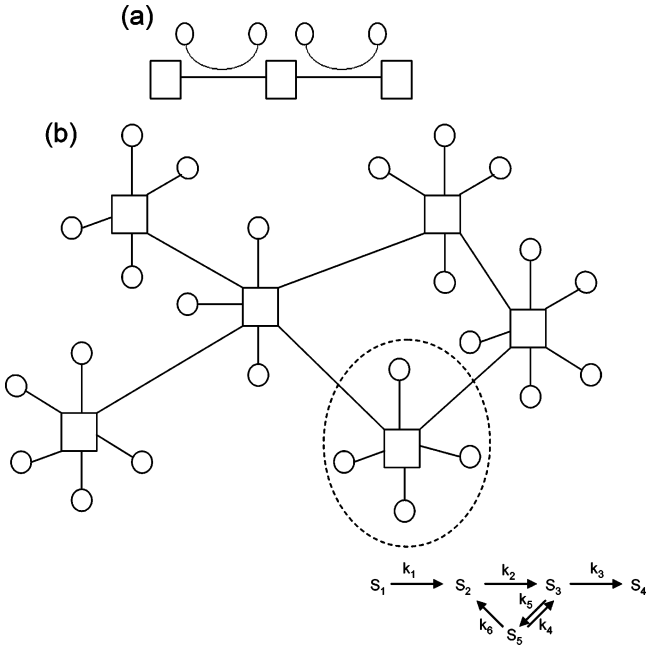

Figure 1. General schematic showing the traditional linear view of a metabolic pathway (a) and a metabolic network model (b) with an expanded view of a section of the network shown in linear view.

Metabolomics: Is It or Isn't It? Analysis of metabolites is not a new field; but historically it has been limited to relatively small numbers of target analytes as in studies of glycolysis or respiration. However, the realization that metabolic pathways do not act in isolation but rather as part of an extensive network has led to the need for a more holistic approach to metabolite analysis. This is illustrated in Figure 1, which compares the traditional linear view of a metabolic pathway with the holistic network model. Metabolite analysis is complicated by the number of analytes, their diversity, and dynamic ranges. It is estimated that the metabolome extends over 7−9 magnitudes of concentration (pmol−mmol),[10] and the number of metabolites in the plant system alone is

* Corresponding author. E-mail: dryan@csu.edu.au. Tel: + 61-2-69334382 Fax: +61-2-69332737.

(1) Sumner, L. W.; Mendes, P.; Dixon, R. A. *Phytochemistry* **2003**, *62*, 817−36.
(2) Fiehn, O.; Kopka, J.; Dörmann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L. *Nat. Biotechnol.* **2000**, *18*, 1157−61.
(3) Raamsdonk, L. M.; Teusink, B.; Broadhurst, D.; Zhang, N.; 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. *Nat. Biotechnol.* **2001**, *19*, 45−50.
(4) Villas-Bôas, S. G.; Rasmussen, S.; Lane, G. A. *Trends Biotechnol.* **2005**, *23*, 385−6.
(5) Oliver, S. G.; Winson, M. K.; Kell, D. B.; Banganz, F. *Trends Biotechnol.* **1998**, *16*, 373−8.
(6) Goodacre, R. *J. Exp. Biol.* **2005**, *56*, 245−54.
(7) Fiehn, O. *Comp. Funct. Genom.* **2001**, *2*, 155−68.
(8) Fiehn, O. *Plant Mol. Biol.* **2002**, *48*, 155−71.
(9) Lindon, J. C.; Holmes, E.; Nicholson, J. K. *Anal. Chem.* **2003**, *75*, 384A−91A.
(10) Dunn, W. B.; Ellis, D. I. *Trends Anal. Chem.* **2005**, *24*, 285−93.

10.1021/ac0614341 CCC: $33.50
Published on Web 10/31/2006

estimated to exceed 200 000.[11] Of more relevance and significance is the number and diversity of analytes likely to be encountered in a single system. We are in unknown territory here, but if we consider a system comprising x metabolites characterized by the substrate concentrations, S_1, S_2, ..., S_x and y fluxes characterized by the rate constants, k_1, k_2, ..., k_y involved in z pathways, where x may be typically 2000−5000 and z may be 20−50, we have some idea of the enormity of the task. This is illustrated in Figure 1 for a simple system involving five substrates, six fluxes, and a single pathway. A complete characterization of this system involves measurement of five concentrations and six rate constants, not all of which can be determined at any one time. The flavonoids illustrate the analytical problems associated with diversity. These commonly occur in glycosidic linkage with 5−10 sugars being commonly involved. The number of known flavonoids now exceeds 5000, leading to a possible 50 000 unique metabolites. For example, more than 200 different aglycons conjugated to glucose have been identified in grapes of *Vitis vinifera*.[12] Thus, a successful analytical approach in metabolomics must be capable of accurately measuring numerous known and unknown compounds that span a diverse chemical spectrum and a large dynamic concentration range. Sample preparation procedures and methods of quantification cannot yet meet these demands.

Historical approaches to metabolite analysis include metabolite profiling, metabolite fingerprinting, and target analysis. Metabolite fingerprinting aims to rapidly classify numerous samples using multivariate statistics, typically *without differentiation of individual metabolites* or their *quantitation*. Target analysis is constrained exclusively to the qualitative and quantitative analysis of a *particular metabolite* or metabolites. As a result, only a very small fraction of the metabolome is focused upon, signals from all other components being ignored.[13] Metabolite profiling involves the *identification* and *quantitation* by a particular analytical procedure of a *predefined set* of metabolites of known or unknown identity and belonging to a *selected metabolic pathway*.[7,10] By their nature, these approaches provide a restrictive noncomprehensive view of the metabolome. Nevertheless, metabolite profiling represents the oldest and most established approach and can be considered the precursor for metabolomics.

The analysis of biophenols in olives using high-performance liquid chromatography with UV detection[14] is an example of profiling that involves a predefined set of metabolites from a single metabolic pathway. It illustrates the notion of dependence of metabolite profiles on detection method in that different profiles are generated at different detection wavelengths. The use of universal detection such as evaporative light scattering or refractive index or mass spectrometry eliminates some but not all of this dependence. Further method dependency is related to the column in both gas chromatography (GC) and liquid chromatography (LC) in that it acts as a "filter" and not all components injected are necessarily eluted. However, it is sample procurement and sample preparation[15] that present the most significant restriction in relation to the need for universality and nonselectivity in metabolomics, and yet, sample handling has received much less attention in the literature than the more glamorous steps of the analysis that involve application of sophisticated instrumentation. Many studies have concluded that biological variation is much higher than that associated with the analytical technique,[16] emphasising the need for large numbers of replicates and standard sampling methods. This applies even to samples at seemingly identical developmental stage grown under highly controlled environmental conditions.

Table 1. Comparison of Selected Papers Against Criteria for Metabolomics

non-selective	universally applicable	identify	quantify	all metabolites	ref
no	no	no	no	no	21
yes	no	no	no	no	38
no	no	yes, selected metabolites	yes (as peak areas only)	no	39
yes	no	no	no	5036 metabolites	40
no	no	yes (limited)	no	no	41
no	no	yes (limited)	no	1227 metabolites	42

Composite metabolite profiling in which separate technology platforms are used to assess different sets of metabolites belonging to different metabolic pathways is one possible approach to metabolomics. Are the individual profiles additive and does the composite profile equate to the metabolome? Indeed, the need to detect as many metabolites as possible in a *single* analysis has been noted.[17] The major limitation of composite profiling is largely temporal in that simultaneous high-throughput analyses of large numbers of metabolites are not generally feasible. As the metabolome is dynamic, this time factor may introduce artifactual problems. The very process of metabolite extraction must disturb the *status quo* of an organism, and chemical changes brought about by exposure to oxygen, solvents, and change of pH are particularly common. Thus, the metabolites isolated from natural sources are not necessarily the metabolites that are present in the living tissue. Moreover, the level of metabolites at any given time represents a composite of both catabolic and anabolic processes and presents a snapshot of the metabolome at that particular time only. Of course, this can be an advantage as time-dependent snapshot sampling will reveal directed correlations of metabolic processes.[18]

Many authors claim to be undertaking metabolomics research, but are these claims sustainable? Using the accepted definition of metabolomics as our starting point, Table 1 compares some recent papers against the criteria required for metabolomics. Papers were selected solely for their use of the term metabolomics in the title, and their inclusion is not intended as a criticism but rather as an acknowledgment of the complexity of the problem. The inability to achieve metabolomics is apparent, and the term metabolome analysis is often used as an alternative. Simple plants

(11) Hartman, T.; Kutchan, T. M.; Strack, D. *Phytochemistry* **2005**, *66*, 1198−9.
(12) Sefton, M. A.; Francis, I. L.; Williams, P. J. *Am. J. Enol. Vitic.* **1993**, *44*, 359−70.
(13) Halket, J. M.; Waaterman, D.; Przyborowska, A. M.; Patel, R. K. P.; Fraser, P. D.; Bramley, P. M. *J. Exp. Bot.* **2005**, *56*, 219−43.
(14) Ryan, D.; Robards, K.; Lavee, S. *Int. J. Food Sci. Technol.* **1999**, *34*, 265−74.
(15) Gullberg, J.; Jonsson, P.; Nordstrom, A.; Sjostrom, M.; Moritz, T. *Anal. Biochem.* **2004**, *331*, 283−95.
(16) Roessner, U.; Wagner, C.; Kopka, J.; Trethewey, R. N.; Willmitzer, L. *Plant J.* **2000**, *23*, 131−42.
(17) von Roepenack-Lahaye, E.; Degenkolb, T.; Zerjeski, M.; Franz, M.; Roth, U.; Wessjohann, L.; Schmidt, J.; Scheel, D.; Clemens, S. *Plant Physiol.* **2004**, *134*, 548−59.
(18) Weckwerth, W. *Ann. Rev. Plant Biol.* **2003**, *54*, 669−89.

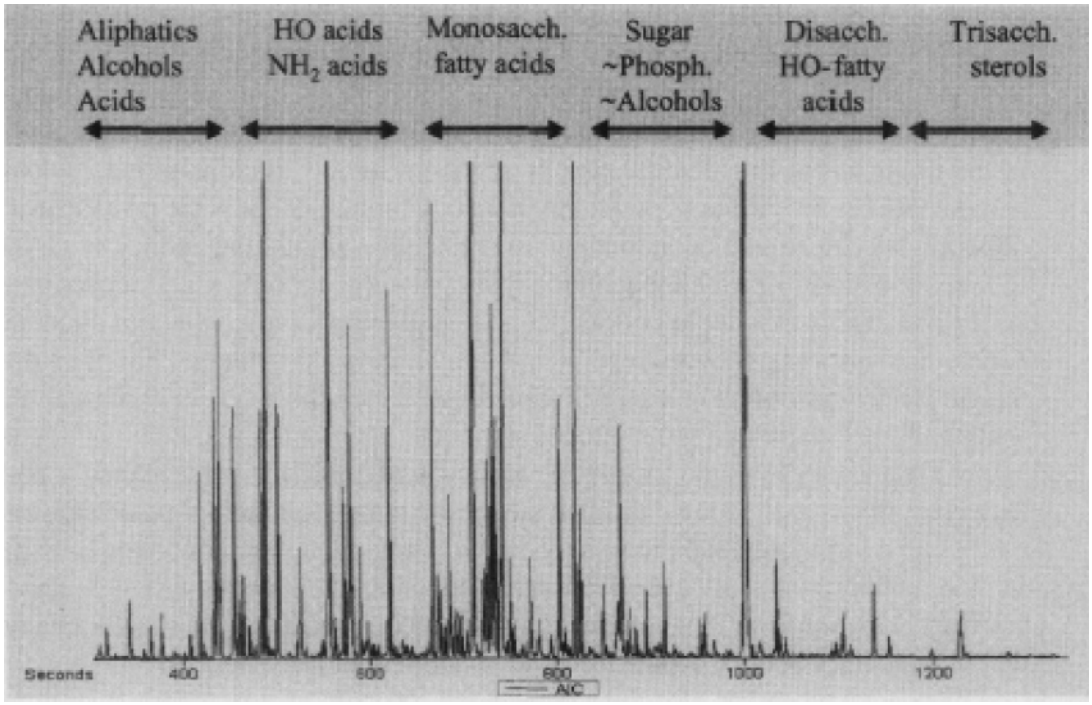

Figure 2. GC/TOF MS analysis of trimethylsilylated derivatives of hydrophilic and lipophilic metabolites extracted from a single *Arabidopsis* leaf sample. Reproduced with permission from Weckwerth et al.[23]

and yeasts with defined genomes have been widely used in metabolome analyses. These simple systems avoid the complications in analyte extraction that arise with increased plant complexity as, for example, with fruits, oily plants, and seeded plants with distinct compartments where analyte compartmentalization becomes a consideration. *Arabidopsis thaliana* is an example of a model plant species[19,20] chosen for its simplicity and for which the complete genome sequence and functional gene annotations are publicly available.

As an example of a typical qualitative analysis using a simple system, extracts of *A. thaliana* were separated on C18 monolithic silica capillary columns coupled to ion trap mass spectrometry.[21] Mass spectral deconvolution software led to the detection of over 700 base peaks. Even if many of the low-intensity base peaks were attributable to reagent impurities or contamination, the number of real metabolites still exceeded several hundred. The authors claimed good coverage of the plant lipophilic and secondary metabolome although carotenes were not detectable by ESI and required UV−visible detection. In another study, over 300 polar metabolites were identified in *A. thaliana* leaf by GC/MS^2 in a single extract. However, the discrepancy between the number of genes and the number of known reactions catalyzed by encoded enzymes has inevitably led to the conclusion that there remain many as yet unidentified metabolites.[22] Mass spectral signals could not be deconvoluted sufficiently to estimate the actual number of metabolites detected in *Arabidopsis*.[22]

Weckwerth et al.[23] developed a novel extraction protocol whereby hydrophilic and lipophilic metabolites, proteins and RNA were sequentially extracted from the same *A. thaliana* sample (Table 2). Hydrophilic and lipophilic metabolites were derivatized and analyzed by gas chromatography/time-of-flight mass spectrometry (GC/TOF MS) (Figure 2). The authors detected 652 metabolites and 297 proteins and clear RNA bands using this procedure. Detection of this number and diversity of metabolites using a single-platform technology is exceptional. What is more interesting from our perspective is what was not detected, namely, cell debris plus insoluble material from each partitioning step in the procedure. We can only guess at the number of undetected metabolites in this rejected material.

Why Persist? Given that metabolomics is not achievable with the current state of development in analytical science, why persist? There are many reasons apart from the simple desire to enhance our knowledge base, and the hype surrounding metabolomics stems from its incredible *potential* to contribute in diverse areas. It is this potential that continues to drive research in the field. Decoding the metabolome will impact on a multitude of areas, a few of which are discussed below.

The chemical diversity of the metabolome offers major opportunities for the discovery of novel drugs and bioactive molecules. This will have a significant impact on human health through the development of new medicines and "functional foods". The latter is directly related to the emerging field of nutrigenomics, which studies how specific genes and bioactive food components interact.[24] Nutrigenomics has the potential to enable individualized diets by prescribing particular functional foods. Given the fact that 150 million Americans spend over $20.5 billion

(19) Britz-McKibbin, P.; Terabe, S. *J. Chromatogr., A* **2003**, *1000*, 917−34.
(20) Taylor, J.; King, R. D.; Altmann, T.; Fiehn, O. *Bioinformatics* **2002**, *18*, S241−8.
(21) Tolstikov, V. V.; Lommen, A.; Nakanishi, K.; Tanaka, N.; Fiehn, O. *Anal. Chem.* **2003**, *75*, 6737−40.
(22) von Roepenack-Lahaye, E.; Degenkolb, T.; Zerjeski, M.; Franz, M.; Roth, U.; Wessjohann, L.; Schmidt, J.; Scheel, D.; Clemens, S. *Plant Physiol.* **2004**, *134*, 548−59.
(23) Weckwerth, W.; Wenzel, K.; Fiehn, O. *Proteomics* **2004**, *4*, 78−83.
(24) Trujillo, E.; Davis, C.; Milner J. *J. Am. Diet. Assoc.* **2006**, *106*, 403−13.

Table 2. References Probing the Metabolome of *Arabidopsis thaliana*

metabolite	extraction	analytical platform	outcome	ref
hydrophilic metabolites: sugars, amino and organic acids	extracted with organic solvent; partitioned into methanol/water	GC/TOF MS of TMS derivatives	652 metabolites	23
lipophilic metabolites: lipids, chlorophyll and waxes	and chloroform			
proteins	residue of initial extract partitioned into buffer	LC/MS	297 proteins and RNA bands	
RNA	and phenol	Northern blot analysis		
amino acids, sugars, organic acids		GC/MS of derivatives	326 metabolites	2
lipophilic and secondary metabolites	extracted with methanol	LC/MS	172 metabolites	21
secondary metabolites such as indoles, nitriles, flavonoids plus amino acids	extracted with aqueous methanol	LC/MS	1415 mass signals in leaf extracts and 827 in root extracts	17
sugars, amino acids, organic acids, sterols, phenols	various	GC/MS of derivatives	extraction efficiency for 66 metabolites	15

on functional foods, nutraceuticals, and dietary supplements,[25] the global impact of nutrigenomics is considerable. Further recognition to the importance of metabolomics in human health is evidenced by the establishment of two metabolomic themes in 2004 by the National Institute of Health (NIH, USA).[26]

Metabolomics represents a way to dissect and modify plant metabolism, physiology, and development.[27] Such capabilities are essential for breeding more robust plant varieties that are disease resistant. Metabolomics will enable the disease status of plants to be better defined through the identification of disease biomarkers and the role of environmental conditions and stresses on plant productivity to be ascertained. The economic ramifications associated with these tasks cannot be underestimated.

Food quality is directly dependent on metabolome composition. For example, poor-quality wheat cannot intrinsically produce high-quality bread. Food authenticity is also related to food quality, particularly with respect to adulteration, for example, when expensive vegetable oils are adulterated with cheap oil. This process represents a major economic fraud, which can have severe health implications depending upon the nature of the adulterant. Metabolomics facilitates product fingerprinting, which is critical for the detection and prevention of adulteration.

Functional genomics, as the name implies, aims to decipher gene function by establishing a better understanding of the correlation between genes and the functional phenotype of an organism.[28] Since the metabolome of a system represents the amplification and integration of signals from other functional genomic levels (e.g., transcriptome and proteome),[29] metabolomics can be considered tool for functional genomics. Functional genomics represents a way to do "smarter" genomics, rather than simply gene mapping and sequencing, and motivation for this research endeavor arises because of the large proportion of open reading frames (typically 20−40%[30]) in a fully sequenced organism that have no known function at the biochemical and phenotype levels. Such genes are referred to as "silent" or "orphan" genes. In the case of *Saccharomyces cerevisiae,* for example, around 6000 protein encoding genes exist; however, there are less than 600 low molecular weight intermediate metabolites (cited in ref 3).

Determining gene function can be achieved through metabolite profiling of specific genetically altered organisms. These metabolite profiles may then be compared to that of a "control" organism to yield information about the metabolic consequence of the altered genome[31] and ultimately assign gene function. This approach was first used by Roessner et al.,[16] whereby several potato tuber transgenic lines, modified in either sugar or starch metabolism, were compared. Based upon their observed results, the authors deemed metabolite profiling to be an important tool for characterizing the metabolic status of a plant with respect to environmental, developmental, or genetic factors. This is in accordance with Fiehn, who theorized that it should be possible to ultimately link metabolomic changes to the underlying genetic alterations.[8] Metabolomics thus has the potential to enable the phenotype of silent mutations to be mapped[29] and gene function to be assigned to the many poorly or unanoted genes.[31]

Systems biology uses an approach similar to that of functional genomics, but has significantly greater aims than the latter. Systems biology represents the ultimate challenge in that is aims to integrate genomics, transcriptomics, proteomics, and metabolomics[32] for a global understanding of biological systems. In essence, systems biology looks at the big picture to obtain a better understanding of how individual pathways or metabolic networks are related. Systems biology does not investigate individual genes, proteins, or metabolites one at a time, but rather investigates the behavior and relationships of all the elements in a particular biological system while it is functioning.[33] The general systems biology approach is a perturbation of the system (biologically, genetically, or chemically), followed by monitoring the impact of the perturbation at the genomic, proteomic, and metabolomic levels. These omic data can then be integrated and ultimately modeled computationally for a complete understanding of system functioning. The potential impact of systems biology is enormous, ranging from metabolite engineering[1] to reshaping medicine toward predictive, preventative, and personalized prevention of cellular dysfunction and disease.[34] What must be realized is that

(25) Burdock, G. A.; Carabin, I. G.; Griffiths, J. C. *Toxicology* **2006**, *221*, 17−27.
(26) Zerhouni, E. *Science* **2003**, *302*, 63−4, 72.
(27) King, J. *Trends Plant Sci.* **2002**, *7*, 278−80.
(28) Bino, R. J.; Hall, R. D.; Fiehn, O.; Kopka, J.; Saito, K.; Draper, J.; Nikolau, B. J.; Mendes, P.; Roessner-Tunali, U.; Beale, M. H.; Trethewey, R. N.; Lange, B. M.; Wurtle, E. S.; Sumner, L. W. *Trends Plant Sci.* **2004**, *9*, 418−25.
(29) Nielson, J.; Oliver, S. *Trends Biotecnol.* **2005**, *23*, 544−6.
(30) Goodacre, R. *Metabolomics* **2005**, *1*, 1−2.
(31) Trethewey, R. N. *Curr. Opin. Biotechnol.* **2001**, *12*, 135−8.
(32) Rochfort, S. *J. Nat. Prod.* **2005**, *68*, 1813−20.
(33) Ideker, T.; Galitski, T.; Hood, L. *Annu. Rev. Genomics Hum. Genet.* **2001**, *2*, 343−72.
(34) Weston, A. D.; Hood, L. *J. Proteome Res.* **2004**, *3*, 179−96.
(35) Kell, D. B. *Curr. Opin. Microbiol.* **2004**, *7*, 296−307.
(36) Ratcliffe, R. G.; Shachar-Hill, Y. *Annu. Rev. Plant Biol.* **2001**, *52*, 499−526.
(37) Fernie, A. R.; Roscher, A.; Ratcliffe, R. G.; Kruger, N. J. *Physiol. Plant.* **2002**, *114*, 172−81.
(38) Kikuchi, J.; Shinozaki, K.; Hirayama, T. *Plant Cell Physiol.* **2004**, *45*, 1099−104.

metabolomics is the core of systems biology, and through systems biology, an iterative improvement in our understanding of the metabolome can be achieved.[35]

The Future. Major advances and a totally new approach to analysis are essential before the holy grail of metabolomics is attainable. This will include developments in data preparation, storage, and curation coupled with a greater understanding and application of bioinformatics. The general paradigm of method development in analytical science is improved sensitivity and selectivity/specificity in recognition of current instrumental limitations. However, success in metabolomics requires that we reverse this approach and employ nonselective universal procedures where specificity is achieved temporally rather than spatially by extension of existing multidimensional platforms. In vivo analysis of intact samples is an attractive proposition, and in this context, NMR-based technologies have a distinct advantage over other analytical techniques.[36,37] However, as noted by Weckwerth,[18] "Whether we model the reality or only a shadow of it, we form a better understanding of the intricate biochemical processes and their scattering in living systems."

(39) Jonsson, P.; Johansson, A. I.; Gullberg, J.; Trygg, J.; Jiye, A.; Grung, B.; Marklund, S.; Sjöström, M.; Antti, H.; Moritz, T. *Anal. Chem.* **2005**, *77*, 5635−42.
(40) Shen, Y.; Zhang, R.; Moore, R. J.; Kim, J.; Metz, T. O.; Hixson, K. K.; Zhao, R.; Livesay, E. A.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **2005**, *77*, 3090−100.
(41) Shellie, R. A.; Welthagan, W.; Zrostliková, J.; Spranger, J.; Ristow, M.; Fiehn, O.; Zimmermann, R. *J. Chromatogr., A* **2005**, *1086*, 83−90.
(42) Welthagen, W.; Shellie, R. A.; Spranger, J.; Ristow, M.; Zimmermann, R.; Fiehn, O. *Metabolomics* **2005**, *1*, 65−73.

Received for review August 2, 2006. Accepted September 22, 2006.

AC0614341