Max-Planck-Institute of Molecular Plant Physiology
Department I, Root Metabolism

Signal-Metabolome Interactions in Plants

Dissertation
Submitted for graduation as
"Doctor rerum naturalium"
(Dr. rer. nat.)
in "Analytical Biochemistry"

A thesis submitted to the
Mathematisch-Naturwissenschaftlichen Fakultät
of the University of Potsdam

by

Claudia Sabine Birkemeyer

Potsdam, 5th November 2005

The work presented in this thesis was carried out between January 2001 and December 2004 at the Max-Planck-Institute of Molecular Plant Physiology in Potsdam, Germany.

Examiner 1: Prof. Dr. Lothar Willmitzer
Max-Planck-Institute of Molecular Plant Physiology, Potsdam/ University of Potsdam, Germany

Examiner 2: Prof. Dr. Elmar Weiler
Institute of Plant Physiology, Ruhr-University of Bochum, Germany

Examiner 3: Prof. Dr. John Pickett
Rothamsted Research, Biochemistry Division, Harpenden, Hertfordshire, United Kingdom

Examiner 4: Prof. Dr. Ivo Feussner
Department of Plant Biochemistry, Georg-August-University of Göttingen, Germany

Statutory Declaration

This Ph. D. thesis is the account of practical work completed between January 2001 and December 2004 in the department of Prof. Willmitzer at the Max-Planck-Institute of Molecular Plant Physiology, Potsdam, Germany. I affirm that the presented document is the result of my own work, without involving illegitimate references, and has not been submitted for any degree at any other university.

Eidesstattliche Erklärung

Diese Dissertation ist das Ergebnis der von Januar 2001 bis Dezember 2004 am Max-Planck-Institut für Molekulare Pflanzenphysiologie in Potsdam, Deutschland, durchgeführten praktischen Arbeiten. Ich versichere, die vorliegende Arbeit selbständig und ohne unzulässige Hilfe angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben. Ich versichere weiter, daß die Arbeit bei keiner anderen Hochschule als der Universität Potsdam eingereicht wurde.

Potsdam, den 05. November 2005

(Claudia Birkemeyer)

Abstract

From its first use in the field of biochemistry, instrumental analysis offered a variety of invaluable tools for the comprehensive description of biological systems. Multi-selective methods that aim to cover as many endogenous compounds as possible in biological samples use different analytical platforms and include methods like gene expression profile and metabolite profile analysis.

The enormous amount of data generated in application of profiling methods needs to be evaluated in a manner appropriate to the question under investigation. The new field of system biology rises to the challenge to develop strategies for collecting, processing, interpreting, and archiving this vast amount of data; to make those data available in form of databases, tools, models, and networks to the scientific community.

On the background of this development a multi-selective method for the determination of phytohormones was developed and optimised, complementing the profile analyses which are already in use (Chapter I). The general feasibility of a simultaneous analysis of plant metabolites and phytohormones in one sample set-up was tested by studies on the analytical robustness of the metabolite profiling protocol. The recovery of plant metabolites proved to be satisfactory robust against variations in the extraction protocol by using common extraction procedures of phytohormones; thus, a joint extraction of metabolites and hormones from plant tissue appears practicable (Chapter II).

Quantification of compounds within the context of profiling methods requires particular scrutiny (Chapter II). In Chapter III, the potential of stable-isotope *in vivo* labelling as normalisation strategy for profiling data acquired with mass spectrometry is discussed. First promising results were obtained for a reproducible quantification by stable-isotope *in vivo* labelling, which was applied in metabolomic studies.

In-parallel application of metabolite and phytohormone profile analysis to seedlings of the model plant *Arabidopsis thaliana* exposed to sulfate limitation was used to investigate the relationship between the endogenous concentration of signal elements and the ‘metabolic phenotype’ of a plant. An automated evaluation strategy was developed to process data of compounds with diverse physiological nature, such as signal elements, genes and metabolites – all which act *in vivo* in a conditional, time-resolved manner (Chapter IV). Final data analysis focussed on conditionality of signal-metabolome interactions.

Table of Contents

Tables

Figures

Introduction

The question of how changes in the environment are recognised by organisms, and how organisms adapt to those changes, is likely to have been one of the first 'simple' questions people were asking when the science of biology was established. Since then, as a matter of course, instruments and experiments designed to answer that question have become more and more advanced. It has become clear, that there will be no easy answer at all. The distinct recognition of a stimulus, the transmission of the signal throughout an organism, and the induction of response processes culminating in acute and permanent adaptation to the original impulse, all seem to be connected in a fashion too complex to be explained by a simple, 'general' model.

Concerning signal transmission in plants, candidate substances have been found acting similar to animal hormones; after them they were named 'phytohormones'. Investigations in this work address phytohormones as highly important messengers for all kinds of developmental and environmental impulses in plants. The terminus 'hormone' originates from Greek language and means 'to stimulate'; thus, phytohormones can be described as chemical transmitters which coordinate and regulate physiology, growth and morphogenesis of plants.

Three characteristics are important in defining a substance as a hormone: translocation, signal transduction, and elicitation of a physiological reaction. In contrast to mammalian hormones, plant hormones miss a clear defined site of biosynthesis. Often, there is only patchy evidence for translocation. Thus, translocation as an imperative for defining a substance as a phytohormone is still under discussion. Other circumstances influencing phytohormone action are that the cell wall acts as a barrier to cell-to-cell communication; plants exhibit often longer reaction times to response to environmental changes and are more dependent on surroundings than mammals due to their immobility (Taiz and Zeiger 1998; Hammond-Kosack and Jones 2000).

Currently, according to the common consensus, nine groups of phytohormones are known: ethylene, auxins (Went 1928), gibberellins (Brian et al. 1954), cytokinins (Letham 1963), salicylates, abscisates (Bennet-Clark and Kefford 1953), brassinosteroids (Grove et al. 1979), jasmonates and polypeptides such as systemine (Pearce et al. 1991). Other compounds have also been claimed to exhibit signalling effects in plants, for example polyamines (Bagni and Serafini-Fracassini 1973).

Despite many years of work, no complete picture has emerged concerning the role of these plant growth regulators; their physiological roles are complex and poorly understood. Although being present only in trace amounts in plant tissue, which makes bioanalytical quantification rather difficult, phytohormones play a major role in plant growth and development.

Often it is not only the mere endogenous concentration of a phytohormone that changes in response to a stimulus but also the availability of a plant receptor or the transmission system for a certain signal compound, and the ratio of phytohormones to each other modify (Trewavas 1982; Trewavas 2000; Weyers and Paterson 2001). The developmental stage of a plant can be as important for the mode and extent of phytohormone action as environmental conditions like water, light or temperature (Bray et al. 2000). The effects caused by a plant hormone are a function of particular circumstances and, consequently, of changes in hormonal balance. Therefore, it is advantageous not only to measure one signalling compound under particular ambient circumstances, but also to observe phytohormone crosstalk by determination of additional, possibly interacting, (signal) compounds.

Phytohormone Analysis

First methods used for quantification of phytohormones were bioassays that quantified rather the extent of phytohormone action than their actual endogenous concentration. In bioassays, known effects of particular phytohormones are exploited to acquire a dose – response curve in control samples and to deduce from the extent of that responses to the 'dose' of the phytohormone present in the sample. Thus, the first well-known bioassay for a phytohormone was developed for auxin, the first invented phytohormone (Went 1928). Went developed the *Avena* test, which uses the capability of auxin to stimulate elongation growth on the light-abandoned side of a decapitated coleoptile.

Bioassays are highly sensitive and can be carried out by using a single plant individual. But soon, the disadvantages of these assays became clear: the performance of bioassays is time consuming and the outcomes exhibit a rather poor precision. All-too often interference with different-type signal compounds affects analysis; also, it is difficult to differentiate between different representatives of the respective phytohormone class. The quantification of compounds that are physiologically inactive under the given experimental conditions is

not possible (Weiler et al. 1983). Therefore, other analytical methods were tested in order to achieve a sensitive and reproducible determination for signal compounds.

Thus, Pengelly and Meins (1977) introduced the radioimmunoassay (RIA) to analysis of auxin and uncovered the potential of this method for phytohormone quantitation. In short succession, RIAs for various phytohormones were developed, including also abscisic acid (Weiler 1979), cytokinins (Weiler 1980) and gibberellic acid (Weiler and Wieczorek 1981). RIA analysis of phytohormones comprises several advantages compared to bioassays. Thus, first of all they enabled a higher sample throughput than bioassays, because they can be performed on crude plant extracts without time-consuming sample purification. RIAs approved to be highly sensitive but do not require a particular sophisticated equipment compared to physicochemical methods (Weiler 1984). However, the applicability of this method is restricted to the availability of a suitable antiserum.

Nowadays, the main field of phytohormone analysis but have become chromatographic methods hyphenated with sensitive detection techniques. Gas-liquid chromatography (GC), for instance, provides the advantage of an unsurpassed separation power in comparison to other chromatographic methods. In combination with mass spectrometry (GC-MS), it is a powerful tool for reproducible, precise, and sensitive detection and identification without compromising specificity for particular compounds. A great advantage of instrument-based versus antibody-based methods is the capability of multi-selective analysis, which means the simultaneous analysis of multiple components within the same sample aliquot.

Signal compounds comprise two groups of substances, volatile and non-volatile signal molecules; the latter are non-volatile mainly due to their relative polarity or their molecular weight. Thus, auxin, abscisic acid, salicylic acid, and jasmonic acid, for instance, are non-volatiles. In comparison, the methyl esters of salicylic acid and jasmonic acid, methyl salicylate and methyl jasmonate, for instance, are known to be highly active volatile signalling compounds (Pickett et al. 2003); the best characterised volatile plant hormone is the gas ethylene.

To apply GC-MS in analysis of signal compounds, the non-volatile compounds have to be modified in order to increase their volatility. This can be done by derivatisation of the compounds, which is a chemical reaction aiming on modification (conjugation) of polar functional groups (e.g. hydroxyl, amino, or carboxyl groups) in order to decrease their polarity prior to GC-MS separation. Different derivatisation methods are used for phytohormones (see Chapter I), but common are methylation for the acid phytohormones (Miersch et al. 1991; Mueller et al. 2002) and trimethylsilylation (Palni et al. 1983;

Badenoch-Jones et al. 1984; Croker and Hedden 2000). The drawback is, however, that in some cases the volatility of the compounds cannot be sufficiently improved for GC-MS analysis, or the molecules may be too unstable and decompose during derivatisation or the evaporation step in the GC injector.

Thus, other hyphenated techniques recently have also been used for analysis of signal compounds. The high selectivity of the mass detector enables the detection of specific compounds even in a complex mixture containing other, more abundant, compounds that might co-elute with trace compounds like the phytohormones. Separation techniques with a lower number of theoretical plates thus can also be used, when hyphenated to mass spectrometry, and so liquid chromatography – mass spectrometry (LC-MS) was introduced in phytohormone analysis (Schneider and Schmidt 1996; Prinsen et al. 1998). Another method suitable for non-volatiles with relatively high separation efficiency that was used for the analysis of signal compounds was capillary electrophoresis, CE (Liu et al. 2002). A coupling of CE and MS is expected to extend the potential of this method. However, GC-MS or LC-MS are nowadays analytical methods that are commonly in use for analysis of phytohormones.

The Concept of Profiling Methods

Prior remarks might have made clear the complexity of facts to consider for investigations addressing signal-metabolome interactions in plants; not only concerning the biochemical analysis of the compounds under investigation, but also the design of appropriate biological experiments. Metabolic effects in consequence to phytohormone signalling are expected to appear highly diverse and tightly dependent on ambient conditions (Trewavas 2000). In order to approach the complexity of metabolic reactions in response to phytohormone signalling, it was decided to use profiling techniques for the analysis of plant metabolites and phytohormones.

When the method of metabolite profiling was introduced as a tool for functional genomics (Trethewey et al. 1999; Roessner et al. 2000; Fiehn et al. 2000b) the profiling concept attracted increasing attention in biochemical analysis, but the profiling concept itself is not new to the field of biochemistry (see Mamer 1994 for review). The acquisition of a profile means to monitor a set of compounds, summarised by a certain, common criteria of these compounds. The outcome is a pattern of analyte concentrations under different

experimental conditions, instead of monitoring only one single target. Thus, profiling methods mean to determine a set of similar compounds at once using only one single sample set-up. Within this context, profiling methods have been used widely (see Mamer 1994), for studying certain pathways (Kopka et al. 1995; Berger et al. 2001), or comparing the entirety of compounds with common biochemical characteristics (Bligh and Dyer 1959; Katona et al. 1999) and functions, such as compounds that are released by plants for communication with the environment (Agelopoulos et al. 1999).

The innovation of the profiling concept, which was introduced by Roessner et al. (2000), was the approach not to restrict analysis to metabolites that meet certain physiological criteria, but to aim for a 'non-selective' model of metabolite profiling, which means no other restriction to the analysis than simply the accessibility of metabolites by the used technique. Therefore, sample preparation and further analysis was designed to be as unspecific as possible under the regime of the used experimental set-up (Roessner et al. 2000).

This profiling concept provides the major advantage to document also unexpected effects that would remain undiscovered, if not included into a targeted analysis. On the other hand, drawbacks of the 'non-selective' approach for metabolite profiling are obvious: a traditional quantitation on basis of collective calibration, namely the absolute quantification of several analytes using an internal standard, or quantification using external calibration curves became rather difficult and complex when the range of envisaged analytes became chemically highly divers. Thus, initially the strategy of a 'relative' quantification was used for metabolite profiling, namely simply to check for (significant) quantitative differences of detector-response signals between analytes present in the sample compared to the respective control.

Further, metabolite profiling did not compensate for discrimination of trace compounds such as phytohormones that simply disappear in the noise of the plant-extract matrix. Among others, it is mainly this cause that legitimates continuing development of profiling methods in the 'classical' sense, meaning the joint analysis of compounds exhibiting similar characteristics or functions with respect to the envisaged investigations as described above.

Concerning a profiling method for phytohormones, it were first Rademacher and Graebe (1984) who introduced the simultaneous analysis of auxins, gibberellins, cytokinins, and abscisic acid in one single sample set-up using GC-MS. Since then, many methods have been developed to measure different sets of signalling compounds together in one sample

set-up, mostly using hyphenated analytical techniques such as LC-MS or GC-MS (Prinsen et al. 1998; Mueller et al. 2002; Dobrev and Kaminek 2002; Liu et al. 2002; Schmelz et al. 2003).

Modern experimental designs narrow microscale and tissue level also for trace compounds such as phytohormones (Weiler 1984; Mueller et al. 2002). In instrumental analysis of phytohormones using hyphenated techniques, highly selective detection methods are applied in order to increase the sensitivity of detection, such as single ion monitoring (SIM), tandem-MS (MS-MS), or multiple reaction monitoring (MRM) techniques (Chen et al. 1988; Moritz and Olsen 1995; Prinsen et al. 1998; Mueller et al. 2002).

Also, there is an increasing demand for ease of sample preparation procedures for phytohormone analysis and in the same time increasing the recovery of the analytes. Earlier, samples were purified using preparative HPLC with fractionation prior to GC/LC-MS analysis (Sembdner et al. 1987; Prinsen et al. 1998). Recently, sample preparation procedures were developed that use solid phase extraction (SPE) for purification of analytes in phytohormone single-targeted and also profiling methods (Chen et al. 1988; Edlund et al. 1995; Mueller et al. 2002, Dobrev and Kaminek 2002). SPE cartridges provide the advantage of little time and less technical instrumentation, a fact which supports the increasing application of SPE-based procedures for the analysis of these trace compounds (Chapter I).

Signal - Metabolome Interactions

In plant physiology, an altered phenotype is the first visible effect in consequence to the biological experiment. In times, where the availability of technical equipment and modern evaluation tools for biochemical analysis was not as self-evident as nowadays, the phenotype was the first hint for the experimenter to conclude, which pathways could be affected or which signalling compound might be involved in regulation of the processes under investigation; but the conclusion from phenotype to hormone action is always an ambiguous matter.

Phytohormones were found to cause a range of various effects on metabolism of plants and are active in tight connection to ambient conditions (Taiz and Zeiger 1998). But between the eliciting event and the actually altered phenotype a range of other, regulating and

regulated, physiological events can be found at molecular level, and signal transmission by phytohormones is only one of those events.

Phytohormone action is embedded in processes, which begin when the recognition of a stimulus causes a change in hormone concentration or in concentration of hormone receptors (Lichtenthaler 1996). After binding of the signal molecule to a specific receptor, the so-called early response is initiated (Taiz and Zeiger 1998). This early response induces an immediate short-term response on the one side, and activates gene transcription for long-term response (e.g. adaptation) of the plant organism on the other side. The present work focuses on the feasibility to deduce from an endogenous signal such as altered phytohormone concentration levels to effects on the concentration levels of primary and secondary metabolites as physiological short- or long-term response, respectively.

In conclusion, tracing back the history of physiological consequences to phytohormone action becomes an obvious imperative in order to understand the causality and interconnectivity between signalling and adaptation on metabolic level. The present work addresses the development of a methodology that will enable the investigation of using the 'metabolic phenotype' of plants for identification and understanding of cross-references of phytohormone action. To explore the impact of signalling on metabolic changes and the conditionality of these coherences, it was aimed to combine analysis and evaluation of the metabolite profiling technique as introduced by Roessner et al. (2000) and Fiehn et al. (2000b) with the analysis of phytohormones.

Chapter I - Phytohormone Profiling - Method Development

Comprehensive Chemical Derivatisation for Gas Chromatography – Mass Spectrometry-Based Multi-Targeted Profiling of the Major Phytohormones

Journal of Chromatography A, Volume 993, Issues 1-2, 18 April 2003, Pages 89-102

Claudia Birkemeyer, Ania Kolasa, Joachim Kopka

Abstract

In the present investigation we report selection of the N-methyl-N-(tert.-butyldimethyl-silyl)-trifluoroacetamide (MTBSTFA) reagent as the most comprehensive derivatisation protocol among 17 tested reactions, i.e. trifluoroacetylation, pentafluorobenzylation, methylations, and trimethylsilylations. MTBSTFA allowed easy-to-perform and robust tert.-butyldimethylsilyl derivatisation of 1-aminocyclopropane-1-carboxylic acid, indole-3-acetic acid, (±)-jasmonic acid, salicylic acid, (±)-abscisic acid, meta-topolin, and trans-zeatin. Detection limits as analysed by selected ion monitoring quadrupole GC-MS were 0.2, 0.01, 1.0, 0.02, 0.3, 0.3, and 0.9 pmol of injected substance, respectively. Analysis of gibberellic acid A3, trans-zeatin riboside and (±)-abscisic acid-β-D-glucopyranosyl ester were best be coupled by splitting extracts and alternative trimethylsilylation. The MTBSTFA derivatisation protocol was optimised and validated. The preparation was insensitive to 2% residual water and to ≤ 1 day storage at room temperature. The final scheme was highly reproducible and successfully applied to extracts from approximately 300 mg fresh weight (FW) of tobacco (*Nicotiana tabacum*) root and *Arabidopsis thaliana* seedlings.

Introduction

Identification of auxin, the first phytohormone discovered by Went (1928), spurred a strong and lasting interest of fundamental research and applied biotechnology in plant growth regulators. In succession, abscisic acid, gibberellins, cytokinins, ethylene, jasmonic acid and salicylic acid were described, identified and demonstrated to exhibit respective

regulatory functions. Even recently, novel signalling substances, such as brassinolides and the oligopeptide systemin (Pearce et al. 1991), were found in plants.

The past and recent analysis of phytohormone action led to the emergence of the concept that none of the crucial biological functions, for example growth rate, growth orientation, development, and water balance, could be completely explained in a mono-causal manner. In contrast, interplay of phytohormone levels nowadays appears to be more important to our understanding of phytohormone function than absolute concentrations of any single substance (Davies 1995; Weyers and Paterson 2001). This novel insight was the incentive for our effort to establish multi-targeted phytohormone profiling as an extension to our recently introduced technology for the GC-MS profiling of primary metabolites (Roessner et al. 2000; Fiehn et al. 2000b). Today novel developments in quantitative phytohormone analysis are directed at either multi-parallel analysis or at increased sensitivity without compromising selectivity of detection. Downscaled sample requirement will increase spatial resolution of our knowledge about phytohormone action, whereas multi-parallel analysis allows novel systems insights into the interplay of phytohormone action. Our final goal is the efficient, sensitive, and comprehensive multi-targeted quantification of phytohormones from a single sample.

Several publications already addressed the challenge of developing a suitable method for phytohormone profiling based on instrumental analytical technologies (Rademacher and Graebe 1984; Prinsen et al. 1989 and 1998; Schneider and Schmidt 1996; Kowalczyk and Sandberg 2001; Dobrev and Kaminek 2002; Liu et al. 2002; Mueller et al. 2002). Phytohormones like most constituents of signal transduction pathways are trace compounds. Thus phytohormone analysis is subject to the common complications in trace analysis, namely laborious multi-step clean-up procedures, strong influence of sample matrix and ambient conditions (Rivier and Crozier 1987). The analytical platform of choice was gas chromatography coupled to mass spectrometry because of unsurpassed instrumental versatility, selectivity, sensitivity, and long-standing previous application in phytohormone analysis. Novel coupling technologies like solid phase micro extraction, GC-GC coupling, and MS-MS techniques extend the already ample instrumental toolbox towards further means of micro-concentration and micro-separation.

Crucial for successful GC analysis is appropriate and stable derivatisation of non-volatile compounds. Indeed nearly all major classes of phytohormones comprise polar compounds with high-boiling points. A wide range of derivatising protocols are available from comprehensive compendium guides (Knapp 1979; Blau and Halket 1993). Some were

already successfully applied to analysis of different phytohormone classes. Trifluoro-acetylation was reported for cytokinins (Ludewig et al. 1982). Trimethylsilylation was described for cytokinin (Palni et al. 1983) and auxin (Badenoch-Jones et al. 1984) analysis. Tert.-butyldimethylsilylation was applied to cytokinins (Hocart et al. 1986). Alkylation with pentafluorbenzylbromide was successful in cytokinin (Letham et al. 1991) and auxin (Prinsen et al. 2000) quantification. Methylation with diazomethane was commonly reported in publications on jasmonic acid (Miersch et al. 1991), auxin, salicylic acid, and abscisic acid (Mueller et al. 2002). Two-step procedures consisting of alkylation with diazomethane and subsequent trimethylsilylation were described for auxin (Edlund et al. 1995) and gibberellins (Croker et al. 1994). A brief summary of further analytical methods for the quantification of the major phytohormones can be found in (Sembdner et al. 1987).

In the present study we reinvestigated and compared those chemical modification schemes which are in frequent use for the GC-MS analysis of phytohormones and which appeared to be versatile. In order of priority, the tested reagents were selected according to ease of handling, comprehensiveness of derivatisation, and molar response ratio. The most promising scheme of a multi-parallel analysis was further optimised, validated, and standardised with a representative selection of phytohormones and other chemically related reference substances. Finally, we applied our method to plant matrices using a previously published extraction and clean-up procedure (Mueller et al. 2002). We introduce a sensitive, robust and easy-to-handle derivatisation scheme appropriate for routine analysis of the major phytohormone classes from single plant samples.

Experimental

Standards and reagents

1-Aminocyclopropane-1-carboxylic acid, ACC (CAS 22059-21-8); myo-inositol, INO (CAS87-89-8); (±)-jasmonic acid, JA (CAS 3572-66-5); DL-tryptophan, TRP (CAS 54-12-6), gibberellic acid A3, GA3 (CAS 77-06-5) and 5α-cholestane, CH (CAS 481-21-0), n-nonadecane (CAS 629-92-5), DL-α-tocopheryl acetate (CAS 7695-91-2) and the pesticide standard mixtures 8081 and EPA 508/508.1 were purchased from Sigma-Aldrich, Germany; meta-Topolin, mT, and 24-epibrassinolide, BL (CAS 78821-43-9), were ordered from Duchefa, Netherlands; trans-zeatin, Z (CAS 1637-39-4), indole-3-acetic acid, IAA (CAS 87-51-4) and salicylic acid, SA (CAS 69-72-7) were from Merck, Germany; (±)-

abscisic acid, ABA (CAS 14375-45-2); (±)-abscisic acid-β-D-glucopyranosyl ester, ABA-GE, and trans-zeatin riboside, ZR (CAS 6025-53-2) were received from Apex Organics, UK. Chemical abstracts system (CAS) registry numbers of the reference substances are provided where available.

The reagents were purchased as follows: N-methyl-N-(tert.-butyldimethylsilyl) trifluoro-acetamide (MTBSTFA); N-methyl-N-(trimethylsilyl) heptafluorobutyramide (MSHFBA); N-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA); N,N-dimethylformamide dimethylacetale (DMF-DMA), trimethylsulphonium hydroxide (TMSH) were from Macherey & Nagel, Germany. N,O-bis(trimethylsilyl)-acetamide (BSA), pentafluorobenzylbromide (pFBBr); N,O-bis(trimethylsilyl)-trifluoroacetamide (BSTFA); N-trimethylsilyl-imidazole (TSIM) ; trimethylchlorosilane (TMCS); hexamethyldisilazane (HMDS); N-methyl-bis(trifluoroacetamide) (MBTFA); trimethylphenyl ammonium-hydroxide (TMAH); methyliodide (MeI)/ potassium carbonate were ordered from Fluka, Germany. Pyridine, methanol, dichlormethane (DCM), chloroform, ethylacetate (EtOAc) and formic acid, all HPLC-grade, were supplied by J.T. Baker (Philipsburg, NJ). Diazomethane was synthesised as described by Schlenk and Gellerman (1960).

Sample preparation

Derivatisation protocols

Standard stock solutions for the comparison of derivatisation protocols were prepared in methanol at concentrations of 1 mg/ mL. Only Z required addition of 1% (v/v) formic acid. 5 µl of a 1:2 (v/v) dilution of each reference substance was combined with an equal volume of a 1:10 dilution of the 5α-cholestane stock solution, dried under nitrogen, and subjected to the derivatisation procedures described below. Final amounts of 0.083 µg of each reference substance were used for GC-MS analysis. In the experiment addressing possible side product formation of MSTFA, BSTFA, MTBSTFA and MSHFBA reactions 0.166 µg SA was used (Table I-2). Reaction parameters, e.g. solvent, volume ratios, incubation time and temperature, were not optimised for the initial screening. Instead general manufacturer's recommendations were applied (1993) unless stated otherwise.

All procedures were carried out in at least three replications and performed at room temperature if not indicated differently.

Trifluoroacetylation with MBTFA
Dissolve in 100 µl EtOAc, add 25 µl reagent, and heat to 120°C for two minutes before analysis.

Methylation with diazomethane (Schlenk and Gellerman 1960):

Add saturated ethereal solution of diazomethane, until yellow colour is persistent, evaporate sample under nitrogen, and redissolve in chloroform for GC-MS.

Methylation with DMF-DMA:

Dissolve in 100 µl EtOAc, add 1000 µl DMF-DMA in pyridine 1:1 (v/v), and inject for GC-MS analysis when solution becomes clear after 0.5 – 3 min.

Methylation with TMAH:

Dissolve in 100 µl EtOAc, add 1 µl reagent, incubate for 10 min, and analyse directly by GC-MS.

Methylation with MeI/ potassium carbonate (Knapp 1979):

Dissolve in 40 µl MeI/ EtOAc 1:1 (v/v), add 1-2 mg potassium carbonate, incubate for 1 h at 90°C, and inject clear supernatant.

Methylation with TMSH:

Dissolve in 100 µl EtOAc, add 50 µl reagent, incubate for 10 min at 100°C, and analyse by GC-MS.

Silylation with TSIM, MSTFA, BSTFA, MTBSTFA, and MSHFBA (Little 1999):

Add 100 µl reagent, incubate for 30 min at 90°C, and analyse by GC-MS.

Silylation with HMDS/ TMCS/ Pyridine 3:1:9 (v/v/v) (Knapp 1979):

Dissolve in 100 µl reagent mixture and incubate 30 min before analysis by GC-MS.

Silylation with BSA/ TSIM/ TMCS and HMDS/ TMCS/ Pyridine 1:1:1 (v/v/v); and HMDS/ TMCS 1:1 (v/v) (Knapp 1979):

Dissolve in 100 µl reagent mixture and incubate 1 h before analysis by GC-MS.

Optimisation of the MTBSTFA protocol

Analysis of reproducibility, incubation time, incubation temperature, and the search for an internal standard substance with improved performance were carried out with 5 µl of 0.5 mg/ mL stock solutions of each reference substance, n-nonadecane, and DL-α-tocopheryl acetate, which were combined with 3 µl of a 0.1 mg/ mL solution of 5α-cholestane. Samples were dried under a stream of nitrogen and incubated in 25 µl MTBSTFA reagent prior to quadrupole GC-EI-MS analysis. Incubation was checked at 40, 60, 80, 100, 120 and 120°C temperature and 30, 60, 120 and 180 min reaction time (n = 7). All subsequent experiments were performed with optimised conditions, namely 1 h at 100°C. Pesticide standard mixtures 8081 and EPA 508/508.1 were tested for potential candidate standard substances by adding 5 µl of each commercial preparation after heating.

Samples for storage stability tests were prepared as described above in 25 µl MTBSTFA and sealed in GC vials until further analysis. Storage was either at room temperature, 20 - 25°C, or at –20°C. A parallel set was sealed with 0.5 µl water added prior to incubation. Three replications of each set were analysed subsequently at 5 h, 12 h, 24 h, 3 days, 7 days, and 14 days after start of incubation. For analysis of variance the 5 - 24 h measurements were combined into the level '≤ 1 day' of the factor storage time, results of day 3 – 14 comprised the alternate level. Analysis of variance (ANOVA) was performed with the statistical software package S-Plus 2000 standard edition release 3 (Insightful, Seattle, USA).

Calibration curves and limits of detection were performed using multiple-component samples which were prepared by dilution of independent stock solutions.

Phytohormone profiling

Tobacco plants, *Nicotiana tabacum* cv. Samsun NN, were grown in sand under optimum growth chamber conditions. Roots were harvested three months after germination, rinsed under tap water until free of sand, and were then snap frozen in liquid nitrogen.

Arabidopsis thaliana seedlings were germinated under sterile conditions on solid support and harvested after two - three weeks. Representative batches were sampled, homogenised in a mortar under liquid nitrogen, and stored at -80°C. 300 mg frozen fresh weight of these samples were extracted in 10 mL/ g FW of Bieleski solvent pre-cooled to -20°C (Bieleski 1964).

Purification of phytohormones from plant extracts was done as described by Mueller et al. (2002), except for omission of the silica-based aminopropyl purification step. No further attempts at optimisation were undertaken in the context of the present work. The final preparation was concentrated by vacuum centrifugation, 1 min at 200 mbar followed by 10 mbar to dryness. The optimised reaction protocol was applied and the MTBSTFA derivatives analysed by quadrupole GC-EI-SIM-MS and ion trap GC-CI-MS with methanol as reactant gas and MS-MS reaction monitoring (GC-CI-MS-MS).

GC-MS analyses

GC-MS systems used in the present work were (1) a MD 800 GC-MS system (ThermoFinnigan, San Jose, USA) of quadrupole technology supplied with split/ splitless injection and MassLab software Version 1.4 and (2) an ion trap Saturn 2000 GC-MS system (Varian, Palo Alto, USA) supplied with programmed temperature vaporisation injection and Saturn Workstation software version 5.4. AMDIS software (Stein 1999) was employed to support peak finding before quantitative analysis and for automated deconvolution of reference mass spectra. Identification of derivatives and side products was performed by co-chromatography and mass spectral fragmentation. The identification was supported by comparison to mass spectra presented by Gaskin and MacMillan (1991) as well as to a commercial mass spectral library in NIST98 format (Ausloos et al. 1999). The quadrupole system was chosen for the analysis of side product formation (Table I-2), because mass spectral deconvolution software in combination with this technology allowed improved automated detection and better mass spectral comparisons with available libraries as compared to ion trap recordings. A data file in the interchange format for AMDIS and NIST98 containing all mass spectra mentioned in Table I-2 and respective ion trap mass spectra is appended as supplementary data for future cross-referencing.

Quadrupole GC-MS chromatograms were monitored by Electron Impact ionisation and either total ion monitoring, m/z 40 – 600, or in the case of determining calibration curves and detection limits via segmental selective ion monitoring (GC-EI-SIM-MS). Selected fragments were m/z = 272 (ACC), m/z = 232 (IAA), m/z = 133 (JA), m/z = 309 (SA), m/z = 190 (ABA), m/z = 469 (mT), m/z = 302 (Z), and m/z = 217 (5α-cholestane).

The ion trap mass spectrometer was operated in the EI-MS mode with total ion monitoring during initial analysis of derivatisation protocols. Phytohormone profiles of plant samples were monitored in the CI-MS-MS mode with methanol gas and positive ion detection.

Maximum reaction time was 128 msec, maximum ionisation time 2 msec, scan rate 0.38 sec $scan^{-1}$, multiplier offset 300 V, emission current 30 µA. The resonant waveform type was adjusted to MS-MS mode with a parent-ion-selection window of three atomic mass units. Parent-ion selections and excitation amplitudes were segmental and changed as follows: ACC $[M+H]^+$, m/z = 330, excitation amplitude 0.6 V; SA $[M+H]^+$, m/z = 367, excitation amplitude 0.6 V; JA $[M+H]^+$, m/z = 325, excitation amplitude 0.5 V; IAA $[M+H]^+$, m/z = 290, excitation amplitude 0.5 V; ABA $[M-H2O+H]^+$, m/z = 361, excitation amplitude 0.6 V.

Arylene type 5% phenyl/ 95% methyl polysiloxane fused silica capillary columns were chosen. A 30 m Rtx-5Sil MS phenyl-arylene fused silica column, 0.25 mm inner diameter, 0.25 µm film thickness, supplied with a 10 m-guard column (Restek, Bad Homburg, Germany) was used for the tests of different reagents and derivatisation protocols. The GC-MS system was preconditioned each time the reagents were changed. Optimisation of the MTBSTFA protocol and phytohormone profiles of plant samples were performed without changes in performance on less expensive 30 m DB 5-MS phenyl-arylene fused silica columns with 0.25 mm inner diameter, 0.25 µm film thickness (Agilent J&W; USA).

Injection was hot splitless at 230°C with an oven temperature ramp of 6°C/ min from 70°C to 350°C, ion source temperature was set to 230°C, and transferline was at 260°C. Helium carrier gas was used at a flow rate of 1 mL/ min. These settings were used for all reagents and represent a compromise of previously described analyses (Ludewig et al. 1982; Palni et al. 1983; Badenoch-Jones et al. 1984; Hocart et al. 1986; Letham et al. 1991; Croker et al. 1994; Edlund et al. 1995; Miersch et al. 1999; Prinsen et al. 2000). The GC method was designed to cover a high temperature range and still separated at least two derivatives from a commercial (±)-jasmonic acid isomer mixture.

The Saturn 2000 System was operated with a temperature program for controlled vaporisation after injection, 0.5 min at 110°C followed by a 250°C/ min ramp to 230°C. This temperature program was optimised for the MTBSTFA derivatives.

Definitions and calculations

Response was defined as chromatographic peak areas derived from mass spectrometric total ion, selected ion, or MS-MS recordings. Molar response was calculated as the quotient of response over mol of substance injected into the GC-MS systems. Molar amount of injected substance was estimated by initial weight, dilution factor before

derivatisation, final volume of derivative and volume injected. Molar response ratios were the quotients of the molar responses of reference substances and a underivatised internal standard substance, 5α-cholestane. The reaction procedure was optimised (Figure I-1) and tested for robustness by monitoring the relative yield of each derivative. The relative yield was calculated as percentage of the maximum molar response ratio of each derivative in the respective experiments.

Results and discussion

Comparison of derivatisation protocols

The reference substances for the following investigations were selected to cover most phytohormone 'classes' by a single commercially available and affordable, natural occurring compound. Thus seven major phytohormone classes were represented by IAA, JA, SA, ABA, Z, mT, GA3, and BL (Table I-1). Systemin had to be excluded because GC-MS technology is clearly unsuited for the analysis of oligopeptides. The ethylene precursor, ACC, was included instead of the gaseous phytohormone. Furthermore we attempted to represent common phytohormone conjugates and typical functional groups by the two reference substances, ABA-GE and ZR. Reference substances of the equally important amino acid and inositol conjugates were not commercially available. Therefore, we included TRP and INO in order to assess their respective chemical behaviour.

The comparative analyses of derivatisation reactions necessitated a common substance for internal volume standardisation. This substance was required to be inert with respect to all tested reagents. Therefore, initial experiments focussed on the use of a range of hydrocarbons (data not shown). 5α-cholestane was the best choice available with respect to inertness, intermediate volatility and distinctive fragmentation.

In Table I-1 we summarised the results of the initial screening for reagents and protocols that are listed in the experimental section, 'derivatisation protocols'. In cases of multiple derivatives only those with highest molar response ratio are shown. IAA, JA, and ABA were easily detected with all protocols, but no strategy of chemical modification allowed analysis of all reference substances.

Table I-1:
Molar response ratios of the main derivatives obtained from 0.083 µg of each reference substance[a].
Molar response ratios were calculated from ion trap EI-MS total ion currents by normalisation to the signal of an equal amount of 5α-cholestane within each preparation. The table was compiled from multiple experiments. Each experiment was performed with aliquots of the same reference substance mixture (n = 3). Values > 0.1 are in bold format. A tert.-butyldimethylsilylation; B trimethylsilylations; C methylations; D combined trimethylsilylation and methylation; E trifluoroacetylation; F pentafluorobenzylation

Reagent[c]	Reference Substance[a]												
	ACC	IAA	JA	SA	ABA	ABA-GE	mT	Z	ZR	GA3	BL	TRP	INO
A													
MTBSTFA	**0.146**	**0.237**	**0.149**	**0.270**	**0.262**		**1.042**	**0.440**	0.034	**0.169**		**0.376**	
B													
30-60 min; room temperature:													
BSA-TMCS-TSIM (1:1:1)[b]		0.005											
HMDS-TMCS-Pyridin (1:1:1)[b]		0.090	0.098	0.096	0.009					0.074			
HMDS-TMCS-Pyridin (3:1:9)[b]		0.073	**0.117**		0.007			0.021		**0.154**			
HMDS-TMCS (1:1)[b]	0.048	0.069	**0.115**	0.015	**0.121**			0.057		**0.588**			
30 min; 90°C:													
TSIM		0.005											
BSTFA	< 0.001	**0.179**	**0.157**	**0.153**	**0.197**	0.010		0.044		**0.907**		0.009	**0.596**
MSTFA	0.069	**0.252**	0.096	0.098	**0.175**	0.020	0.014	**0.131**	**0.311**	**0.801**		**0.123**	**0.404**
MSHFBA	0.087	**0.231**	**0.133**	**0.149**	**0.146**	0.029	0.040	0.095	**0.343**	**0.691**		0.088	**0.231**
C													
MeI		0.066	0.026	< 0.001						**0.138**			
Diazomethane		0.055	**0.133**	0.007	**0.146**					0.083	0.012		
DMF-DMA		0.050	0.088		**0.130**			0.002		0.088			
TMSH		**0.360**	0.056	0.001	**0.149**			0.066		0.009			
TMAH		**0.116**	**0.105**	0.002	**0.147**			0.009		0.072			
D													
Diazomethane-MSTFA	0.015	**0.179**	0.083	0.047	**0.139**			0.008		**0.643**	0.020		**0.184**
E													
MBTFA		0.087	0.051		0.004					0.012			
F													
PFBBr	**0.288**	0.086	**0.115**	**0.259**	**0.317**							**0.200**	

[a] ACC, 1-aminocyclopropane-1-carboxylic acid; IAA, indole-3-acetic acid; JA, (±)-jasmonic acid; SA, salicylic acid; ABA, (±)-abscisic acid; ABA-GE, (±)-abscisic acid-β-D-glucopyranosyl ester; mT, meta-topolin; Z, trans-zeatin; ZR, trans-zeatin riboside; GA3, gibberellic acid A3; BL, 24-epibrassinolide; TRP, DL-tryptophan; INO, myo-inositol
[b] Volume ratios
[c] Refer to section 2.1. for full identification of reagents

Best coverage and molar response ratios were obtained with silylating reactions. Combined silylation and methylation increased the number of side products but allowed detection of BL. However, cytokinins, ABA-GE, and ZR were lost. This observation was also made for all stand-alone methylation reactions. Trifluoroacetylation exhibited strong selectivity and low molar response ratios. PFBBr was communicated as a highly sensitive reagent and robust to residual water (Epstein and Cohen 1981). In addition, pentafluorobenzylation is ideally suited for negative chemical ionisation NCI-MS (Knapp 1979; Letham et al. 1991). In our hands, PFBBr-derivatives exhibited high sensitivity even when monitored with EI-MS, but allowed detection only of ACC, IAA, JA, SA, ABA, and TRP.

Four silylating reagents with high donor strength, BSTFA, MSTFA, MSHFBA, for trimethylsilylation and MTBSTFA, the latter transferring tert.-butyldimethylsilyl-groups, appeared to be comprehensive. Comparison with less reactive silylating reagents under mild reaction conditions demonstrated that high reactivity was essential for this

observation. Therefore, only the highly reactive reagents, BSTFA, MSTFA, MSHFBA, and MTBSTFA, were further investigated with regard to the formation of side products and compared in a single large-scale experiment using a quadrupole GC-EI-MS system (Table I-2).

Most molar response ratios were increased as compared to previous ion trap results (Table I-1). This effect was caused to a large extent by a reduced molar response of 5α-cholestane. Interestingly, the quadrupole GC-MS system appeared also to discriminate the INO derivative and to be more sensitive for the derivatives of mT, Z, TRP, and ACC. The overall relative standard deviation (RSD) of these experiments including all minor products was 21, 31, 27, and 24% (n = 3) including all minor products of the reactions with MTBSTFA, BSTFA, MSTFA, and MSHFBA, respectively.

The most comprehensive derivatisation was trimethylsilylation. All trimethylsilyl-reagents, BSTFA, MSTFA, and MSHFBA, generated a single main product and identical side products except for ABA-GE and BL. In the case of ABA-GE low signal intensity was likely caused by instability of the conjugate as concluded by occurrence of free silylated glucose (data not shown). Compared to trimethylsilyl reagents, MTBSTFA was slightly more sensitive and less prone to formation of side products (Table I-2; see TRP and mT). In some cases, MTBSTFA exhibited a preference for a lower degree of substitution, namely IAA and TRP. Unfortunately, this property of MTBSTFA did not allow analysis of GA3, ZR, ABA-GE, BL, or INO. In the case of GA3 and ZR we detected minor signals of derivatives with low degree of substitution, but the bulk derivative was lost.

In view of the ultimate goal of our efforts - the sensitive close to comprehensive multi-targeted quantification of phytohormones - we decided on an in-depth analysis of the MTBSTFA derivatisation reaction. This decision took into account firstly the prospective sensitivity, namely the combined aspects of low side product formation, high molar response, and low complexity of fragmentation. Secondly, we expected higher selectivity of detection, because the mass spectral fragmentation pattern of MTBSTFA derivatives generates typical $[M-57]^+$ and $[M-15]^+$ fragments from, in most cases, still detectable molecular ions. Finally, we took into account the purity of the reagent and the stability of derivatives (Corey et al. 1972).

Table I-2:
Molar response ratios of all observed derivatives generated from 0.083 µg of each reference substance[a].
Molar response ratios were calculated from quadrupole EI-MS total ion currents of a single experiment (n = 3) by normalisation to the signal of an equal amount of 5α-cholestane within each preparation. Mass spectra of all tert.-butyldimethylsilyl (TBS) and trimethylsilyl (TMS) derivatives mentioned in the table are available via download of the mass spectral library included as supplementary data. The number of silyl-groups (#) is listed, brackets indicate an estimated number.

Reference Substance[a]			**TBS - Derivatives**			**TMS - Derivatives**				
	Molecular Mass ($g\ mol^{-1}$)	Amount[c] (nmol)		#	MTBSTFA		#	BSTFA	MSTFA	MSHFBA
ACC	101.1	0.82	ACC TBS 1	2	0.688	ACC TMS 1	2	0.421	0.380	0.460
IAA	175.2	0.47	IAA TBS 1	2	0.005	IAA TMS 1	2	1.152	1.060	1.058
			IAA TBS 2	1	1.369	IAA TMS 2	1	0.140	0.002	0.001
JA	210.3	0.39	JA TBS 1	1	0.675	JA TMS 1	1	0.436	0.293	0.408
			JA TBS 2	2	0.002	JA TMS 2	2	0.002	0.163	0.010
SA	138.1	1.20[b]	SA TBS 1	2	1.464	SA TMS 1	2	0.613	0.473	0.699
ABA	264.3	0.31	ABA TBS 1	1	1.229	ABA TMS 1	1	0.834	0.746	0.693
			ABA TBS 2	2	0.050	ABA TMS 2	2	0.001	0.087	0.018
ABA-GE	426.5	0.19		-	-	ABA-GE TMS 1	(4)	0.023	0.008	0.015
mT	241.4	0.34	mT TBS 1	2	1.421	mT TMS 1	2	0.595	0.620	0.619
				-	-	mT TMS 2	1	0.030	0.039	0.025
				-	-	mT TMS 3	3	0.005	0.017	0.018
Z	219.2	0.38	Z TBS 1	2	1.856	Z TMS 1	2	1.058	1.057	0.966
			Z TBS 2	3	0.007	Z TMS 2	3	0.050	0.079	0.077
ZR	351.4	0.24	ZR TBS 1	(3)	0.016	ZR TMS 1	4	1.548	1.494	1.561
GA3	346.4	0.24	GA3 TBS 1	1	0.100	GA3 TMS 1	3	0.999	1.066	0.950
				-	-	GA3 TMS 2	3	0.081	0.077	0.057
BL	480.8	0.17		-	-		-	-	-	-
TRP	204.2	0.41	TRP TBS 1	3	0.005	TRP TMS 1	3	0.329	1.175	1.270
			TRP TBS 2	2	1.254	TRP TMS 2	2	0.053	0.191	0.083
				-	-	TRP TMS 3	1	0.030	< 0.001	< 0.001
				-	-	TRP TMS 4	2	0.160	< 0.001	< 0.001
INO	180.2	0.46		-	-	INO TMS 1	6	0.009	0.137	0.068

[a] ACC, 1-aminocyclopropane-1-carboxylic acid; IAA, indole-3-acetic acid; JA, (±)-jasmonic acid; SA, salicylic acid; ABA, (±)-abscisic acid; ABA-GE, (±)-abscisic acid-β-D-glucopyranosyl ester; mT, meta-topolin; Z, trans-zeatin; ZR, trans-zeatin riboside; GA3, gibberellic acid A3; BL, 24-epibrassinolide; TRP, DL-tryptophan; INO, myo-inositol
[b] 0.166 µg per preparation
[c] Amount per analysis

The drawback of MTBSTFA derivatisation, however, is low sensitivity for gibberellins and monosaccharide conjugates. This deficiency may be circumvented by splitting extracts and alternative analysis with trimethylsilylation. Trimethylsilylation has a more comprehensive potential (Fiehn et al. 2000b), but showed interference with residual water. Moreover, mass spectral fragmentation patterns were clearly more complex and less specific. Therefore, a higher demand on pre-purification and concentration from plant matrices and thus lower overall sensitivity was expected.

Optimisation of the MTBSTFA protocol

Repeatability of GC-MS analysis

The MTBSTFA protocol was optimised using a quadrupole GC-EI-MS system because of the higher sensitivity in the EI-MS mode and comparative ease of handling and data processing.

For this purpose we performed experiments with a reference mixture of ACC, IAA, JA, SA, ABA, mT, and Z and varying internal standard substances (n = 9) in the course of 10 h. The molar response ratios of the derivatives using our initial choice of the 5α-cholestane standard had 6.0 – 13.1% RSD (Table I-3). In our hands this level of repeatability is typical for GC-MS analysis, when MTBSTFA or MSTFA are used as solvents for injection.

In an attempt to assess improvement of GC-MS reproducibility, we tested internal standardisation by n-nonadecane, DL-α-tocopheryl acetate, and each of the components of the pesticide standard mixtures 8081 and EPA 508/508.1. The pesticide mixtures allowed the fast screening of a large range of different compound classes which were in part also derivatised by MTBSTFA. None of the tested compounds exhibiting higher as well as lower boiling points qualified for a better internal standard of any of the reference derivatives than 5α-cholestane. Therefore, we continued internal standardisation with this compound.

Table I-3:
Repeatability of the molar response ratio of the main derivative synthesised from 0.083 µg of each reference substance.
Molar response ratios were calculated from quadrupole EI-MS total ion currents (n = 9) set to a scanning range of m/z = 40 – 600 using the signal of 5α-cholestane within each preparation for normalisation.

Reference Substance			**Total Ion Response**	
	Amount (nmol) per analysis	Derivative	Molar ResponseRatio AVG	RSD
1-aminocyclopropane-1-carboxylic acid	0.82	ACC TBS 1	1.004	6.0
indole-3-acetic acid	0.47	IAA TBS 2	1.281	13.1
(±)-jasmonic acid	0.39	JA TBS 1	0.857	6.2
salicylic acid	0.60	SA TBS 1	0.920	6.8
(±)-abscisic acid	0.31	ABA TBS 1	1.438	8.7
meta-topolin	0.34	mT TBS 1	1.621	11.8
trans-zeatin	0.38	Z TBS 1	2.040	13.0
DL-tryptophan	0.41	TRP TBS 2	1.300	10.7
5α-cholestane	0.48	-	-	-

Optimisation of reaction conditions

Figure I-1 summarises the effects of permuted incubation time and temperature on relative yield. The relative yield was calculated separately for the main derivatives of each phytohormone. The maximum molar response ratio obtained, respectively, was set to 100% relative yield. Side products did not accumulate under any of the conditions tested. Derivatives were grouped according to similarity of behaviour.

These groups were related to the GC elution sequence of derivatives. Cytokinins exhibited almost identical behaviour; IAA and ABA were highly similar, whereas JA, SA, and ACC

showed a similar tendency. In general, incubation temperature exhibited a stronger influence on reaction yield than incubation time. The optimum compromise for all phytohormones was 1 h incubation at 100°C.

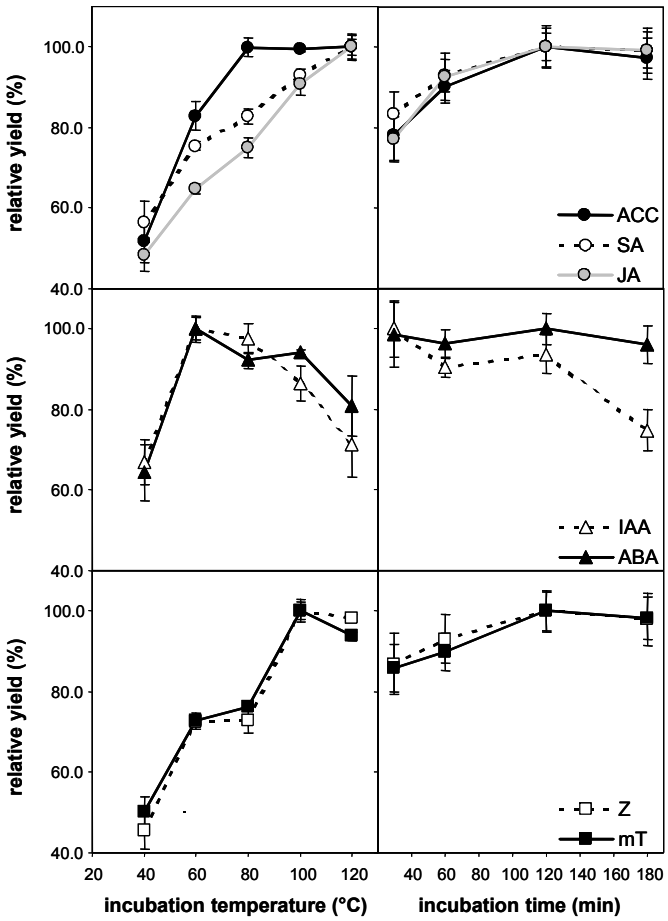

Figure I-1:

Optimisation of the MTBSTFA reaction protocol.
Relative yield was defined as % of maximum molar response ratio of each main derivative when monitored by quadrupole GC-EI-MS in total ion monitoring mode. Amounts per analysis were 2.5 µg of reference substance and 1.5 µg of 5α-cholestane. A, permutation of incubation temperature (n = 7); B, permutation of incubation time (n = 7).
ACC, 1-aminocyclopropane-1-carboxylic acid; IAA, indole-3-acetic acid; JA, (±)-jasmonic acid; SA, salicylic acid; ABA, (±)-abscisic acid; mT, meta-topolin; Z, trans-zeatin

Stability and Storage

An experiment in factorial 2^3 design was performed to detect parameter which might influence robustness of analysis after final derivatisation. Two levels of three factors were investigated by 24 experiments comprising three replications of each possible combination of factors. The stability of the derivative was tested for a typical time of analysis, ≤ 1 day, as compared to storage for 3 – 14 days. Influence of residual water, a common problem in phytohormone preparations, was checked by addition of 2% (v/v) of water before reaction with MTBSTFA. Storage temperature was the third parameter tested. Typically samples are exposed to room temperature before GC analysis. Therefore, we compared storage at -20°C with the exposure to room temperature.

A three-way ANOVA was performed for each phytohormone. IAA, SA, and ABA analysis was not influenced by any of the tested challenges to robustness. Each derivative was stable in the presence of trace amounts of water. All relative yields were 90 – 100%. Storage time increased relative yield of ACC by a factor of 1.5 ($p < 0.001$) and relative yield of JA by a factor of 2.0 ($p = 0.004$). In contrast, relative yield of mT was reduced by a factor of 0.65 ($p < 0.001$). This finding may be indicative of incomplete derivatisation in the case of ACC and JA and shows slight long-term instability of the mT derivative. Lowering storage temperature to -20°C appeared not to be beneficial. In contrast, the relative yield of Z was reduced by a factor of 0.9 ($p = 0.006$). This effect increased in combination with long storage time ($p < 0.001$). No other interaction of factors was detectable. ANOVA clearly demonstrated general robustness of the selected protocol. Furthermore, we checked effects on side product formation and found their relative occurrence to be invariant.

Calibration and limits of detection

Quadrupole GC-EI-MS in selected ion monitoring mode was used for analysis of detection limits and respective signal to noise ratios (S/N). Only 1/25 of final the sample volume was analysed by GC-MS with splitless injection. The noise value was determined as the maximum amplitude of the background signal in a range of ±5 times the respective peak width. Peak height was calculated from the average noise level to peak apex. The detection limits are presented as amount required before derivatisation with MTBSTFA reagent (Tab I-4).

Table I-4:
Detection limits of phytohormones expressed as amount required before derivatisation by MTBSTFA[a].
Quadrupole GC-EI-MS in selected ion monitoring. Signal to noise at the limit of detection was calculated based on the maximum amplitude of the background signal in the vicinity of the respective peak.

Phytohormone	Selected Ion Monitoring					
	Derivative	Fragment (m/z)	Relative Abundance %[c]	Detection Limit[b] (ng)	Detection Limit[b] (pmol)	S/N
ACC	ACC TBS 1	272	7.3	0.50	5.0	4:1
IAA	IAA TBS 2	232	19.4	0.05	0.3	5:1
JA	JA TBS 1	133	7.0	5.00	24.0	5:1
SA	SA TBS 1	309	18.5	0.05	0.4	15:1
ABA	ABA TBS 1	190	6.0	2.00	7.5	4:1
MT	mT TBS 1	469	8.9	2.00	8.3	5:1
Z	Z TBS 1	302	7.9	5.00	23.0	4:1

[a] ACC, 1-aminocyclopropane-1-carboxylic acid; IAA, indole-3-acetic acid; JA, (±)-jasmonic acid; SA, salicylic acid; ABA, (±)-abscisic acid; mT, meta-topolin; Z, trans-zeatin
[b] Total derivatisation volume was 25 µl. 1 µl was applied to GC-MS analysis.
[c] The relative abundance of selected ion fragments was determined in-parallel by total ion monitoring experiments with m/z = 40 – 600.

Sensitivity varied within two orders of magnitude among the different compounds analysed. Detection of IAA and SA was highly sensitive, 0.3 and 0.4 pmol, respectively, whereas JA and Z exhibited detection limits of 24 pmol and 23 pmol per sample. ACC, ABA, and mT had intermediate detection limits of 5.0, 7.5, and 8.3 pmol. The fragments chosen for selected ion monitoring had relative abundances of 7 - 19% and were mostly in the high-mass range. Both, high relative abundance and high mass of available fragments, contributed to the considerable sensitivity of tert.-butyldimethylsilyl derivatives as compared to TMS derivatives.

Calibration curves of the phytohormones are presented in double logarithmic scale (Figure I-2). The smallest amount shown corresponds to the respective detection limit (Table I-4). In the sub-pmol to pmol range the reference substances, ACC, IAA, SA, and JA all exhibited clear linear behaviour. The response functions of ABA, mT, and Z were sigmoid. Calibration curves that were extended into the nmol range all had a sigmoid shape and indicated upper detection limits of 10 - 100 nmol injected.

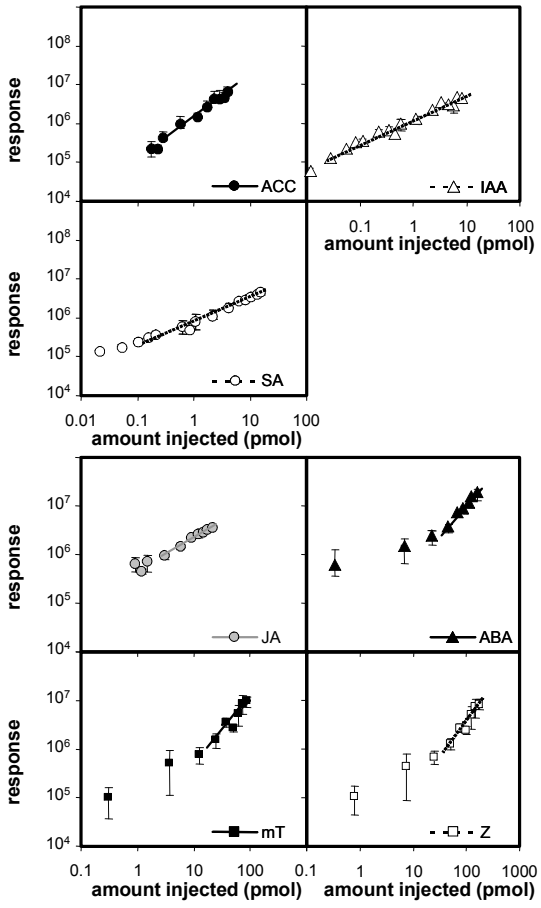

Figure I-2:

Calibration curves of ACC, SA, JA, IAA, ABA, mT and Z for the optimised derivatisation protocol.
Calibration curves (n = 6) demonstrate the working range of the MTBSTFA derivatisation protocol of 1 h incubation at 100°C in a volume of 25 µl (quadrupole GC-EI-MS in selected ion monitoring mode). The smallest amounts shown represent the detection limits. Fragments and the signal to noise ratios at the detection limits were as listed in Table I-4. Error bars represent the RSD.
ACC, 1-aminocyclopropane-1-carboxylic acid; IAA, indole-3-acetic acid; JA, (±)-jasmonic acid; SA, salicylic acid; ABA, (±)-abscisic acid; mT, meta-topolin; Z, trans-zeatin

Phytohormone profiling

The MTBSTFA protocol was successfully applied to analysis of extracts from tobacco root (Figure I-3) and seedlings of *A. thaliana*. A published extraction and purification method (Mueller et al. 2002) was adopted and phytohormone fractions prepared accordingly from representative samples of approximately 0.3 g fresh weight.

Analysis with quadrupole GC-EI-MS in selective ion monitoring mode exhibited in part intense peaks which were identified by spiking experiments. Moreover, each phytohormone was monitored using four different fragments in four consecutive runs. These experiments and total ion scanning analysis indicated inadequate sample purity for routine analysis with quadrupole GC-EI-MS. For the unequivocal demonstration of the presence of endogenous phytohormones we employed ion trap GC-CI-MS-MS. Analysis of MS-MS spectra allowed identification of ACC, IAA, SA, JA, and ABA from *A. thaliana* seedlings (data not shown). In shoot organs of *A. thaliana* IAA and JA are typically found in concentrations of 100 - 1000 pmol g^{-1} FW, whereas SA and ABA mostly range from 10-100 pmol g^{-1} FW (Mueller et al 2002). Successful identification of IAA, SA, JA and ABA was therefore in agreement with expectations.

ACC was previously not noticed in similar preparations (Figure I-3). We did not try to monitor Z in this experiment because the Z concentration reported to occur in tobacco seedlings did not exceed 0.25 pmol g^{-1} FW (Werner et al. 2001).

Conclusions

We present a novel method appropriate for comprehensive chemical derivatisation and subsequent gas chromatography-mass spectrometry of phytohormones. However, current means of joint extraction and preparation of phytohormone fractions from plant samples restrict the potential of our analysis to currently five endogenous target compounds: ACC, IAA, SA, JA, and ABA. Future efforts will focus on extending the preparation protocol for full exploitation of our method. Currently, ion trap GC-CI-MS-MS analysis is mandatory and quantitative standardisation is restricted to the use of stable-isotope labelled reference standards.

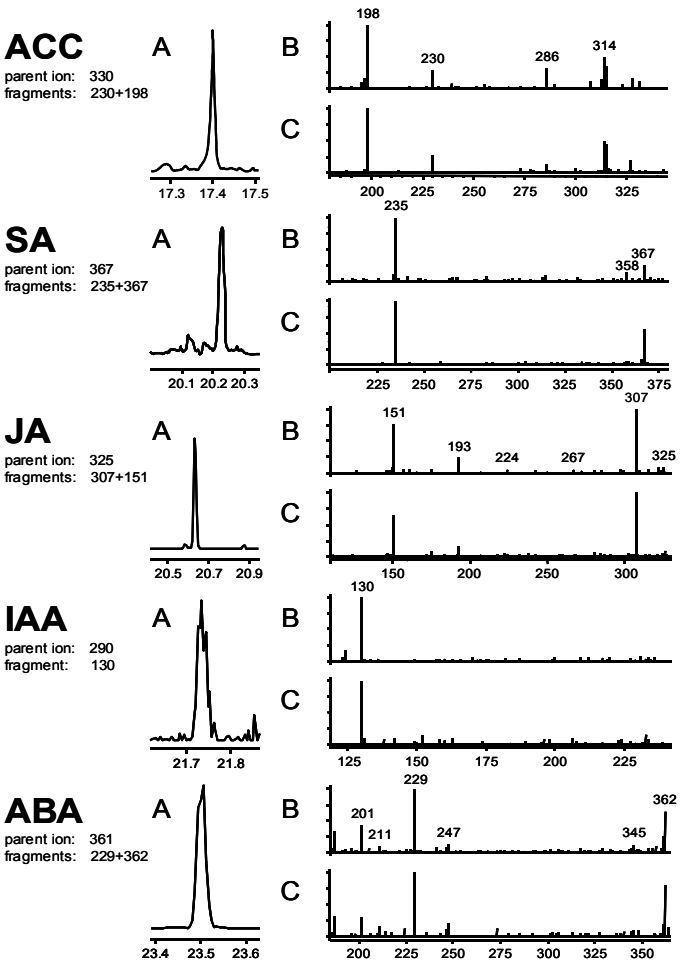

Figure I-3: MTBSTFA phytohormone profile of 0.3 g tobacco root
MTBSTFA phytohormone profile of 0.3 g tobacco (*Nicotiana tabacum*) root on ion trap GC system in positive CI-MS-MS mode with methanol as reactant gas.
Panels A show the specific MS-MS fragment traces. MS-MS fragmentation of reference substances, panels B, is compared to the MS-MS spectra of endogenous plant compounds, panels C. MS-MS spectra were taken from the peak apexes. ACC, 1-aminocyclopropane-1-carboxylic acid; IAA, indole-3-acetic acid; JA, (±)-jasmonic acid; SA, salicylic acid; ABA, (±)-abscisic acid.

Acknowledgements

The authors thank all members of the Max-Planck-Institute of Molecular Plant Physiology, Golm, Germany, especially Cornelia Wagner for patient support in performing GC-MS measurements, and Dr. Alisdair Fernie and Prof. Lothar Willmitzer for fruitful discussions and close scrutiny of our work.

Our special gratitude belongs to Dr. Hilde Boiten, Dr. Els Prinsen (University of Antwerp, Belgium), Dr. Axel Müller, and Prof. Elmar W. Weiler (University of Bochum, Germany) for their immediate readiness to give insight into their methods of phytohormone analysis and further invaluable support.

Chapter II - Robustness of Metabolite Extraction from Plant Tissue

Design of Metabolite Recovery by Variations of the Metabolite Profiling Protocol

Proceedings of the Third International Congress on Plant Metabolomics in 2004, Ames (Iowa, USA); Kluwer Academic Press 2005

Claudia Birkemeyer, Joachim Kopka

Abstract

More than 670 GC-MS metabolite profiles were performed in the course of three years in an effort to probe robustness and reproducibility of metabolite profiling and to design metabolite recovery and the range of metabolite classes which are finally analysed with GC-MS based metabolite profiling. Experiments were performed with two highly important plant organs, namely root and leaf, using the model plant tobacco, *Nicotiana tabacum L.* var. Samsun NN (SNN). We investigated solvent composition, pH and temperature during metabolite extraction and subsequent liquid partitioning of extracts. All permutations of the metabolite profiling protocol were directly compared to the initially published standard protocol. In agreement with the fundamental approach of profiling analyses, results were reported relative to this standard condition. Thus the consistency of results was maintained in the course of years. The resulting set of chromatograms was screened for mass spectral tags (MSTs), which represent identified as well as still unidentified metabolites. An unsupervised mass spectral and retention-time-index library (MSRI_NS) of these MSTs was constructed for the future discovery of hitherto unidentified metabolic components and will be made available at the Golm Metabolome Database GMD (http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html). Cluster analyses and multivariate statistical techniques were applied to obtain insight into general trends of protocol variants. A choice of representative metabolites was analysed in-depth for potential analytical improvement and final protocol optimisation.

Introduction

The essence of metabolite profiling is the screening of biological samples for changes of metabolite levels relative to reference samples (Fiehn et al. 2000b; Sumner et al. 2003). The use of reference samples in biological experimentation and throughout relative quantitative analysis allows control of arbitrary changes in apparent metabolite levels. This approach of profiling analyses is essential, because unwanted influences can not always be avoided. For example, experimental investigations of biological systems often cannot exactly be reproduced in all aspects, and artefacts of the chemical analysis are common experience.

In consequence, it can be argued that - provided the profiling approach is chosen - any type of comparative chemical analysis or experiment may be performed, if (i) reference samples are included and if (ii) these references and the samples under investigations are treated identically throughout the complete analytical process. Thus, only a small number of in-depth analyses aimed to optimise metabolite recovery have been performed so far (e.g. Fiehn et al. 2000b; Roessner et al. 2000; Roessner-Tunali et al. 2003; Maharjan and Ferenci 2003; Gullberg et al. 2004). However, metabolite profiling may be described as the art of making as many metabolites as possible amenable to simultaneous analysis. This aspect is one of the key aspects for improved metabolite profiling analyses. Two strategies allow extension and modulation of the scope of GC-MS based metabolite profiling: (i) the choice of chemical derivatisation and (ii) the method of metabolite preparation.

In the present report we focussed on chemical derivatisation by MSTFA after methoxamination (Roessner et al. 2000) only and, furthermore, kept subsequent GC-MS analysis invariable. As a consequence, the crucial factor for the modulation of chemical metabolite classes and the dynamic range of metabolite concentrations was the selected extraction and pre-fractionation protocol.

To further our understanding of possible, more efficient variants of profiling analyses we investigated temperature, pH, solvent composition and liquid partitioning. We first generated an inventory of identified metabolites which were present under the respective profiling regimes and constructed a mass spectral and retention time index library for future qualitative investigations of hitherto unidentified components (Wagner et al. 2003; Kopka et al. 2005; Schauer et al. 2005).

Then, we selected key metabolites, which represented the predominant metabolite classes covered by routine GC-MS based metabolite profiling and performed exemplary analyses on metabolite recovery and robustness of analysis. In agreement with the general principle of

metabolite profiling, all investigations were performed in direct comparison with standard, or in other words, reference samples. These reference samples were defined to be samples, which were analysed in parallel using the initially published protocol of GC-MS based metabolite profiling (Roessner et al. 2000; Fiehn et al. 2000b).

Experimental

Plant Material

Nicotiana tabacum L. var. Samsun NN (SNN) plants were cultivated on quartz sand under 16h/ 8h long day conditions as described previously (Birkemeyer et al. 2003). Plant organs of 12 simultaneously grown plants were harvested three months after germination. All non-senescent leafs were immediately shock-frozen in liquid nitrogen. The root systems were washed sand-free under tap water, shortly dried on filter paper, and subsequently frozen in liquid nitrogen. The complete root and leaf batches of 12 plants were homogenised separately in pre-cooled mortars while permanently being kept under liquid nitrogen. Aliquots of ~50 mg each were prepared and stored for up to three years at -80°C until further processing.

Extraction

The general extraction scheme started with ball-mill homogenisation of the deep-frozen plant material in 2 mL micro vials as described earlier (Fiehn et al. 2000a; Fiehn et al. 2000b; Roessner et al. 2000). Non-sample controls were entered into analysis at this step of the protocol. Extraction was performed in three steps:

(i) The first extraction step was initiated by adding a 300 µL volume of pre-cooled polar water miscible solvent, which contained the internal standard substances, ribitol, D(-)-isoascorbic acid for monitoring the recovery of GC-separated, oxidation-sensitive vitamin C, L(+)-ascorbic acid, and 2,3,3,3-D4 alanine, to the deep-frozen powder without removal of the steel ball.

(ii) After initial incubation a 200 µL volume of chloroform was added and shortly incubated 5 min at 37°C.

(iii) Finally, a polar and a lipid phase was separated by addition of 400 µL of bi-distilled water, 1 min vigorous mixing, and 10 min centrifugation at room temperature in a micro vial

centrifuge set to maximum speed. Cellular debris accumulated mainly below the bottom chloroform layer and to a small extent at the interphase boundary. 80, 120, or 160 µL aliquots were carefully drawn from the top polar and the bottom lipid layers of each experiment.

The different aliquot volumes, which were dried by Speedvac vacuum concentration before chemical derivatisation, were adjusted to obtain a reproducible number and intensity of analyte peaks by introducing equal amounts of fresh weight 'aliquots' to GC-MS analysis irrespective of the particular sample preparation protocol. The total volume of each liquid fraction from all protocol variations was calculated.

Four protocol variations were performed. In the following we use a four-character code to address respective variants and 'L_' or 'R_' prefix to distinguish between leaf and root matrix, respectively. Each position of the code characterises an altered parameter. The codes are explained in Table II-1.

(i) We modified the major polar solvent of the first extraction step from pure methanol to acetone : water (2:1, v:v), first position code characters 'm' (methanol) and 'a' (acetone).

(ii) The pH of the first solvent was not adjusted or either acidified by 1.67% formic acid or adjusted to basic pH by saturation with solid sodium carbonate, second position code characters, 'n' (neutral), 'a' (acid), and 'b' (basic).

(iii) Incubation of the first extraction step was either hot, 15 min at 70°C, or cold, over night at -20°C, third position code character, 'h' (hot) or 'c' (cold).

(iv) Extraction was performed including a liquid/liquid partitioning into a polar and lipid phase, fourth position code characters, 'p' (polar) and 'l' (lipid), or phase separation was omitted, missing the last code character. The preparations without liquid/liquid partitioning are in the following called 'combined'. The reference protocol was defined as the polar fraction of the hot methanol extraction without pH adjustment, protocol code, mnhp.

Chemical Derivatisation

Dried extracts were derivatised in two consecutive steps as described earlier (Wagner et al. 2003). Carbonyl moieties were converted into methoxyamine groups (MEOX) by 90 min incubation at 30°C with 40 µL of freshly prepared methoxyamine hydrochloride (Sigma, Munich, Germany) solution in pure pyridine (20 mg/ mL) (Merck, Darmstadt, Germany). Subsequently acidic protons were substituted by trimethylsilyl-groups (TMS) using N-Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA, Macherey & Nagel, Düren,

Germany). Silylation was performed by adding a 70 µL volume of MSTFA followed by 30 min incubation at 37°C. Samples were subsequently kept at room temperature on the GC-MS injector tray for 2 - 12 h.

Retention-time-index standard mixture of n-alkanes was added in a 10 µL volume prior to silylation. n-Alkanes were dissolved in pyridine at a final concentration of 0.22 mg/ mL each: n-dodecane (RI 1200; CAS 112-40-3), n-pentadecane (RI 1500; CAS 629-62-9), n-octadecane (RI 1800; CAS 593-45-3), n-nonadecane (RI 1900; CAS 629-92-5), n-docosane (RI 2200; CAS 629-97-0), n-octacosane (RI 2800; CAS 630-02-4), n-dotriacontane (RI 3200; CAS 544-85-4), n-hexatriacontane (RI 3600; CAS 630-06-8). All above substances were obtained from Sigma, Munich, Germany, if not indicated otherwise.

GC-MS Analysis

A MD 800 quadrupole GC-MS system (ThermoQuest, Manchester, UK) was equipped with a RTX-5Sil MS capillary column, 30 m length, 0.25 mm inner diameter, 0.25 µm film thickness and a 10 m IntegraGuard pre-column (Restek GmbH, Bad Homburg, Germany). The system was operated in constant flow mode with 1 mL min^{-1} Helium 5.0 carrier gas (Air Liquide, Magdeburg, Germany).

GC-MS analysis was performed essentially as described earlier (Fiehn et al. 2000a; Fiehn et al. 2000b; Roessner et al. 2000; Roessner-Tunali et al. 2003; Colebatch et al. 2004). Injection was 1 µL in splitless mode at 230°C with a 2 min valve closure. The temperature program comprised 1 min isothermal time at 70°C, a 6 min ramp to 76°C, a 45 min ramp to 350°C, 1 min isothermal at 350°C, and further isothermal heating for 10 min at 330°C. The quadrupole mass selective detector was operated with Electron Impact ionisation. The transfer line was set to 250°C, and the ion source operated at 200°C. Scan rate was 2 spectra s^{-1}, with the m/z range set to 40 - 600.

Mass Spectral Tags for Analysis of Metabolite Recovery

In GC-MS metabolite profiling analyses metabolites are represented by mass spectra of metabolite derivatives which occur in highly reproducible retention time index windows. Thus, in a more precise phrasing, the chemical derivative and not the genuine, endogenous, metabolic chemical is quantified by GC-MS. Exceptions are those metabolites, which are

volatile and not susceptible to derivatisation, for example nicotine. The majority of analytes in GC-MS profiles of plants is still not identified. For these cases we created the expression 'mass spectral tag' or 'MST' (Colebatch et al. 2004). MSTs are defined by mass spectrum and retention behaviour and can thus be reproducibly deconvoluted and identified without knowledge about the chemical nature of the represented metabolite. Each analyte or MST can be quantified by fragments which constitute the respective full mass spectrum of MSTs. In Electron Impact ionisation analyses the relative intensities of fragments from a single compound are constant and independent of concentration, except if mass detectors are operated beyond the linear range of detection. We analysed each analyte and MST using multiple fragment masses as noted in Table II-2.

Response Calculations

Peak areas, X, of selected ion traces (Table II-2) were retrieved using the find algorithm of the MassLab software 1.4 (ThermoQuest, Manchester, UK). For each metabolite a single, specific and selective fragment mass was selected. Correct peak integration was monitored manually. Peak areas with low intensity were rejected. The resulting peak area values were defined to be 'fragment responses' (X_i of fragment i). Fragment responses were normalised to the fresh weight of the sample. In this investigation we did not use the internal standard substance, ribitol, for volume correction. Instead, we determined the final volume of each extract variation and performed a numerical correction to ensure analysis of equal fresh weight equivalents from all extraction regimes, resulting in relative response values N_i ($N_i = X_i$ * extract volume^{-1} * fresh weight^{-1}).

In a further step, the relative response of a fragment, N_i, is divided by the averaged relative response of the same fragment that was analysed with the initially developed standard protocol, mnhp, as a reference sample ($R_i = N_i$ * $avgN_{i(mnhp)}^{-1}$). The resulting quotient is subsequently called response ratio R_i. R_i describes the x-fold change in metabolite recovery relative to the standard extraction protocol. Response ratios were calculated for root and leaf samples separately. As a consequence, leaf and root mnhp overlapped in PCA analyses and indicate the origin of plots (Figure II-1 and II-4).

Each protocol variation, v, was analysed in 5-20 fold replication (Table II-1). In order to reduce the complexity of the data set we averaged the response ratios for the replicate analyses of each protocol, $avgR_{i,v}$. Fragments that exhibited a high relative standard deviation (RSD > 35%) when applying the standard protocol, mnhp, were excluded from further

analysis. We observed an increased RSD of all metabolites among replicate experiments performed in the course of the three year storage period of our sample batches. For this reason we had to accept a 35% threshold (Gullberg et al. 2004) rather than the reported 10% average analytical RSD of all analytes for repeated injections of GC-TOF-MS profiles (Weckwerth et al. 2004). Some fragments exhibited low RSD but were frequently missed by the MassLab find algorithm. For these cases we required successful peak finding in more than 50% of the replicate standard analyses.

Finally, all $avgR_{i,v}$ were combined into a single matrix that describes the complete set of metabolite recoveries under the different extraction regimes employed in this experiment. The majority of metabolites which exclusively occurred in only one plant matrix or only under specific conditions, except for dehydroascorbic acid, were excluded from the present version of our comparative investigation. The currently available set of analysed metabolites is reported in Table II-2.

Principal component analysis using the covariance model (PCA) and hierarchical cluster analysis using complete linkage of a Euclidian distance matrix (HCA) was performed without log10 transformation of the $avgR_{i(\text{protocol variation})}$ matrix. The S-Plus 2000 software package standard edition release 3 (Insightful, Berlin Germany) was used for HCA, PCA and visualisation.

Mass-Spectral Retention-Time-Index Libraries

Mass spectral metabolite tags (MSTs) were generated by automated deconvolution of the GC-MS chromatograms using the public available deconvolution-software package AMDIS (http://chemdata.nist.gov/mass-spc/amdis/; National Institute of Standards and Technology, Gaithersburg, USA) (Stein 1999). Mass spectra were collected with component width of 20, adjacent peak subtraction set to two, low resolution and shape requirement, and medium sensitivity.

The currently processed chromatograms (‘library entries’) and resulting number of MSTs are overviewed in Table II-1. RIs of all MSTs were determined and thus annotated MSTs were uploaded into an unsupervised custom-made NIST02 mass-spectral library (NIST02 mass spectral search program, http://chemdata.nist.gov/mass-spc/Srch_v1.7/index.html; National Institute of Standards and Technology, Gaithersburg, USA) (Ausloos et al. 1999).

Table II-1:
Summary of the mass spectral library for the protocol variations
Summary of protocol variations, specification of protocol codes, number of experiments, number of replicas that entered a unsupervised mass spectral and retention-time-index library (library entries), and number of mass spectral tags (MSTs) per respective protocol variation, generated by automated peak deconvolution using AMDIS software

Protocol variants					***Nicotiana tabacum* leaf**			***Nicotiana tabacum* root**			**non-sample control**		
Code	Solvent	pH	Tempe-rature	Fraction	Experi-ments	Library entries	MSTs/ entry	Experi-ments	Library entries	MSTs/ entry	Experi-ments	Library entries	MSTs/ entry
					309	16	9112	368	16	5796	48	8	1466
aacl	acetone	acidic	cold	lipid	7	-	-	7	-	-	-	-	-
aacp	acetone	acidic	cold	polar	10	-	-	10	-	-	-	-	-
abcl	acetone	basic	cold	lipid	7	2	496	7	2	247	4	1	152
abcp	acetone	basic	cold	polar	10	2	599	10	2	388	4	1	163
ancl	acetone	non-adjusted	cold	lipid	7	2	551	7	2	243	4	1	159
anc	acetone	non-adjusted	cold	mixed	8	-	-	10	-	-	-	-	-
ancp	acetone	non-adjusted	cold	polar	10	2	597	10	2	432	4	1	210
aahl	acetone	acidic	hot	lipid	5	-	-	13	-	-	-	-	-
aahp	acetone	acidic	hot	polar	10	-	-	20	-	-	-	-	-
abhl	acetone	basic	hot	lipid	4	-	-	17	-	-	-	-	-
abhp	acetone	basic	hot	polar	10	-	-	17	-	-	-	-	-
anhl	acetone	non-adjusted	hot	lipid	6	-	-	7	-	-	-	-	-
anh	acetone	non-adjusted	hot	mixed	8	-	-	10	-	-	-	-	-
anhp	acetone	non-adjusted	hot	polar	11	-	-	10	-	-	-	-	-
macl	methanol	acidic	cold	lipid	10	-	-	7	-	-	-	-	-
macp	methanol	acidic	cold	polar	10	-	-	10	-	-	-	-	-
mbcl	methanol	basic	cold	lipid	5	2	573	5	2	372	4	1	199
mbcp	methanol	basic	cold	polar	10	2	511	20	2	379	4	1	204
mncl	methanol	non-adjusted	cold	lipid	9	-	-	5	-	-	-	-	-
mnc	methanol	non-adjusted	cold	mixed	8	-	-	10	-	-	-	-	-
mncp	methanol	non-adjusted	cold	polar	11	-	-	19	-	-	-	-	-
mahl	methanol	acidic	hot	lipid	10	-	-	5	-	-	-	-	-
mahp	methanol	acidic	hot	polar	10	-	-	10	-	-	-	-	-
mbhl	methanol	basic	hot	lipid	5	-	-	5	-	-	-	-	-
mbhp	methanol	basic	hot	polar	10	-	-	10	-	-	4	1	167
mnhl	methanol	non-adjusted	hot	lipid	20	2	614	7	2	410	-	-	-
mnh	methanol	non-adjusted	hot	mixed	8	-	-	10	-	-	-	-	-
mnhp[a]	methanol	non-adjusted	hot	polar	70	2	617	90	2	429	20	1	212

[a]standard protocol

The approach of constructing unsupervised mass spectral libraries and applications of these libraries was described by Wagner et al. (2003).

Automated Mass Spectral Identification

All chromatograms processed for MST library construction were also screened for already known analytes using AMDIS software and our Q_MSRI_ID library. Results of these identifications are listed in Table II-2. The Q_MSRI_ID mass spectral library contains repeatedly observed MSTs and those analytes which were identified by standard addition experiments; it is available at http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html (Kopka et al. 2005). Identifications were automated using the AMDIS option to generate tab-separated report files about the deconvolution results. Thresholds for the acceptance of positive identifications were: signal to noise ratio (S/N) > 20, reverse mass spectral match value (Match) > 65 and retention time index (RI) deviation < 5.0.

Results and Discussion

Inventory of Analytes and MSTs

Automated and reliable mass spectral identification of deconvoluted mass spectra is one of the major challenges in GC-MS based metabolite profiling, not least because this process is prerequisite for the discovery of those MSTs and analytes, which have hitherto not been found in a specific type of biological sample, or have not yet been archived in mass spectral and retention-time-index libraries (MSRI). We applied the concept of unsupervised mass spectral libraries (MSRI_NS) (Wagner et al. 2003) and a library search of identified metabolites and frequently observed MSTs (MSRI_ID) (Kopka et al. 2005), which was developed in the course of recent years, to identify components found in different extracts of the model plant tobacco, *Nicotiana tabacum* L. var. Samsun NN (SNN). We manually selected 16 diverse and representative GC-MS chromatograms of polar and lipid fractions from non-pH-adjusted and basic extraction protocols including eight complementary non-sample controls (library entries in Table II-1). In total we obtained 16374 MSTs using AMDIS deconvolution. Non-sample controls had on average 183 MSTs per chromatogram, which comprised some laboratory contaminations, such as benzoic acid or lactic acid, and a range of short to long chain fatty acids (refer to Table II-2 for details), as well as a few unavoidable side products of the chemical reagents, that were used for GC-MS profiling.

Besides side products of the reagents, for example hydroxylamine, the majority of contaminations, approximately > 50%, belonged to the class of linear and cyclic polysiloxanes. These compounds accumulate over time in the MSTFA reagent, when exposed to traces of ambient air, but are also mobilised by MSTFA from GC-MS septum and crimp cap material. Capillary GC columns are an additional source of MSTFA mobilised polysiloxanes, especially toward the end of their life time, and may contain mixed methyl-phenyl- or methyl-aryl polysiloxanes.

When analysing the complexity of plant extracts, we found significant differences in retrieved numbers of MSTs. The following results were obtained after adjustment of major peaks to the upper detection limit. As a rule of thumb, these dominant components did not show peak deformation due to chromatographic overloading, but may in cases exhibit slight mass spectral distortions due to saturation of the quadrupole mass detector used in this investigation.

Table II-2:
List of 146 identified metabolites from leaf or root extracts
Identification was performed by automated comparison with the MSRI_ID library of identified mass spectra available at http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html using AMDIS software (MPIMP-ID; analyte identifier). Thresholds for accepting identifications were: total-ion-current signal to noise ratio (S/N) > 20; reverse mass spectral match (Match) > 65; and retention time index (RI) deviation < 5.0. Relative quantitative analysis of changes in metabolite recovery (RA that is recovery analysis) was performed on a subset of metabolites that were present under all tested protocol regimes. All suggested selective quantification masses (QM) for quantitative metabolite analysis are indicated.

Analyte				**RI**		**Match**		**S/N**	**QM**					**RA**
MPIMP-ID[h]	Metabolite	Derivative												
		TMS	MEOX	Expected	Deviation	Simple	Reverse	Ratio	m/z	m/z	m/z	m/z	m/z	
176002-101	Aconitic acid, cis-	3		1762.8	-0.2	33	84	40	229	285	375	211	215	X
188005-101	Adenine	2		1872.4	-0.1	44	88	110	264	279	192	165	237	-
151006-101	Adipic acid	2		1509.0	2.5	80	90	64	275	111	141	172	159	-
144001-101	Alanine, beta-	3		1431.4	0.5	86	93	131	248	290	174	160	100	X
138002-101	Alanine, DL-	3		1363.9	-0.2	96	99	52	188	262	290	100	114	X
138002-211	Alanine,DL-,2,3,3,3-D_4-[i]	3		1359.9	0.6	96	99	440	192	266	294	104	117	-
167002-101	Arabinose	4	1	1675.3	0.7	93	99	185	307	217	160	103	189	X
168001-101	Asparagine, DL-	3		1683.3	0.9	88	97	176	116	188	231	258	159	-
152002-101	Aspartic acid, DL-	3		1525.0	1.4	76	99	255	232	218	306	202	334	X
128003-101	Benzoic acid	1		1256.7	-1.6	65	95	109	179	105	135	77	194	X
164003-101	Benzoic acid, 4-hydroxy-	2		1639.5	-0.4	83	93	157	267	223	282	193	126	-
184001-101	Benzoic acid, p-amino-	2		1841.1	-0.4	40	887	26	266	281	222	192	126	-
117002-101	Butyric acid, 2-amino-, DL-	2		1169.9	2.5	67	89	28	130	204	218	232	142	-
153003-101	Butyric acid, 4-amino	3		1530.7	0.4	98	100	397	174	304	216	246	100	X
126002-101	Butyric acid, 4-hydroxy	2		1242.6	2.3	49	84	48	233	117	204	143	133	-
214001-101	Caffeic acid, trans-	3		2141.4	-0.8	81	95	157	396	381	219	307	205	X
311001-101	Caffeoylquinic acid, 3-trans-	6		3126.0	-2.3	76	98	1293	345	255	397	324	219	X
317001-101	Caffeoylquinic acid, 4-trans-	6		3188.8	-4.1	79	100	840	307	489	324	255	219	X
319001-101	Caffeoylquinic acid, 5-trans-	6		3209.8	-2.5	68	89	1407	307	447	345	255	219	X
329001-101	Campesterol	1		3262.1	0.4	97	99	506	472	382	343	367	129	-
319002-101	Cholesterol	1		3156.9	-1.6	97	98	415	458	368	353	329	129	-
148001-101	Citramalic acid, D(-)-	3		1473.8	-0.2	66	96	194	247	349	259	321	203	X
182004-101	Citric acid	4		1827.8	0.5	98	100	375	273	375	211	183	257	X
184011-101	Citric acid, 2-methyl-, DL-	4		1841.9	-1.6	43	67	66	287	479	389	197	225	-
194005-101	Coniferylalcohol, trans-	2		1945.7	0.0	84	97	64	324	293	235	309	219	-
195001-101	Coumaric acid, p-, trans-	2		1944.6	1.8	51	89	27	249	293	308	219	179	-
156002-101	Cysteine, DL-	3		1560.7	0.6	72	95	82	294	220	218	100	116	-
144002-101	Cysteine, S-methyl-, DL-	2		1427.3	0.0	54	77	36	162	218	236	100	115	-
147004-101	Decanoic acid, n-[a]	1		1462.1	2.3	67	89	81	229	244	117	201	145	-
185002-101	Dehydroascorbic acid, dimer[c]	+		1852.6	-1.1	79	86	121	316	173	157	245	231	X
166003-101	Dodecanoic acid, n-[a]	1		1662.5	2.2	90	99	96	257	272	117	201	145	-
245001-101	Eicosanoic acid, n-	1		2456.3	-0.2	72	96	186	369	384	117	201	145	-
150002-101	Erythritol	4		1510.2	1.8	91	98	221	217	293	307	205	320	X
154001-101	Erythronic acid	4		1548.7	1.6	97	100	605	292	220	117	319	205	-
146002-101	Erythrose	3	1	1459.1	2.5	28	75	24	205	117	161	233	262	-
128002-101	Ethanolamine	3		1269.1	0.7	70	97	52	174	86	100	188	262	-
210001-101	Ferulic acid, trans-	2		2098.6	2.4	33	89	29	338	249	323	293	308	-
187002-101	Fructose	5	1	1874.6	1.0	97	100	684	307	217	277	364	335	X
232002-101	Fructose-6-phosphate	6	1	2321.4	-0.6	80	99	161	459	315	357	217		X
175001-101	Fucose	4	1	1746.5	-1.4	89	92	77	117	160	364	277	321	X
137001-101	Fumaric acid	2		1359.8	1.8	92	99	333	245	115	217	143		X
299002-101	Galactinol	9		2993.5	-2.5	80	96	226	204	191	433	305	169	X
194001-101	Galactitol	6		1941.8	0.1	32	73	41	319	307	157	217	331	-
199002-101	Galactonic acid	6		1997.9	-0.3	95	99	404	333	292	319	305	157	X
189002-101	Galactono-1,4-lactone, DL-	4		1890.5	1.2	54	65	524	217	451	466	334	305	-
188001-101	Galactose	5	1	1892.3	0.1	96	99	534	160	319	229	343	305	-
194003-101	Galacturonic acid	5	1	1946.9	-2.5	34	77	37	333	160	423	292	364	X
283004-101	Gentiobiose	8		2828.3	-0.5	89	95	112	160	480	390	204	361	-
200001-101	Gluconic acid	6		2002.7	-1.5	88	99	334	333	292	319	305	157	-
189008-101	Gluconic acid-1,5-lactone	4		1887.9	1.2	68	86	68	220	229	319	451	129	-
190007-101	Glucopyranoside, 1-O-methyl-, beta-D-	4		1898.3	4.7	59	73	239	133	204	377	231	290	-
189002-101	Glucose	5	1	1897.3	0.4	96	100	594	160	319	229	343	305	X
172001-101	Glucose, 1,6-anhydro, beta-D-	3		1715.1	0.3	96	99	164	204	217	333	243	317	-
233002-101	Glucose--phosphate	6	1	2334.5	-1.4	92	99	120	160	387	299	471	357	X
193004-101	Glucuronic acid	5	1	1937.4	-1.6	85	98	288	333	160	423	292	364	X
163001-101	Glutamic acid, DL-	3		1631.4	0.6	89	99	1183	246	363	128	348	156	X
178001-101	Glutamine, DL-	3		1785.1	-0.4	96	99	482	156	245	347	362	203	-
143001-101	Glutaric acid	2		1414.6	1.6	53	87	37	158	261	233	116	186	-
158004-101	Glutaric acid, 2-oxo	2	1	1588.7	0.9	91	99	302	198	288	304	186	229	-
135003-101	Glyceric acid, DL-	3		1339.6	1.9	99	99	365	292	189	307	205	133	-
129003-101	Glycerol	3		1282.5	1.6	97	99	237	293	205	117	103		X
174002-101	Glycerol-2-phosphate	4		1741.3	1.2	52	77	47	243	299	389	211	445	-
177002-101	Glycerol-3-phosphate, DL-	4		1775.1	1.1	86	100	621	357	445	299	315	211	-
133001-101	Glycine	3		1311.9	0.0	91	99	216	174	248	276	100	86	X
214002-101	Guanine	4		1232.7	2.7	71	97	56	352	367	264	202	99	X
278001-101	Guanosine	5		2781.8	-0.3	81	91	76	324	245	280	368	410	-
196001-101	Gulonic acid	6		1964.2	-2.2	79	95	70	333	292	423	433	319	-
205001-101	Hexadecanoic acid, n-	1		2050.2	0.1	98	100	1041	313	328	117	201	145	X
106001-101	Hexanoic acid, n-[a]	1		1064.0	-2.3	70	88	69	173	188	117	129	145	-
146001-101	Homoserine, DL-	3		1454.0	1.8	52	89	38	218	128	292	230	202	X
209002-101	Inositol, myo-	6		2091.9	-1.8	83	100	756	305	265	318	191	507	X
243003-101	Inositol-phosphate, myo	7		2429.0	-2.2	69	99	199	299	318	387	315	217	X
182003-101	Isocitric acid	4		1831.6	-2.1	82	96	254	245	319	390	83		X
132002-101	Isoleucine, DL-	2		1300.6	0.8	75	100	91	158	232	218	102	260	X
291002-101	Isomaltose	8	1	2907.0	-3.8	67	78	52	160	480	204	319	361	-
135004-101	Itaconic acid, DL-	2		1351.7	2.0	65	87	37	259	215	133	147	230	-

105001-101	Lactic acid, DL-[a]	2		1048.9	-0.2	94	99	173	219	117	191	133	234	-
129002-101	Leucine, DL-	2		1278.8	0.9	54	100	294	158	232	102	260		-
192003-101	Lysine, DL-	4		1920.8	-0.1	86	98	325	156	174	317	230	434	-
133003-101	Maleic acid[a]	2		1314.7	0.9	96	100	490	245	147	170	215		X
137003-101	Maleic acid, 2-methyl-	2		1358.0	-2.5	58	83	31	259	184	122	157	231	-
149001-101	Malic acid, DL-	3		1492.3	1.1	97	99	495	233	245	335	307	217	X
122003-101	Malonic acid	2		1211.4	-0.1	96	99	333	233	248	147	133	109	-
284001-101	Maltitol	9		2839.2	-1.6	74	77	99	204	361	345	525	305	-
277002-101	Maltose	8	1	2768.4	-0.8	81	95	136	160	204	361	319	271	X
355003-101	Maltotriose	11	1	3550.2	-0.8	65	91	38	204	361	217	480	169	-
193002-101	Mannitol	6		1928.8	-2.0	94	99	161	319	307	157	217	331	X
189001-101	Mannose	5	1	1899.5	-2.2	43	79	1389	160	319	229	343	305	-
231001-101	Mannose-6-phosphate	6	1	2323.5	-0.4	35	76	24	160	471	387	357	299	X
346001-101	Melezitose	11		3475.7	-1.2	84	95	202	361	451	271	204	217	X
290002-101	Melibiose	8	1	2903.5	2.3	66	89	52	160	480	204	319	361	X
152001-101	Methionine, DL-	2		1521.2	0.6	95	99	128	176	128	250	293	202	-
259003-101	Nicotianamine	4		2606.3	-1.3	74	89	74	186	218	246	345	232	-
139002-101	Nicotine			1357.5	0.1	95	99	290	84	133	161	162	92	X
133004-101	Nicotinic acid	1		1303.7	1.7	83	95	85	180	136	106	78	195	-
138003-101	Nonanoic acid, n-[a]	1		1369.2	2.1	73	97	89	215	230	117	129	145	-
231003-101	Octadecadienoic acid	1		2315.1	0.4	78	85	108	337	352	262	129	117	-
221003-101	Octadecadienoic, 9,12-(Z,Z)-	1		2211.1	0.1	83	99	802	337	335	262	129	117	-
225002-101	Octadecanoic acid, n-	1		2247.0	-0.2	98	100	431	341	356	117	129	145	X
222003-101	Octadecatrienoic acid, 9,12,15-(Z,Z,Z)-	1		2219.1	-0.6	98	99	450	335	350	93	129	117	-
223003-101	Octadecenoic acid	1		2225.1	-0.4	56	65	108	339	354	117	129	145	-
222001-101	Octadecenoic acid, 9-(Z)-	1		2217.4	0.4	48	84	55	339	354	117	129	145	-
127006-101	Octanoic acid, n-[a]	1		1270.6	2.4	65	87	54	201	216	117	129	145	-
182002-101	Ornithine, DL-[d]	4		1821.9	-1.3	60	91	79	142	174	420	200	258	X
195004-101	Pentadecanoic acid, n-[a]	1		1949.0	0.0	48	95	92	299	314	117	201	145	-
164001-101	Phenylalanine, DL-	2		1635.4	-1.5	88	97	338	192	266	218	91	294	X
129001-101	Phosphoric acid	3		1281.9	0.0	95	100	665	314	299	211	283	225	X
348004-101	Phylloquinone	2		3487.2	-2.3	63	77	201	596	591	331	356	371	-
132003-101	Proline, DL-	2		1303.4	0.0	95	99	180	142	130	117	244		X
175002-101	Putrescine[e]	4		1741.6	-2.0	98	100	232	174	361	214	100	200	X
114003-101	Pyridine, 3-hydroxy-[a]	1		1137.0	-0.3	59	84	35	152	167	136	122	92	-
153002-101	Pyroglutamic acid[f]	2		1528.1	0.9	90	98	102	156	258	230	140	273	X
104002-101	Pyruvic acid	1	1	1036.6	2.5	78	76	65	174	189	115	89	158	-
185001-101	Quinic acid	5		1862.7	-1.1	86	100	1048	255	345	334	537	419	X
337002-101	Raffinose	11		3396.0	-4.7	94	99	110	437	451	361	217	204	-
172002-101	Rhamnose	4	1	1727.8	-0.4	87	93	113	117	160	364	277	321	-
173001-101	Ribitol[i]	5		1734.7	0.1	92	99	1359	319	307	422	217	205	-
168002-101	Ribose	4	1	1690.9	0.0	92	99	255	307	217	160	103	189	X
138001-101	Serine, DL-	3		1369.3	1.6	98	100	542	204	218	278	306	100	X
181002-101	Shikimic acid	4		1820.9	-0.8	84	98	492	204	462	372	255	357	X
209005-101	Sinapyl alcohol	2		2094.6	-1.2	80	97	119	354	234	323	339	293	-
338002-101	Sitosterol, beta-	1		3355.5	-1.6	79	91	810	396	357	486	471	129	-
193001-101	Sorbitol	6		1935.8	-1.1	80	91	102	319	307	157	217	331	X
226002-101	Spermidine	5		2252.6	-1.3	77	94	187	174	144	156	116	491	-
332001-101	Stigmasterol	1		3289.8	1.1	98	99	899	484	394	255	129	379	-
171007-101	Suberic acid	2		1710.8	0.7	44	79	34	187	303	169	117	129	-
134001-101	Succinic acid	2		1326.0	-0.2	97	100	360	247	172	147	262	129	X
264001-101	Sucrose	8		2653.4	0.3	90	89	1811	437	451	361	319	157	X
284005-101	Tetracosanoic acid	1		2838.7	0.7	70	90	211	425	440	117	132	145	-
185004-101	Tetradecanoic acid, n-[a]	1		1852.6	0.4	91	98	179	285	300	117	201	145	-
149002-101	Threitol	4		1501.7	2.5	77	94	130	217	293	307	205	320	-
156001-101	Threonic acid	4		1568.2	3.5	98	100	492	292	220	205	217	245	X
140005-101	Threonic acid-1,4-lactone	2		1382.5	2.2	96	99	166	247	147	262	217	101	X
140001-101	Threonine, DL-	3		1394.0	1.2	98	100	358	219	291	218	117	320	-
147008-101	Threose	3	1	1459.7	1.9	29	78	24	205	117	161	233	262	-
142006-101	Thymine[g]	2		1407.7	3.1	50	74	89	255	270	113	239	140	-
316001-101	Tocopherol, alpha-	1		3145.3	-0.9	88	99	759	502	503	236	237	277	-
300002-101	Tocopherol, gamma	1		3002.9	1.7	42	95	110	488	489	222	223	263	-
274002-101	Trehalose	8		2749.1	0.3	65	87	151	191	169	361	243	331	X
223001-101	Tryptophan, DL-	3		2218.6	-1.0	89	99	452	202	291	218	303	130	-
191004-101	Tyramine	3		1910.4	-0.8	84	96	258	174	338	86	100	264	X
194002-101	Tyrosine, DL-	3		1941.4	-0.8	87	98	283	218	280	354	179	100	X
156004-101	Undecanoic acid, n-[a]	1		1559.9	2.8	33	68	22	243	117	129	132	145	-
136001-101	Uracil	2		1346.6	2.1	40	94	34	241	255	99	113	126	-
127002-101	Urea	2		1260.1	0.7	60	99	122	189	204	171	87	99	-
247002-101	Uridine	3		2468.0	0.7	46	77	47	217	259	243	230	169	-
122001-101	Valine, DL-	2		1220.2	-0.2	97	99	126	144	218	156	246	100	X
171001-101	Xylitol	5		1717.6	-0.9	78	75	157	307	319	332	217	205	-
166001-101	Xylose	4	1	1669.2	0.1	91	99	390	307	217	160	103	189	X

[a] May occur in non-sample controls
[b] Methoxyamines form E- and Z- isomers in stable ratios; only the major isomer is reported
[c] In the presence of ambient air dehydroascorbic acid dimer may be generated from ascorbic acid
[d] Arginine and citrullin may decompose and form ornithine
[e] Agmatine may decompose and form putrescine
[f] Glutamine and, to a less extent, glutamic acid may form pyroglutamic acid
[g] Thymidine readily decomposes and forms thymine
[h] Detailed mass spectral information may be obtained by submitting the MPIMP-ID at http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html (Kopka et al. 2005)
[i] Internal standard compounds

A few general trends were observed (Table II-1). Profiles of root material had less components compared to leaf samples, a tendency which in general was more obvious for the lipid than for the polar fractions. Acetone appeared to yield less lipid components as compared to methanol extracts. Other trends such as differences caused by pH adjustment or temperature may be revealed after further in-depth analysis.

The identification process resulted in 470 identified analytes and MSTs. We chose a S/N threshold >20, a setting that drastically reduced false positives but unavoidably created a small fraction of false negative identifications. A RI deviation of ± 5.0 units was applied based on the two observations that (i) the amount of analyte may influence RI by approximately 2.5 units and (ii) the batch to batch reproducibility plus aging processes of the chosen capillary column had an approximately equal contribution to the variability of absolute RI.

Setting mass spectral match thresholds faced the fundamental problem that mass spectral deconvolution unavoidably results in increasing numbers of chimeric results, in other words mixed mass spectra of a major compound and a co-eluting (trace) compound, when the complexity of samples increases. In addition, GC-MS systems with low mass spectral scanning rates, such as quadrupole or ion trap type of systems, are also prone to chimeric deconvolution. In order to accommodate this inherent problem we choose the reverse match to be > 65 on a scale of 100, because reverse matching allows identification of known mass spectra also in chimeric MSTs. Simple match values are reported, but not thresholded, to distinguish between hits based on perfect deconvolution and hits based on chimeric deconvolution (Table II-2).

As a final result of automated identification, we currently know 172 analytes and 298 MSTs in extracts from *Nicotiana tabacum* L. var. Samsun NN (SNN). In Table II-2 we report only one major analyte for each of the identified 147 metabolites and two of 11 added internal standard compounds. Additional information on these compounds, currently available from our MSRI_ID libraries, can be accessed using the MPIMP-ID identifiers or analyte names reported in Table II-2 with web query pages at the Golm Metabolome Database (GMD; http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html).

For further investigations we focussed on metabolites that were - at least in detectable traces - present in all prepared fractions. We avoided compounds that were close to detection limits in the reference protocol, mnhp, and those compounds that were frequently missed by the used peak finding algorithm. We followed the reasoning of Roessner et al. (2001) for a generalised comparative analysis which is biased by condition-specific components and concentrated on common analytes for detecting general trends of protocol variants. Analysis and discovery of known metabolites and novel analytes that occur only under specific protocol regimes is of course a major aspect of this project, and also one of the prerequisites for method development and optimum design of extended and diverse complementary GC-MS profiles from single samples. Estimation and identification of novel analytes is an ongoing project in our laboratory.

General Trends Caused by Modifications of Extraction

We tested five possible influences that may change qualitative and quantitative composition of metabolite profiles. For the respective protocol codes please refer to Table II-1 and experimental section, 'Extractions'.

(i) The effect of different sample matrices is well known in chemical analysis. Different types of biological samples may not only contain specific sets of metabolites, but may influence the recovery of metabolites and thus introduce - without proper analytical standardisation - apparent changes in metabolite levels as artefacts and not due to genuine effects. Here, we compared root to leaf samples.

(ii) The choice of solvent exerts effects on metabolite extraction yield and enzyme inactivation. We compared the polar protic solvent methanol with polar aprotic acetone, which is completely water-miscible and known to effectively precipitate proteins at high and low temperatures.

(iii) pH is known to stabilise metabolites or to induce acid- and base-catalysed hydrolysis of labile compounds. Further, pH is expected to have an influence on the partition coefficient of the compounds. The initial protocol of metabolite profiling used non-pH-adjusted organic solvents (Roessner et al. 2000). We tested acidified solvent mixtures similar to classical Bieleski mixtures (Bieleski 1964) and carbonate saturated basic medium.

(iv) Two general temperature regimes exist for the extraction of labile compounds. One regime uses hot short extraction for highly effective inactivation of enzymes in organic solvents and short exposure of temperature labile metabolites to heat. The second regime uses cold but long extraction to avoid loss of temperature labile compounds and reduce enzyme activity. Cold extraction has slow extraction kinetics and usually requires extended extraction times. This requirement, as a trade-off, allows the influence of residual enzyme activity, an effect that has been described previously for sucrose cleavage in plant material possibly due to residual invertase activity (Weckwerth et al. 2004). We tested 15 min at 70°C versus -20°C over night.

(v) The liquid/liquid partitioning allows differentiation into polar and lipophilic fractions (Fiehn et al. 2000b; Weckwerth et al. 2004). Combined analysis of both fractions should effectively increase the complexity of metabolite profiles (Gullberg et al. 2004) and avoid effects of irreproducible partitioning of amphipolar compounds. We tested lipid and polar liquid layers separated and compared both with combined extracts obtained by omitting liquid partitioning from the standard protocol.

We currently have accessed data on 64 metabolites (Table II-2). The derivatised GC-MS analytes of these metabolites were selected according to the criteria stated above. PCA was performed on the complete data set and the two first principal components are shown for a general insight into the main variance within the data set (Figure II-1). Figure II-2 demonstrates the major profile classes observed with root and leaf extracts and the high similarity of classification between these two biological matrices. Separate PCA analyses of on one hand the lipid extracts (Figure II-3) and on the other hand polar and combined extracts (Figure II-4) not only demonstrated specific trends but also revealed the underlying analytes that contributed the major variance to each of the calculated principal component.

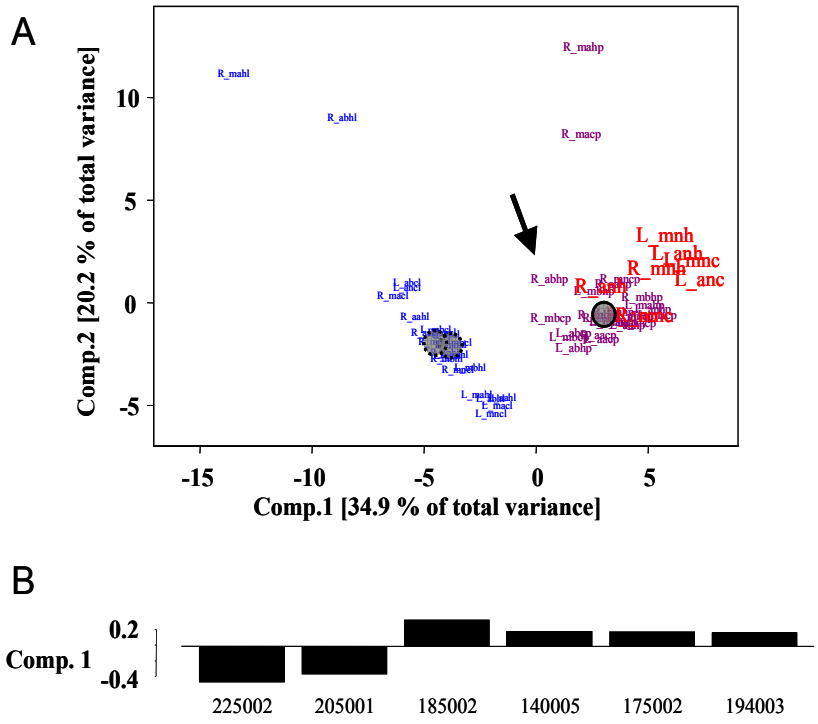

Figure II-1:

PCA analysis of the complete data set of 64 analytes which represents those metabolites that were observed in GC-MS metabolite profiles of all tested protocol variations.
A … Component scores.
B … Ranked analyte loadings.
Comp.1: 225002, octadecanoic acid, n-; 205001, hexadecanoic acid, n-; 185002, dehydroascorbic acid; 140005, threonic acid-1,4-lactone; 175002, putrescine and 194003, galacturonic acid.
Protocol codes are listed in Table II-1 with L_- or R_-prefixes to indicate leaf and root preparations.
Circles indicate positions of the standard protocol mnhp (closed circle) and mnhl (dotted circle) of leaf and root. Lipid preparations (small, blue font), polar preparations (intermediate, violet font), combined preparations (large, red font). The arrow indicates the position of basic pH protocols.

Influence of Biological Matrix

Influences of the biological matrix on metabolite profiles must be avoided. We used the leaf and root standard preparations for normalising the analyte responses obtained from all other extracts. Root and leaf standard preparations, R_mnhp and L_mnhp, therefore, co-localise and are centred to the origin within PCA analyses. The matrix effects on protocol variations can only be estimated relative to these preparations (Figure II-1A, II-3A, II-4A). The lipid fraction clearly exhibited a systematic difference between root and leaf samples (Figure II-3A), for

example the standard protocol applied to the lipid fractions, mnhl. Acidified cold and hot acetone extracts appeared to have smaller matrix effects in the lipid fractions.

In contrast the majority of polar fractions did not separate, implementing that these fractions are subject to the same matrix effects which occur under the conditions of our standard preparation. Caution is, however, recommended if combined lipid and polar fractions of leaf and root are analysed. Cold extracts were most diverse when analysed in combined mode. In addition we found evidence for a clear matrix effect on dehydroascobic acid. Dehydroascorbic acid was - with a few infrequent exceptions - absent from leaf extracts (data not shown). Further examples of matrix specific metabolite recovery are to be expected from future analyses.

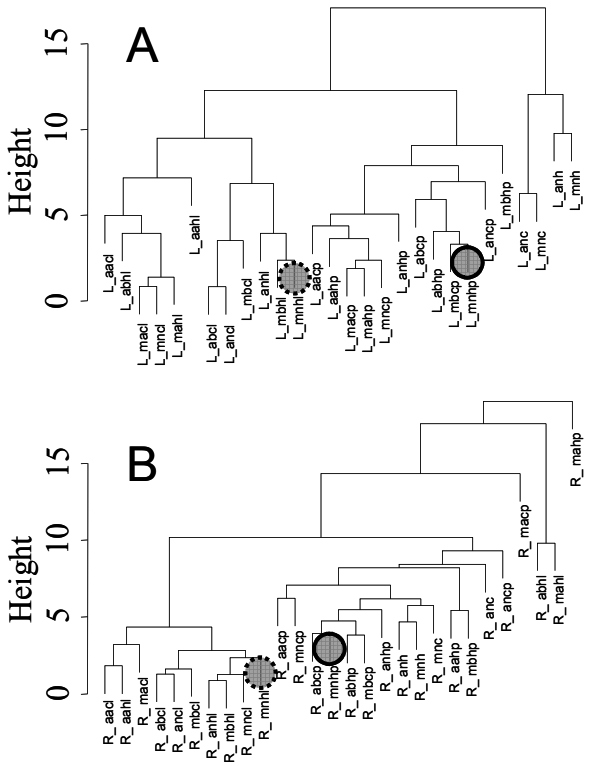

Figure II-2:

HCA analysis of the leaf analyses (A) compared to the root subset (B).
Protocol codes are listed in Table II-1 with L_- or R_- prefixes to indicate leaf and root preparations.
Circles indicate positions of the standard protocol mnhp (closed circle) and mnhl (dotted circle) of leaf and root.

Influence of Extract Polarity

Predominance of the influence of extract polarity was as expected independent of the biological matrix (Figure II-2); it was the major influence in all of our experiments (Figure II-1A). Highly lipophilic compounds, such as octadecanoic and hexadecanoic acid, had the strongest influence on sample clustering in PCA (Figure II-1B). Interestingly, dehydro-ascorbic acid, threonic acid-1,4-lactone and putrescine also ranked among those compounds that were more sensitive to polarity changes of the extraction protocol.

Combined analysis of polar and lipid fractions unexpectedly did not result in an intermediate behaviour between lipid and polar fractions. Dehydroascorbic acid, for example, was strongly increased in combined analyses of leaf samples. Especially leaf samples differed, as was demonstrated by Figure II-4A. Metabolite recovery in combined preparations, however, was highly similar to the respective polar preparation.

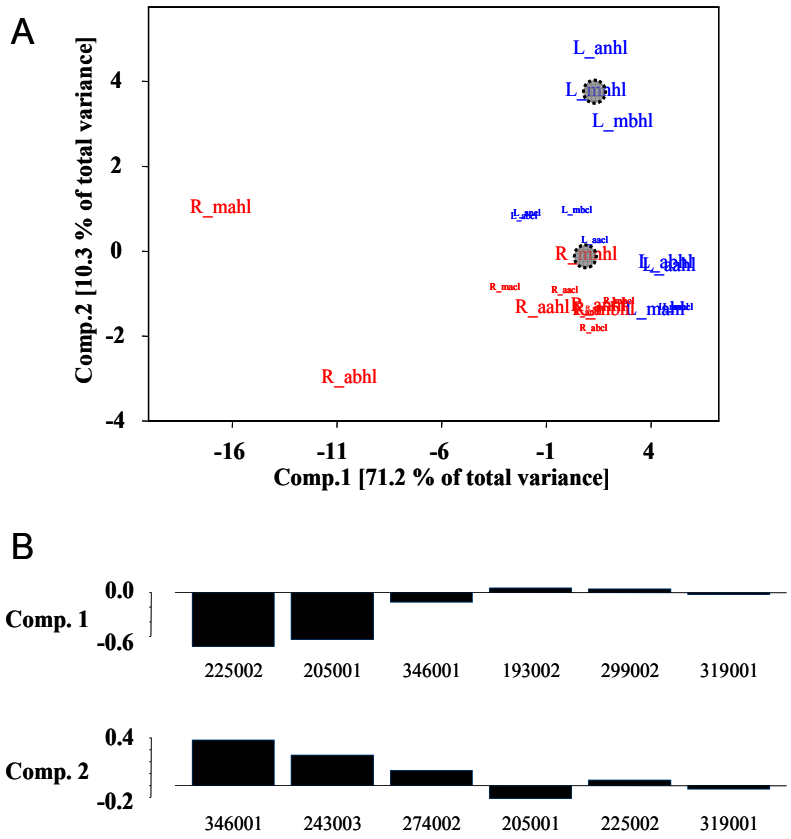

Figure II-3:

PCA analysis of lipid protocol variations.
A … Component scores.
B … Ranked analyte loadings.
Comp.1: 225002, octadecanoic acid, n-; 205001, hexadecanoic acid, n-; 346001, melezitose; 193002 mannitol; 299002, galactinol and 319001, caffeoyl-quinic acid, 5-trans-;
Comp.2: 346001, melezitose; 243003 inositol-phosphate, myo-; 274002, trehalose; 225002, octadecanoic acid, n-; 205001, hexadecanoic acid, n- and 319001, caffeoylquinic acid, 5-trans-.
Protocol codes are listed in Table II-1 with L_- or R_- prefixes to indicate leaf and root preparations.
Circles indicate positions of the lipid fraction of leaf and root analysed with the standard protocol, mnhl (dotted circle). Cold preparations (small), hot preparations (large), leaf preparations (red), and root preparations (blue).

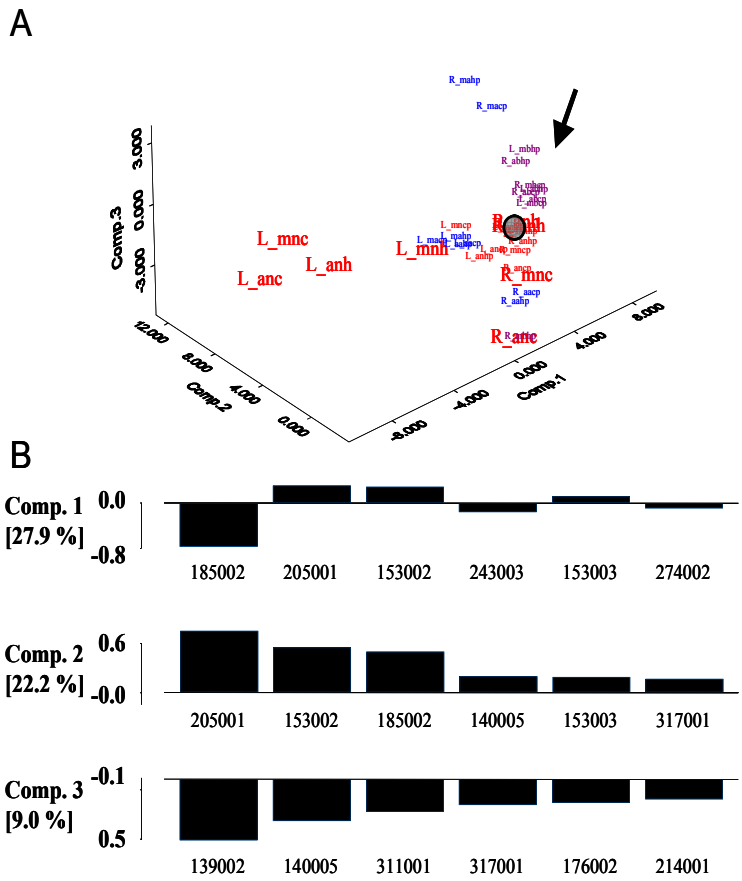

Figure II-4

PCA analysis of polar and combined protocol variations
A…Component scores.
B…Ranked analyte loadings.
Comp.1: 185002, dehydroascorbic acid; 205001, hexadecanoic acid, n-; 153002, pyroglutamic acid; 243003 inositol-phosphate, myo-; 153003; butyric acid, 4-amino- and 274002, trehalose;
Comp.2: 205001, hexadecanoic acid, n-; 153002, pyroglutamic acid; 185002, dehydroascorbic acid; 140005, threonic acid-1,4-lactone; 153003; butyric acid, 4-amino- and 317001, caffeoylquinic acid, 4-trans-;
Comp. 3: 139002, nicotine; 140005, threonic acid-1,4-lactone; 311001, caffeoylquinic acid, 3-trans-; 317001, caffeoylquinic acid, 4-trans-; 176002, aconitic acid, cis- and 214001, caffeic acid, trans-. Protocol codes are listed in Table II-1 with L_- or R_- prefixes to indicate leaf and root preparations.
Circles indicate positions of the standard protocol mnhp (closed circle). Polar preparations (small), combined preparations (large), non-pH-adjusted preparations (red), acidic preparations (blue), and basic preparations (violet). The arrow indicates the position of basic-pH protocols.

Component 1 and 2 (Figure II-4) allowed a clear differentiation of combined analyses of leafs from both root and all non-combined preparations. Metabolite loadings indicated strong influences on the recovery of dehydroascorbic acid, hexadecanoic acid, pyroglutamic acid, inositolphosphate, 4-amino-butyric acid, and threonic acid-1,4-lactone. These differences are to be expected, when changing the standard protocol to combined analyses of lipid and polar metabolites (Figure II-4B). Interestingly, the combined but hot variations of the standard protocol, L_mnh and R_mnh, were least affected, and were in the case of root samples highly similar to the standard protocol samples (Figure II-4A).

Influence of pH

One of the controversial features of the metabolite profiling protocol may have been the absence of buffering during extraction. All tested buffer substances interfered with either chemical derivatisation or chromatographic performance (unpublished data). Determination of the native pH of extracts demonstrated, however, an almost constant pH at approximately pH 6.3. Because slightly acidic pH is the most favourable condition for many pH labile compounds we deemed non-pH-adjusted metabolite profiling to be an

acceptable procedure. Here, we show that only minimal differences are caused by acidifying extracts. In contrast, a basic pH introduced a clear difference (Figure II-1A and II-4A). Substances that responded to basic conditions were revealed by component 3 (Figure II-4B): nicotine, threonic acid-1,4-lactone, caffeoylquinic acids, aconitic acid, and caffeic acid. In lipid analyses acidic pH appeared to be even beneficial. In general, acidification introduced less variance as compared to non-adjusted or basic pH (Figure II-3A).

Influence of Temperature

Temperature had only a minor influence on the pattern of metabolite recovery; strong temperature effects were only observed in lipid profiles (Figure II-3). In this case cold extractions were, as a rule of thumb, less variable than hot extractions. In our hands the previously described hydrolysis of sucrose during prolonged cold extraction in methanol (Weckwerth et al., 2004) was minor except for acidified acetone extractions (data not shown). This observation may be explained by a difference in the stability of invertase(s) from *Arabidopsis thaliana*. *Arabidopsis* was used in the previous study, whereas this study inherently tested the homologous enzyme(s) from tobacco.

Influence of Primary Solvent

The choice of primary solvent, methanol or acetone, did not fundamentally affect metabolite recovery patterns (Figure II-5 to II-7). More detailed analyses of single metabolites may, however, reveal better criteria for the choice of primary solvents.

Analysis of Metabolite Recovery

We chose three metabolites, nicotine, myo-inositol, and myo-inositol-phosphate, which may serve as examples for the detailed analyses of metabolite recovery.

The alkaloid nicotine that was present in all polar fractions, was almost excluded from the basic polar fractions (Figure II-5). Instead, nicotine accumulated in the basic lipid fractions and was partially present in non-pH-adjusted lipid fractions. Cold extraction had the potential to improve nicotine partitioning into the lipid phase. Because partitioning into the lipid layer did rarely correlate with reduced presence in the polar layer, we concluded that basic pH acted on both, extraction from the sample and liquid partitioning of nicotine.

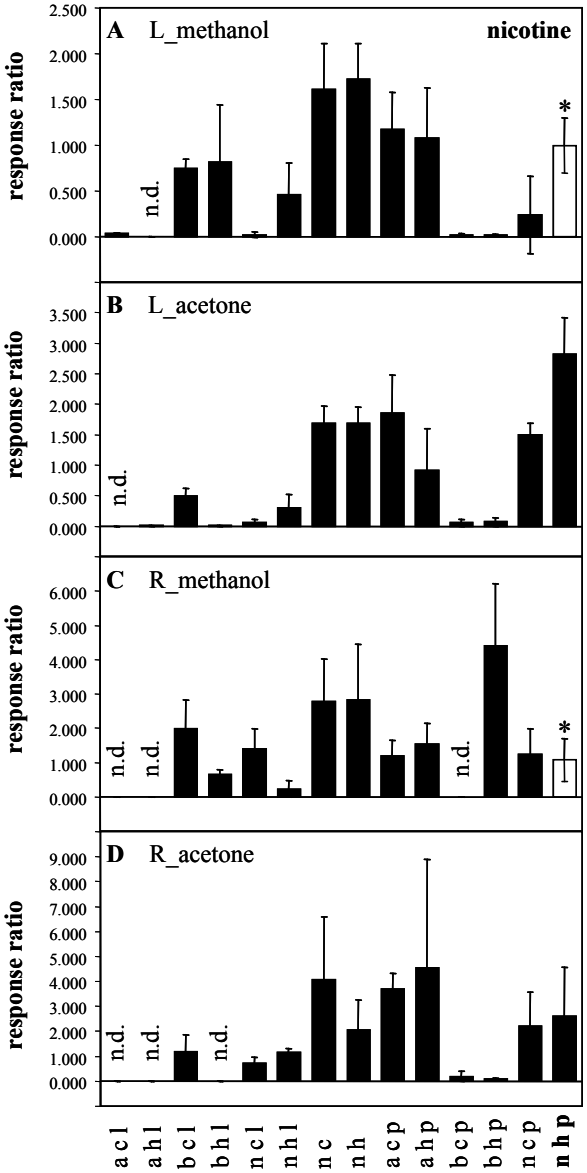

Figure II-5:
Recovery pattern of nicotine.
Leaf preparations (A, B), root preparations (C, D), methanol preparations (A, C) acetone preparations (B, D), n.d. (not detectable). Star indicates standard preparation set to response ratio 1.0.

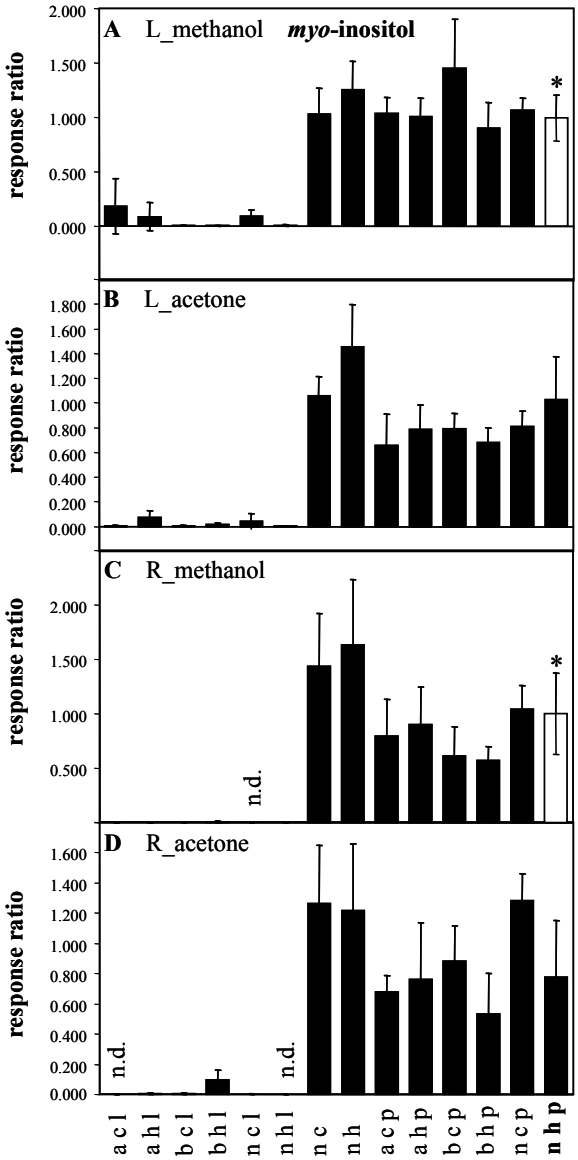

Figure II-6:
Recovery pattern of myo-inositol.
Leaf preparations (A, B), root preparations (C, D), methanol preparations (A, C) acetone preparations (B, D), n.d. (not detectable). Star indicates standard preparation set to response ratio 1.0.

Increased nicotine recovery in R_mbhp (Figure II-5C) was contrary to the generally observed trends. Thus, a hitherto elusive factor was indicated, which may influence nicotine recovery.

Myo-inositol, as expected, exhibited clear partitioning into the polar fractions (Figure II-6). Recovery was almost constant and robust. Temperature and pH effects were small. Improved recovery appeared to be possible by combined analysis of lipid and polar fractions.

Myo-inositol-phosphate exhibited unexpected partial or even dominant presence in lipid fractions (Figure II-7A). We interpret this observation as possible breakdown of phosphatidylinositol lipids, especially under hot basic or non-pH-adjusted conditions. Recovery of this compound consistently increased in combined hot analyses (Figure II-7).

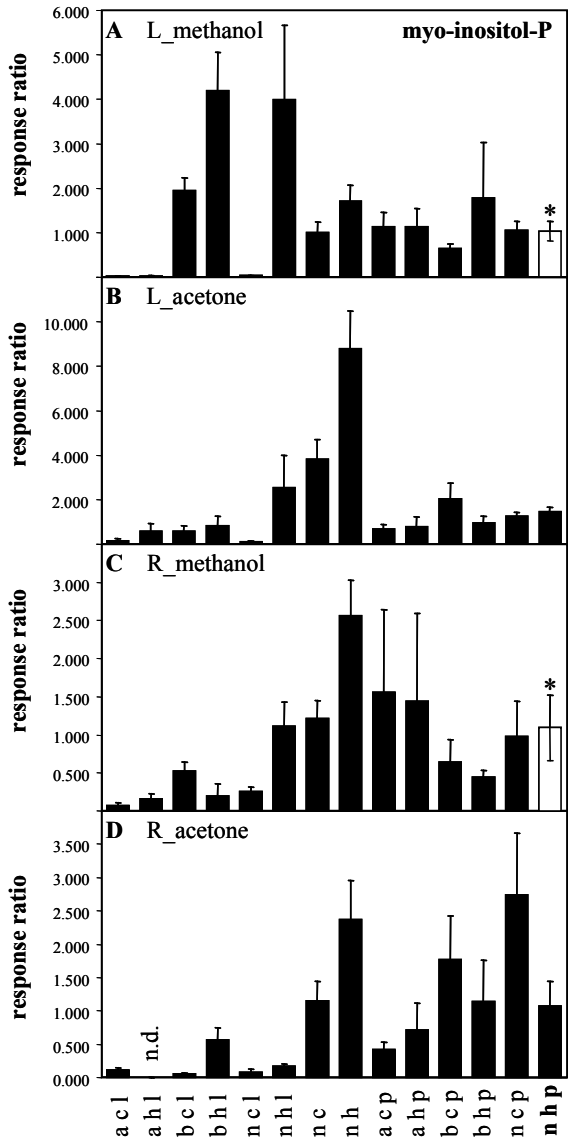

Figure II-7:

Recovery pattern of myo-inositol-phosphate.
Leaf preparations (A, B)
Root preparations (C, D)
Methanol preparations (A, C)
Acetone preparations (B, D)
n.d. (not detectable)

Star indicates that the standard preparation value was set to response ratio 1.0.

Conclusion

Investigations of extraction conditions were clearly proven to be highly useful for obtaining detailed information on the analysis of metabolites in complex mixtures. Information on metabolite recovery as previously suggested and analysed for other plant matrices (Fiehn et al. 2000b; Roessner et al. 2000; Roessner-Tunali et al. 2003; Gullberg et al. 2004) is essential for metabolite profiling and a deeper understanding for the design of

increasingly robust and complex extracts for metabolite profiling. As cautionary remarks we would like to stress that analysis of any novel biological matrix or change to a novel protocol has to be closely checked for possibly arising matrix effects. Our investigation clearly demonstrated that these effects may unexpectedly occur and introduce major variances as shown for most of the analyses of lipid fractions (Figure II-3A). Unfortunately, we had at the beginning of this project only restricted means to test the matrix effect of root versus leaf material under standard conditions more closely, for example, by adding stable-isotope labelled internal standard compounds.

The use of synthetic, labelled internal standards was recommended earlier (Fiehn et al. 2000b; Gullberg et al. 2004). For discussion of an extended standardisation concept and of the potential of *in vivo* stable-isotope labelling we would like to refer the reader to Birkemeyer et al. (2005). The partial ^{13}C-labelling of yeast for the use in yeast metabolomic studies (Mashego et al. 2004) and the *in vivo* labelling of metabolites by ^{15}N (Harada et al. 2004) are first promising applications.

Finally, we expect subsequent in-depth analysis of our data set, especially the inventory and identification of novel and condition-specific analytes, will yield further valuable insight into improved range of metabolite classes, potential breakdown of complex metabolites, possible designs of GC-MS profiles with improved robustness, and the complementary design of multiple GC-MS profiles as compared to the standard protocols of polar and lipid GC-MS metabolite profiles, which are currently in use.

Acknowledgements

We would like to thank A.R. Fernie, A. Erban, Max-Planck-Institute of Molecular Plant Physiology, Potsdam, Germany, and U. Roessner-Tunali, Australian Centre for Plant Functional Genomics, School of Botany, University of Melbourne, Australia, for critically reading and discussing our manuscript.

Chapter III - Normalisation of Profiling Data

Metabolome Analysis: The Potential of *in Vivo* Labelling with Stable Isotopes for Metabolite Profiling

TRENDS in Biotechnology Vol.23 No.1 January 2005

Claudia Birkemeyer, Alexander Luedemann, Cornelia Wagner, Alexander Erban, Joachim Kopka

Abstract

Metabolome analysis technologies are still in early development, because unlike genome, transcriptome, and proteome analyses, metabolome analysis has to deal with a highly diverse range of biomolecules. Combinations of different analytical platforms are therefore required for comprehensive metabolomic studies. Each of these platforms covers only part of the metabolome. To establish multi-parallel technologies, thorough standardisation of each measured metabolite is required. Standardisation is best achieved by addition of a specific stable-isotope labelled compound, a mass isotopomer, for each metabolite. This suggestion, at first glance, seems unrealistic because of cost and time constraints. A possible solution to this problem is discussed in this article. In *in vivo* labelling, saturation with stable isotopes enables the biosynthesis of differentially mass-labelled metabolite mixtures, which can be exploited for highly standardised metabolite profiling by mass isotopomer ratios.

Introduction

The field of analytical biochemistry has recently received a novel extension: metabolomics, a concept that has been defined as the science of the comprehensive monitoring of the metabolic complement in biological systems (Oliver et al. 1998; Fiehn et al. 2000b; Glassbrook et al. 2000; Roessner et al. 2000; Fiehn 2002; Goodacre et al. 2004). Metabolomics provides information about biological systems that cannot be obtained by the classical '–omics' approaches, genomics, transcriptomics and proteomics. Metabolomics could be viewed, therefore, as the missing fourth Rosetta stone (Trethewey

et al. 1999), which fills the metabolic gap within the previously developed systems-wide approaches towards the global analysis of biological processes.

Metabolomic approaches are under dynamic development and several synonyms have been suggested, such as metabonomics (e.g. Nicholson et al. 1999), metabolite profiling (e.g. Mamer 1994) or fingerprinting; refer to Fiehn (2002) and Goodacre et al. (2004) for recent overviews. A multitude of analytical platforms have been introduced (Sumner et al. 2003): spectroscopy fingerprints at infrared (IR), near infrared (NIR) or ultraviolet (UV) wavelength ranges, gas chromatography-mass spectrometry (GC-MS) (Mamer 1994; Fiehn et al. 2000b; Roessner et al. 2001); liquid chromatography-electro spray ionisation-mass spectrometry (LC-ESI-MS) (Goodacre et al. 2002; Huhman and Sumner 2002; Castrillo et al. 2003); capillary electrophoresis coupled to mass spectrometry (CE-MS) (Soga et al. 2002); or liquid chromatography with nuclear magnetic resonance spectroscopy (LC-NMR) (Nicholson et al. 1999; Raamsdonk et al. 2001), just to mention a few. There is no single analytical platform currently conceivable that would enable the multi-parallel analysis of the complete metabolome (Sumner et al. 2003; Kopka et al. 2004), which comprises the full range of chemically diverse biomolecules from low molecular weight volatiles to storage polymers, such as starch or triacylglycerols. The diversity of required methods is in stark contrast to genome, transcriptome, and proteome profiling technologies, which monitor molecules of highly similar chemical properties, such as DNA, RNA and proteins, respectively.

Metabolome analyses not only need to accommodate the high diversity of biomolecules but also need to cover the vast dynamic range of metabolite concentrations. These encompass highly abundant nutrients or primary metabolites and equally important trace compounds that might carry biological signals. In addition, the metabolome is formed by a complex network of reactions, which are subject to rapid enzymatic turnover (Sauer et al. 1999; Christensen and Nielsen 2000; Fischer and Sauer 2003; Forster et al. 2003; Barabasi and Oltvai 2004). Extreme care and fast inactivation of all biochemical reactions during sampling is therefore vital (De Koning and Van Dam 1992; Roessner et al. 2001; Buchholz et al. 2002; Castrillo et al. 2003). Although the sequence information embedded within protein and RNA structure enables unequivocal identification of the source organism, metabolites *per se* do not carry information on their respective origin. Thus, metabolite measurements need to be controlled for artefact chemical contamination that might arise during biological experimentation or chemical analysis.

In spite of considerable technical obstacles metabolome analyses are in high demand and have been widely proposed for studies in molecular physiology (Roessner et al. 2000; Roessner et al. 2001; Kopka et al. 2004), functional genomics (Oliver 1998; Trethewey et al. 1999; Fiehn et al. 2000b; Raamsdonk et al. 2001), clinical chemistry (Nicholson et al. 1999), bio-marker discovery, research on the mode of drug action and monitoring of drug therapy (Raamsdonk et al. 2001; Harrigan and Goodacre 2003; Ilyin et al. 2004). This interest necessitates a short discussion of the properties of novel metabolomic approaches compared with classical chemical analytics.

Analytical Approaches of Metabolome Analysis: General Variants, Properties and Applications

Four major variants of analytical approaches are currently conceivable, fingerprinting, profiling, absolute quantification of pool sizes and, finally, flux analysis, recently suggested as an '-omics' approach in its own right ('fluxomics', e.g. Sauer 2004). Table III-1 gives a short overview of the typical characteristics of these variants. We are aware that all shades of intermediate analytical set-up might exist and that single analytical technologies, such as GC-MS or LC-MS, might enable all four levels of information to be obtained, depending on the chosen experimental set-up.

Quantification of concentrations predates the '-omic' era and was the first means to characterise the metabolic make-up of biological samples via metabolite pool sizes, which was called the 'metabolic phenotype' (Roessner et al. 2001; Fiehn 2002). The availability of pure metabolites is a prerequisite for quantitative analytical methods. Thus, quantification of concentrations might be considered as preconceived: only information on the metabolite under scrutiny is retrieved. The reason for this immanent 'bias' is obvious – quantitative analytics requires calibration of detector signals to metabolite amount. Unequivocal identification and determination of detection limit, linear range, upper loading limit, and of inherent method variability are only possible when pure reference metabolites are available. In addition, recovery analyses are necessary to check the influence of the sample composition on quantitative detection, the so-called matrix effect (Kopka et al. 2004; Matuszewski et al. 2003; Guo et al. 2002).

Table III-1:
Overview of the four general variants in metabolome analysis.
Properties of fingerprinting, profiling, pool size and flux analysis are described for typical analyses.

	Fingerprinting	Profiling	Pool size analysis	Flux analysis
Major field of application	Functional genomics, diagnostics	Functional genomics, molecular physiology	Biochemistry, biotechnology, molecular physiology	Biotechnology, modelling, molecular physiology
Major result	Sample classification based on apparant metabolite pattern	Relative quantification of changes in metabolite pool size, identification and discovery of novel metabolites	Absolute quantification of metabolite pools	Quantification of metabolite flux
Sample composition	High complexity (minimal pre-purification)		Low complexity (partial or highly selective purification)	High complexity (minimal prepurification) possible
Sample throughput	High	High - medium	Low (may be extremely high when dedicated to a single metabolite)	Medium – low
Analytical technology	Non-hyphenated technologies possible	Hyphenated technologies required	Combination of hyphenated or non-hyphenated technologies (dependent on the means of prepurification)	
Metabolite coverage	Limited only by choice of metabolite extraction and analytical technology		Preconceived, limited to a predefined set of targeted metabolites	
Metabolite identification	Identification of metabolites not required	Identification of as many metabolites as possible by chemical properties	Unambiguous metabolite identification required	Unambiguous metabolite and mass isotopomer identification required
Metabolite concentrations	The concentration of the most abundant metabolite determines the highest possible sample load. The dynamic range of the instrument defines the detection limit of co-analysed minor metabolites.		Prepurification enables concentration of trace metabolites and that way an adaptation to the sensitivity range of the analytical instrument. The dynamic range of instrumental analysis is thus not limiting.	
Required control experiments	Detector response is corrected for the initial amount of sample and total losses of material during sample preparation and handling.	In addition to fingerprints, analysis of recovery, detection limits and linearity of detector response of all known metabolites	In addition, quantitative calibration of the detector response by dilution series of pure metabolites.	In addition, tracer experiments with radioactive or stable-isotope labelled metabolites.
Analytical trade-off	The precision of metabolite identification and quantification is sacrificed for optimised sample through-put.	Absolute quantification is substituted for relative quantification in exchange for full metabolite coverage and medium to high sample throughput.	The number of analysed metabolites is restricted in exchange for precise quantification.	The number of analysed metabolites is restricted in exchange for precise quantification of metabolite mass isotopomers.

The discovery of radioactive and stable isotopes, the development of specific detectors for radioactive decay and the single unit mass differences of isotopes sparked investigations of metabolite flux (Szyperski 1998; Sauer et al. 1999; Wittmann and Heinzle 2001; Wiechert 2001; Wittman 2002; Fischer and Sauer 2003; Hellerstein 2003; Nielsen 2003; Sauer 2004). Flux analysis is again targeted to the number of preconceived metabolites, requires application of an isotope tracer that leads to a partial incorporation of isotopes into metabolite pools (Szyperski 1998; Wiechert 2001; Wittman 2002; Hellerstein 2003; Nielsen 2003. Demands on subsequent analytical resolution are high. Resolution is necessary not only for each metabolite but also each isotope variant, the mass isotopomers.

By contrast, profiling and fingerprinting technologies are aimed at analysis of all metabolites that fall within the range of the chosen technologies (e.g. Roessner et al. 2000; Fiehn et al. 2000b). Accordingly, the concept of metabolite purification before analysis is reverted into combinatorial approaches that aim to combine as many metabolites as possible into one analysis. Thus, profiling and fingerprinting could be termed 'unbiased', that is, limited only by the scope of the chosen means of metabolite extraction and analytical technology (Fiehn 2002). The major appeal of these approaches is the potential of discovery. Novel or unexpected metabolites can now be linked to physiological processes or gene function and used as biomarkers (Harrigan and Goodacre 2003; Ilyin et al. 2004).

The 'unbiased' approach, however, comes at a cost. Typical fingerprinting analyses use uncalibrated detector readings obtained from complex metabolite mixtures; theses analyses are often carried out for simple sample classification and biomarker screening. Analysis is, as a rule, thoroughly checked by non-sample controls and calibrated to reference samples, so that relative changes in signals, as compared with the reference sample, can be calculated. This set-up, although highly efficient in screening and classifying high numbers of samples, could be deemed to be insufficient for four main reasons. Firstly, little or no effort is put into assigning metabolite identity to detector signals. Applications are thus restricted to sample classification without the potential to unravel the underlying metabolic and physiological cause. Secondly, single metabolites could be represented by multiple detector readings, such as diverse NMR signals or wavelengths. Therefore, it is conceivable that a single or few abundant metabolites might unknowingly dominate sample classification. Thirdly, artefacts caused by laboratory contamination cannot be completely ruled out and will have an impact on fingerprints. This may become increasingly important in replicate analysis. Fourthly, and most importantly, metabolite

recovery is not controlled in fingerprints. Thus, observed apparent changes do not necessarily reflect direct metabolic changes within the sample but might actually be caused by matrix effects.

For the above reasons, metabolite profiling was suggested, which aims to identify as many metabolites as possible. This concept has become feasible with the advent of hyphenated technologies that allow joined measurement of compounds with multiple chemical properties and that exploit these properties, for example mass and chromatographic retention, for separation and metabolite identification. The increase in resolution and chemical information represents a substantial improvement; values that were obtained from hyphenated techniques and that provide with information on metabolite identity can now be exchanged between laboratories. As a testing ground, an open exchange of metabolite identification based on GC-MS mass spectra and chromatographic retention was envisioned (Wagner et al. 2003), has now been initiated: mass spectral libraries at CSB.DB [http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html], and will hopefully extend by efforts on other technology platforms.

The fundamental advantage of profiling is the opportunity to meet quantitative standards for all identified metabolites within profiles, especially the highly important aspect of metabolite recovery. The technological means for control of matrix effects in metabolite profiles have been suggested previously, namely the use of stable-isotope labelled internal standards (Fiehn et al. 2000b; Fiehn et al. 2000a). However, the high costs of chemical synthesis and the apparent lack of availibility of standard-synthesis protocols for as yet unidentified metabolites made this suggestion appearingly unfeasible for general application.

Recent advances in *in vivo* labelling of micro organisms, specifically of yeast, that use and modify the experimental concepts of flux analysis, open up novel perspectives for avoiding the analytical pitfalls of metabolite profiling (Mashego et al. 2004).

Quantitative Metabolite Profiling by Isotopomer Mass Ratios

Efforts towards quantitative profiling technologies were essential for the general acceptance of transcriptomics and proteomics as semi-quantitative methods (Schena et al. 1995; Lockhart et al. 1996; Unlu et al. 1997; Gygi et al. 1999; Aebersold and Mann 2003; Van den Bergh et al. 2004). Approaches that introduce a differential label in the course of

chemical analysis, currently prevail in quantitative transcriptome (Schena et al. 1995; Lockhart et al. 1996) and proteome (Unlu et al. 1997; Gygi et al. 1999; Aebersold and Mann 2003; Van den Bergh et al. 2004) analyses, for example isotope-coded protein-tagging techniques (Gygi et al. 1999; Aebersold and Mann 2003), protein fluorescence labelling (Unlu et al. 1997; Van den Bergh et al. 2004) and two-colour nucleotide labelling by fluorescent probes (Schena et al. 1995; Lockhart et al. 1996). Although all RNA or protein molecules have common chemical moieties that can be exploited for directed chemical labelling, in metabolome analyses comprehensive chemical tagging technologies are impossible, not least because of the high chemical diversity of metabolites. The most elegant solution to this problem is introduction of label at atomar level through *in vivo* labelling of the biological reference sample (Mashego et al. 2004). The two elements found most abundantly in living organisms are carbon and hydrogen. ^{13}C-carbon can be easily supplied in the form of pure carbon sources to synthetic defined media of microbes or as carbon dioxide to photosynthetic organisms. In contrast, complete hydrogen replacement requires deuterated water and nutrients. For this reason, ^{13}C-carbon labelling appears most promising. *In vivo* labelling can be performed in a similar way to a typical flux experiment but must be directed towards complete labelling, for example by feeding pure U-^{13}C-glucose (99 atom %) starting with colony plating (Figure III-1). Replacement of other elements essential to life appears feasible but is either experimentally more difficult or restricted to limited parts of the metabolome.

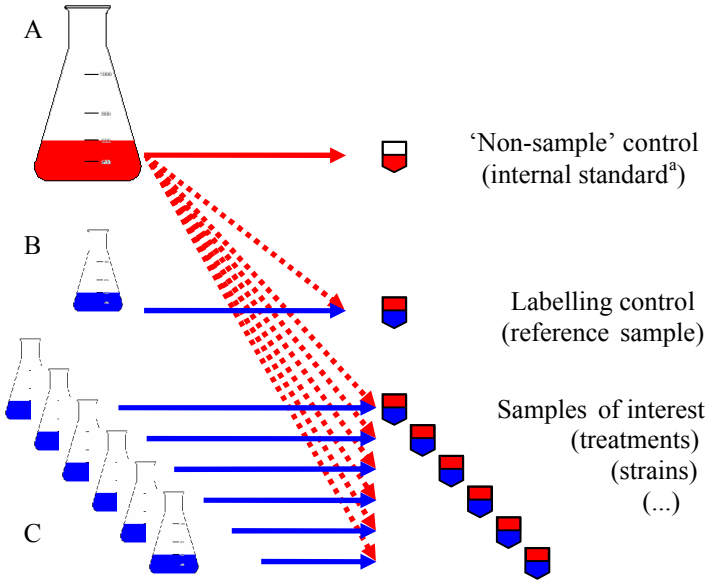

Figure III-1:

Experimental set-up of isotopomer mass ratio profiling.
(A) A yeast parent strain is grown on pure U-^{13}C-glucose (99 atom %) in synthetic defined media (red).
(B) An identical culture is prepared with unlabelled glucose (blue).
(C) Experiments on different strains or treatments are performed with unlabelled carbon sources (blue).
Equal amounts of culture A are combined with samples of B or C. Labelled samples serve as analytical internal standards and are typically monitored by 'non-sample' controls. The labelling control B checks for inherent changes due to ^{13}C-labelling (see Figure III-3 part B). Relative changes in metabolite pool size are determined by mass isotopomer ratio as exemplified in Figure III-3 part A.

[a] The vitamin and auxotrophic supplements may be unlabelled (Figure III-2, note 3).

Novel Applications of ^{13}C-Saturated Microbial Metabolomes

Several fascinating uses of ^{13}C-saturated microbial metabolomes are envisioned and some have already been pursued. The most essential application is the internal standardisation of metabolite profiling experiments (Figure III-1 and III-3) by addition of standardised extracts from ^{13}C-saturated microbial metabolomes. This procedure enables correction for the recovery of each metabolite (Mashego et al. 2004) that are present in these extracts. Moreover, when the same metabolite is profiled using different MS-based technology platforms, the isotope mass ratio will be identical and independent of suppression effects, as occured for example in ESI-MS (Matuszewski et al. 2003) or matrix-assisted laser desorption – time of flight (MALDI-TOF)-MS (Guo et al. 2002; Wittman 2002) experiments. Thus, isotope-mass-ratio profiling has the potential to finally unify measurements obtained from the multitude of relevant profiling technologies.

Although this issue alone justifies efforts to establish ^{13}C-isotope-mass-ratio metabolite profiling, having isotope mass ratios at our disposal opens the door for the use of MS-based methods for the purpose of quantification of pool size, which absolutely require standardisation by stable-isotope reference compounds, such as MALDI-TOF-MS (Guo et al. 2002; Wittman and Heinzle 2001; Kang et al. 2001). In addition, the enrichment of trace compounds or unstable metabolites that is typically accompanied by high and highly variable metabolite losses is now a feasible procedure in metabolite profiling.

The mass spectra of ^{13}C-labelled molecules provide information about the number of carbon atoms in each fragment when mass shifts of highly labelled and unlabelled metabolites are compared (Figure III-2). This knowledge enhances interpretation of mass spectral fragmentation and elucidation of molecular sum formulas, both essential means to narrow down possible chemical structures of yet unidentified compounds (Fiehn et al. 2000a).

A variant of typical flux analyses and tracing experiments for pathway identification can be pursued. A stable-isotope labelled metabolome enables analysis of the fate of unlabelled chemicals. Thus the multitude of cheap and commercially available unlabelled compounds can now be used for tracer and pulse experiments within a ^{13}C-saturated metabolome. This approach appears feasible because we have demonstrated that only minor changes in metabolite levels occur upon ^{13}C-labelling (Figure III-3B).

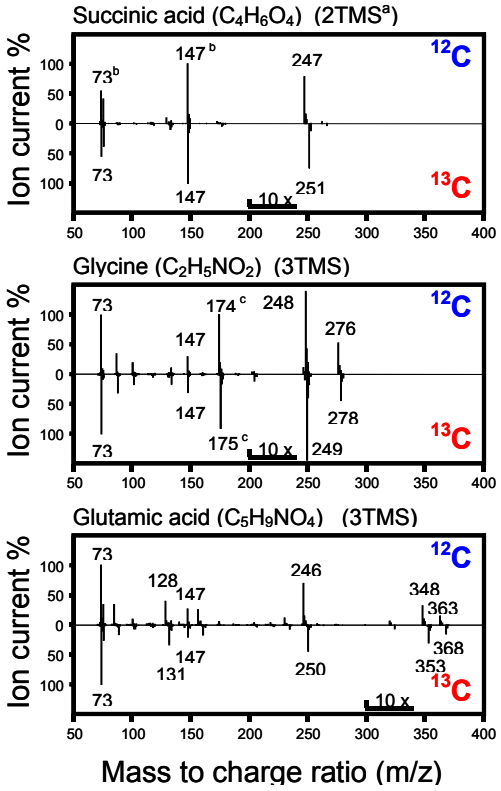

Figure III-2:

Head-to-tail comparison of GC-MS spectra from separate ^{13}C-labelled and ^{12}C metabolite preparations.
Mass spectra show the number of carbon atoms in all those mass fragments that originate from metabolites.

[a] The GC-MS metabolite profiling requires chemical derivatisation by N-methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA). This reagent introduces a specific number of trimethylsilyl moieties (TMS) to each metabolite molecule, as indicated in brackets.
[b] Mass Fragments at 73 and 147 atomic mass units are generated from TMS moieties exclusively.

Note 1: High labelling efficiency is essential, because the chances to obtain a fully labelled mass isotopomer decrease when atom numbers increase (please refer to the glossary). Up to C28 we found unambiguous mass isotopomer distribution in metabolites from yeast grown on pure U-^{13}C-glucose (99 atom %).

Note 2: Incomplete labelling, while insufficient for the determination of high carbon numbers, still allows quantification by mass isotopomer ratios. For high molecular weight metabolites, *in vivo* labelling of less abundant elements, for example N, chemical tagging or analysis of low molecular weight constituents is advisable, such as are employed in proteome analysis (Gygi et al. 1999; Aebersold and Mann 2003).

Note 3: Addition of unlabelled essential vitamins and auxotrophic supplements to microbial cultures causes respective products to be unlabelled. For example, we found NAD^+ to be fully labelled at the 15 carbon atoms which are ultimately synthesised from glucose. The residual six carbon atoms resulting from the nicotinic acid vitamin supplement were unlabelled.

A final, and possibly trivial, but highly effective advantage of *in vivo* stable-isotope labelling is the very fact that a metabolite is labelled *in vivo*. This fact is direct proof that the compound is indeed a metabolite and not one of the possible laboratory contaminations, which have hitherto been tedious to detect and to avoid.

In conclusion, we are convinced that isotopomer-mass-ratio metabolite profiling will not only enhance accurate and quantitative monitoring of the metabolome but also enable comparison of quantitative results from diverse analytical sources and thus take into account the fact that metabolome data need to be generated through a set of diverse analytical platforms.

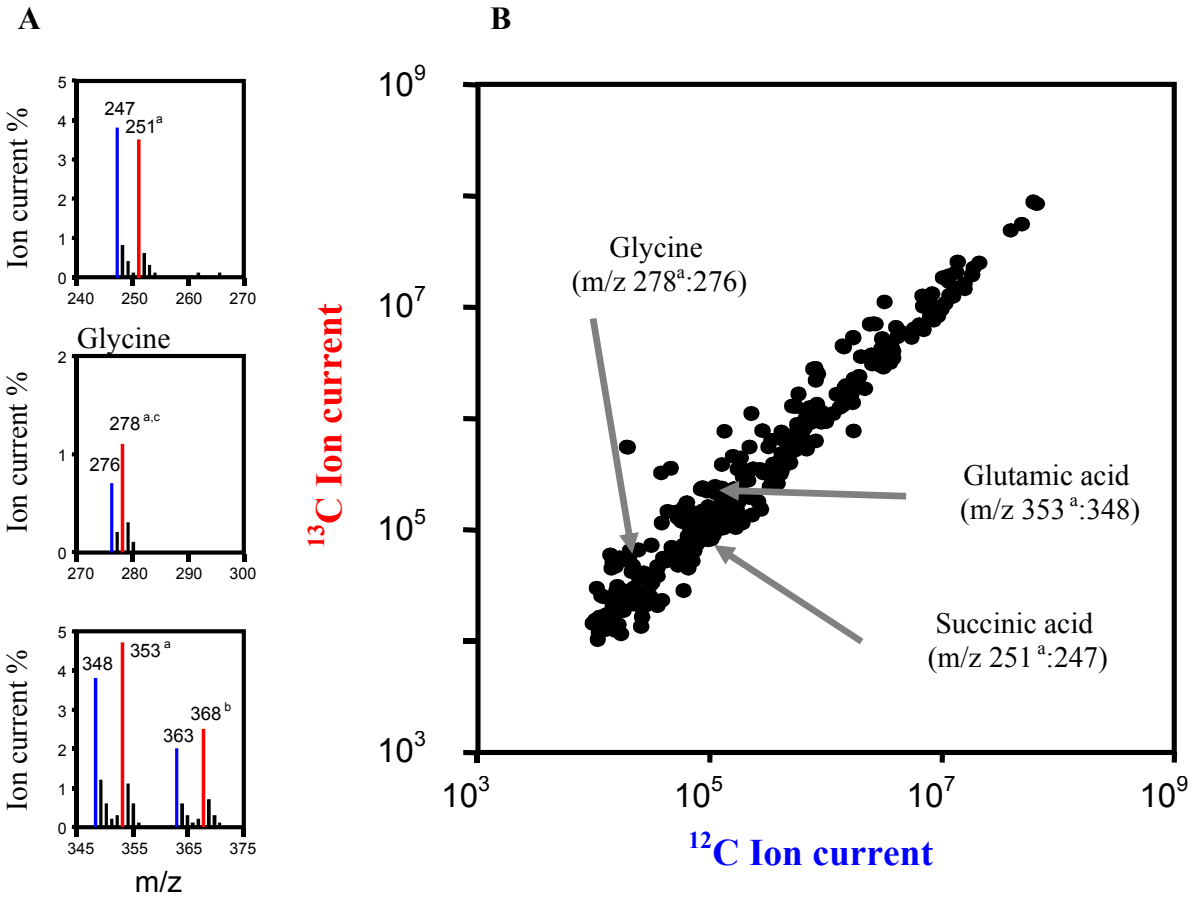

Figure III-3:
Quantification by GC-MS isotopomer-mass-ratio profiling.
(A) Fragment pairs of labelled and unlabelled isotopomers representing the same metabolite. Ion currents reflect the relative changes in metabolite abundance.
(B) Plot of labelled over unlabelled metabolite fragments from a mass-isotopomer-ratio profile, demonstrating that yeast cultures- in this case overnight batch cases- exhibited changes in metabolite levels upon ^{13}C *in vivo* labelling (This plot represents the labelling control experiment in Figure III-1).

[a] Mass fragments which represent the ^{13}C-labelled isotopomer, the specific internal standard for this metabolite.
[b] Metabolites can be monitored by one or multiple mass isotopomer pairs for quantification and confirmation.
[c] Labelled mass isotopomers, especially those with less than three carbon atoms, are best corrected for natural stable mass isotopes.

Glossary

Dynamic range: The range of concentrations, between detection limit and maximum amount of a substance to be quantified by one analytical technology.

Hyphenated analytical technologies: Hyphenation stands for the combination of at least two principles of chemical separation in a single instrument, such as gas chromatography – mass spectrometry (GC-MS), capillary electrophoresis – electrospray mass spectrometry (CE-ESI-MS) or high performance liquid chromatography – nuclear magnetic resonance spectroscopy (HPLC-NMR). Exploiting two chemical properties for compound separation is prerequisite for in-parallel analysis of multiple compounds from a complex mixture.

Mass isotopomer: Chemical substances are composed of naturally occurring (or technically enriched) mixtures of atomic isotopes, which have, as a rule, the same chemical

properties but exhibit different mass. Modern mass spectrometers can resolve the mass differences of a single isotope substitution within a molecule. Each mass variant of a chemical substance is called a mass isotopomer.

Mass isotopomer distribution: The mass isotopomer distribution of molecules can be precisely calculated and is dependent: (i) on the number of atoms present in a molecule and (ii) on the natural or technically enriched isotope abundances of each element (Platzner 1997). High isotope enrichment and low atom numbers result in highly abundant, fully labelled mass isotopomers. Low enrichment and high atom numbers favour partial labelling and cause a broad distribution. For example, a four-carbon molecule may carry a ^{12}C- or ^{13}C-label at either of the positions. The chances for a fully ^{13}C-labelled four-carbon mass isotopomer at ambient 1.1% enrichment are negligible, $0.011^4 = 1.46 \times 10^{-8}$, whereas chances are intermediate, $0.70^4 = 0.24$, or high, $0.99^4 = 0.96$, at 70% and 99% isotope enrichment, respectively. In comparison, the chances of a fully labelled 30-carbon mass isotopomer at 99% enrichment are clearly reduced, $0.99^{30} = 0.74$.

Matrix effect: The matrix effect is a long-standing observation in chemical and enzymatic analyses of complex biological samples. Namely, the nature or composition of complex samples can influence the apparent amount of metabolites and thus might lead to false quantitative results. Matrix effects can either stabilise labile compounds (matrix stabilisation) or suppress compound measurements (matrix suppression). Matrix effects can occur at any step during chemical analysis, from extraction through clean-up, to final instrumental analysis. Well known examples are the matrix suppression effects of ESI-MS (e.g. Matuszewski et al. 2003) or MALDI-TOF-MS (e.g. Guo et al. 2002) technologies or the oxidation of labile metabolites, such as vitamin C.

Recovery: Recovery measurements are the analytical means to control and standardise metabolite measurements for matrix effects. Recovery is routinely expressed as percentage or ratio. The comparison made between equal amounts of metabolites either supplied as a pure reference sample or added to the biological sample under scrutiny. Recovery analyses are most elegantly performed by using chemically synthesised mass isotopomers, which can be distinguished from the respective naturally occurring counterparts by high resolution mass analysis.

Metabolic phenotype: The qualitative and quantitative inventory of all metabolites in a biological sample (Roessner et al. 2002; Fiehn et al. 2000b; Fiehn 2002).

Chapter IV - Metabolite and Hormone Profiling Data: Development of an Evaluation Concept

Automated Pathway Search Using Transcript and Metabolite Data Indicates Spheres of Influence for Signalling Compounds

Submitted to Functional Plant Biology September 2005

Claudia Birkemeyer, Alexander Luedemann, Victoria J. Nikiforova, Ellen Zuther, Petra Birth, Holger Hesse, Rainer Hoefgen, Joachim Kopka

Abstract

A case study of joint evaluation of gene expression and metabolite data from higher plants with data of concentration levels of several signal molecules is introduced. *Arabidopsis thaliana* seedlings exposed to sulfur starvation were the biological test case under evaluation. Five signal compounds, 1-amino-cyclopropanoic acid (ethylene precursor), salicylic acid, jasmonic acid, indole-3-acetic acid, zeatin, and three polyamines including putrescine, were measured and evaluated together with available gene expression and metabolite profiles of the same plant material. A simple data processing strategy eased handling and analysis of the different, complex profiling data. Aided by the use of publicly available databases and the pathway search tool PaVESy, this data evaluation strategy efficiently identified pathways affected under conditions of sulfur starvation.

Introduction

The potential, and difficulty, of a joint evaluation of data from gene expression (Schena et al. 1995) and metabolite profile analyses (Roessner et al. 2000) has been emerged, ever since metabolite profiling was first used as a tool for functional genomics (Fiehn et al. 2000b). The studies of Urbanczyk-Wochniak et al. (2003), Hirai et al. (2004), Gibon et al. (2004), Kopka et al. (2005), and Nikiforova et al. (2005b) are examples, where an attempt was made to combine data that originated from two different profiling methods and targeted on compounds which act on different levels of physiological regulation.

More tools have become available to aid and support an automated data evaluation and pathway search on the basis of profiling data (Mueller et al. 2003; Luedemann et al. 2004;

Thimm et al. 2004; Tokimatsu et al. 2005). Databases such as the Stanford Microarray Database (SMD; Gollub et al. 2003), The *Arabidopsis* Information Resource (TAIR; Rhee et al. 2003), Genevestigator (Zimmermann et al. 2004), and the Comprehensive System Biology Database (CSB.DB; Steinhauser et al. 2004) provide public access to gene expression and metabolite profiles for generating and testing hypotheses. Other profiling methods have been described recently that expand the range of options for a comprehensive analysis of biological systems, such as differential display (Unlu et al. 1997) or the simultaneous measurement of plant hormones (Mueller et al. 2002; Schmelz et al. 2003; Birkemeyer et al. 2003). However, there are still only a few examples available, which provide the possibility of looking at a whole network of determined components and testing the potential of these resources.

For example, gene expression and metabolite data from the same plant material of a sulfur starvation experiment have recently become available; the relationship of gene expression and metabolism was discussed (Nikiforova et al. 2003; Nikiforova et al. 2005a and b). In conclusion of results from a transcript profile (Nikiforova et al. 2003), an involvement of phytohormones in response to sulfur starvation was suggested. One of the most promising candidates seemed to be jasmonate (Nikiforova et al. 2003; Maruyama-Nakashita et al. 2004; Hirai et al. 2004). Nikiforova et al. (2003) found auxin-responsive elements also induced and thus further proposed the involvement of auxin signalling under conditions of sulfur starvation.

In this work, the analysis of the signalling compounds 1-amino-cyclopropanoic acid (ACC), salicylic acid (SA), jasmonic acid (JA), indole-3-acetic acid (IAA), zeatin (Z) and the three polyamines putrescine (Put), spermidine (Spd), and spermine (Spm) was used to generate a test case for the use of automated pathway search on three different kinds of profiling data. A data evaluation strategy is proposed to generate hypotheses on system regulation under distinct experimental conditions and to reduce the focus of interest to a reasonable extent using a methodology easy to follow and to apply. The results are introduced as an example of how to manage the vast amount of data and knowledge that can be collected to describe a biological system, using the data set describing physiological processes in sulfur-starved *Arabidopsis thaliana* seedlings as a template.

Experimental

Plant Growth Conditions

A. thaliana seedlings were grown in Petri-dish tissue culture, on filters placed on the surface of Murashige-Skoog medium (Nikiforova et al. 2003). For sulfur deficiency, the Murashige-Skoog medium was changed to a medium, in which sulfur containing salts were replaced with equimolar amounts of sulfur-free salts. The resulting medium (referred as '–S') contained 89% less sulfur, 8% more nitrogen and 4% more chlorine.

For 'constitutive' sulfur starvation (experiment 1), seeds were sown directly on –S medium; for 'induced' starvation (experiment 2), the seeds were first sown on normal medium, and after growing for eight days, seedlings were transferred to –S medium. In both experiments, samples were taken twice: before (10 days for experiment 1 and six days for experiment 2) and after the appearance of visible changes (13 days for experiment 1 and 10 days for experiment 2). Five biological replicates were performed following the sampling procedure described in Nikiforova et al. (2003).

Values concerning mineral content, gene expression profiles, GC-MS, and LC-MS metabolite profiling data, were taken from analysis of the same plant material, published by Nikiforova et al. (2003 and 2005a).

Phytohormone Analysis

Extraction and purification. For phytohormone quantification the following stable-isotope labelled standards were used: $^{2}H_2$-indole-3-acetic acid, $^{2}H_4$-salicylic acid and $^{13}C_2$-jasmonic acid, kindly provided by Axel Mueller, formerly University of Bochum (Germany); $^{2}H_4$-1-aminocyclopropanoic acid (Sigma, Germany) and ^{15}N-zeatin (Apex Organics, UK). N-methyl-N-(tert-butyldimethylsilyl) trifluoroacetamide (MTBSTFA) was purchased from Macherey-Nagel, Germany. Methanol, chloroform, diethyl ether and formic acid, all HPLC grade, were supplied by J.T. Baker, Philipsburg, NJ, USA. The following solid phase extraction cartridges were used: 200 mg aminopropyl columns (J.T. Baker, Philipsburg, NJ, USA), 200 mg ENV^{+} cartridges (Separtis, Germany) and 500 mg C18-u cartridges (Phenomenex, Germany).

80 µL from a mixture of labelled standards in appropriate concentration for GC-MS analysis was added to 500 mg of the frozen plant tissue, which was immediately subjected

to incubation in 1 mL / 100 mg Bieleski extraction solvent (Bieleski 1964) for 30 min at 70°C. Sample extract was then pushed through C18-u columns, preconditioned with first methanol and second Bieleski solvent. The eluate was dried in a vacuum concentrator, redissolved in 2 x 500 µL diethyl ether, and successively applied to a aminopropyl cartridge (preconditioned in diethyl ether), washed with 2 mL chloroform : 2-propanol (2:1, v/v) and eluted twice with 500 µL 10% formic acid in diethyl ether. This protocol was used in modification of a procedure of Mueller et al. (2002).

Dried extract of the C18-u column effluate (see above) was redissolved a second time in HCl, pH3 (Dobrev and Kaminek 2002) and then applied to ENV^+ columns (preconditioned with HCl, pH3). Sample was eluted with 500 µL 1% formic acid in methanol : water (8:2, v/v) after Mueller et al. (2002). Eluates were evaporated to dryness and the residue was derivatised with 7 µL MTBSTFA (Birkemeyer et al. 2003).

GC-MS conditions. Derivatised samples were measured on an ion trap GCMS system Saturn 2000 (Varian inc. Palo Alto, CA, USA) in EI-MRM (Multiple Reaction Monitoring) mode under the following conditions: DB 5-MS fused-silica column with 30 m length, 0.25 mm inner diameter, 0.25 µm film thickness (Agilent Technologies, Waldbronn, Germany); injection temperature 230°C; injection volume 1 µL, splitless mode; carrier gas helium, flow 1 mL / min; temperature program: 5 min at 70°C, with 10°C / min to 200°C, with 15°C / min to 310°C, then hold for 10 min; transfer line temperature 250°C; ion trap temperature 200°C. Parent ions and excitation amplitudes were set as presented in Table IV-1.

Table IV-1:
Parent ions and excitation amplitudes for phytohormone GCMS-MRM analysis

analyte	parent ions	excitation
ACC	272; 276	0.55
SA	309; 313	0.55
JA	267; 269	0.60
Auxin	232; 234	0.60
Zeatin	315; 316	0.70

Polyamine Analysis

Polyamines were analysed following a method adopted from Smith and Davis (1985). 150 mg of frozen plant material were extracted in 1 mL of 0.2 M $HClO_4$. After 1 h incubation at 4°C and 30 min of centrifugation, 100 µL of the supernatant was used for

dansylation of the soluble polyamine fraction. 10 µL of 0.5 mM 1,6-hexane-diamine was added as internal standard. After adding 100 µL of 1.5 M Na_2CO_3 and 200 µL of 27.8 mM dansyl chloride, samples were incubated with vigorous shaking for 1 h at 60°C in the dark. The remaining dansyl chloride was removed by incubation with 50 µL proline (0.8 M) for 30 min at 60°C in the dark. Polyamines were then extracted with 250 µL toluene; the organic phase was collected and evaporated to dryness. The polyamine containing residue was resuspended in 100 µL methanol and measured by HPLC on a Hypersil ODS reverse phase C18 column (Fa. Supelco) using a 70 to 100% methanol gradient over 20 min duration at a flow rate of 1 mL / min. Fluorescence detection used an excitation/emission wavelength of 365 / 510 nm.

Data Processing

Genes and metabolites responsive to signalling molecules and sulfur. Existing datasets of metabolite and gene expression profiles available for the plant material under investigation (Nikiforova et al. 2003 and 2005a) were employed to identify co-regulations between responses of 6454 genes, 120 metabolites and the measured signal elements, namely five plant hormones and three polyamines. For data processing, relative response values (namely the ratio of absolute response values expressed in area units of the endogenous compound divided by the absolute response value of an appropriate internal standard) of each identified compound of five sample replicates were averaged and normalised to the untreated control for all signal element, metabolite and gene expression data.

A general scheme illustrating the workflow of data processing is introduced in Figure IV-1. The averages of the relative response values (AVG) of the responsive signal elements (four time points and the untreated controls resulted in eight data points) were subjected to analysis of co-regulation to gene expression and metabolite response by using a rule-based approach. For this, a matrix was calculated by the following procedure: increased values compared to the untreated control were set to one, preserved values to zero and decreased values were substituted with minus one. A comparable classification strategy was also chosen by Kaplan et al. (2004).

To consider a substance as decreased or increased under the experimental conditions, thresholds were set relatively low, namely ≥ 1.35 and ≤ 0.74 of the sample/control ratio, respectively +/-0.13 of the logarithm of the sample/control ratio. These thresholds were

chosen to achieve a comprehensive overview by also including minor changes; this way we aimed to place emphasis on the occurrence of trends rather than on the magnitude of changes in concentration levels of the compounds.

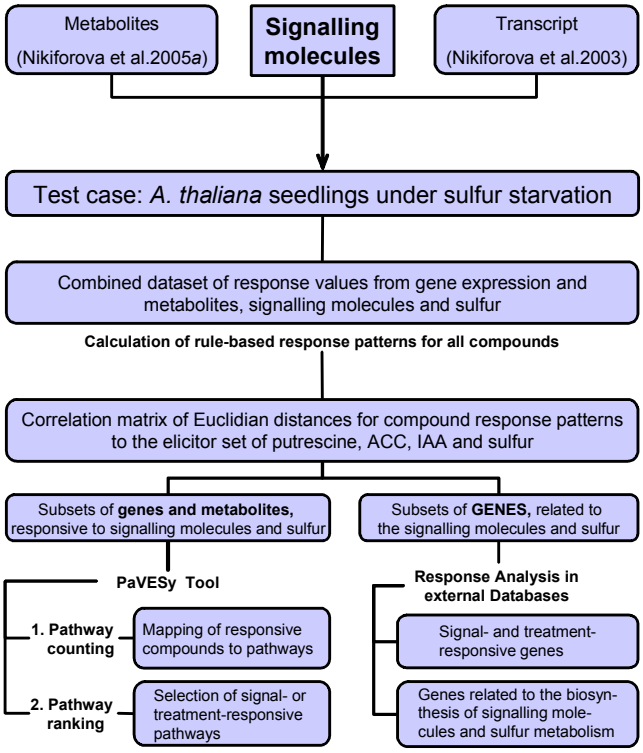

Figure IV-1:

Workflow.
Schematic description of data preprocessing. Compound concentration ratios (treatment/control) are classified in three categories (increased, preserved, decreased) Compounds with altered concentration ratios were compiling subsets that were subjected to pathway search and response analysis in public accessible data resources

All compounds that exhibited strictly the same or contrary behaviour as the set of the relevant signal compounds (those that were observed to be either consistently decreased or increased in all four sample/control ratios) were selected for compiling subsets of genes and metabolites (Tables IV-2 and IV-3), putatively regulated in response to both, sulfate limitation and the signal compounds IAA, ACC and Put.

Pathway search with PaVESy and data reduction. We used the software tool PaVESy (Pathway Visualisation and Editing System), a new data managing system for editing and visualisation of biochemical pathways (Luedemann et al. 2004), on top of the KEGG pathway database (Kanehisa 1997) for processing the subsets of co-regulated metabolites and genes (Table IV-2 and IV-3).

By searching the KEGG database for our investigations, EC numbers were assigned to the ATG identifiers of the genes that were included in the experimental setup of Nikiforova et al. (2003). These EC numbers were subjected to the same evaluation procedure as used for metabolites and as described below.

Table IV-2:
Responsive Metabolites
Names and KEGG compound IDs for metabolites responding to applied conditions. Average response values of five replicates (AVG) were used for calculations. Average ratios to the sample control were calculated: 'Constitutive' starvation for 10 days (AVG2) versus 10 days control (AVG1); 'induced' starvation for 6 days on –S medium after 8 days sufficient S supply (AVG4) versus 14 days control (AVG3); 'constitutive' starvation for 13 days (AVG6) versus 13 days control (AVG5); 'induced' starvation for 10 days on –S medium after 8 days sufficient S supply (AVG8) versus 18 days control (AVG7). 'Rule' columns show the rule-based classification procedure and list the classified response values to the elicitor set of ACC, IAA and sulfur, in application of the thresholds of +/-0.13 lgAVG(ratio) (refer to experimental section: Genes and metabolites responsive to signalling molecules and sulfur).

						Average values ratio				log (Average ratios)				Rule: -0.16>Lg(AVG ratio)>+0.16			
Metabolite	**Method**	**Reference**	**RSD in %**	**KEGG ID**	**Group**	**AVG2/ AVG1**	**AVG4/ AVG3**	**AVG6/ AVG5**	**AVG8/ AVG7**	**AVG2/ AVG1**	**AVG4/ AVG3**	**AVG6/ AVG5**	**AVG8/ AVG7**	**AVG2/ AVG1**	**AVG4/ AVG3**	**AVG6/ AVG5**	**AVG8/ AVG7**
N-acetyl-serine	GC-MS	[Nikiforova 2005a]	75		Sulfur	9.21	2.63	10.60	2.22	0.96	0.42	1.03	0.35	1	1	1	1
O-acetyl-serine	GC-MS	[Nikiforova 2005a]	92	C00979	Sulfur	24.22	8.53	14.48	3.71	1.38	0.93	1.16	0.57	1	1	1	1
DL-Cysteine	HPLC-FD	[Nikiforova 2005a]	48	C00736	Sulfur	0.03	0.02	0.03	0.03	-1.46	-1.75	-1.48	-1.46	-1	-1	-1	-1
Sulfur	ICP-AES	[Nikiforova 2005a]	52	C00087	Sulfur	0.30	0.43	0.27	0.30	-0.52	-0.37	-0.57	-0.52	-1	-1	-1	-1
Glucosinolate_indolyl (m/z 478)	LC-MS	[Nikiforova 2005a]	84		Sulfur	0.16	0.27	0.04	0.03	-0.79	-0.57	-1.44	-1.46	-1	-1	-1	-1
Glucosinolate_aralkyl (m/z 493)	LC-MS	[Nikiforova 2005a]	102		Sulfur	0.05	0.07	0.00	0.00	-1.26	-1.15	-2.52	-2.85	-1	-1	-1	-1
Sulfolipid (m/z 834)	LC-MS	[Nikiforova 2005a]	87		Sulfur	0.12	0.31	0.08	0.17	-0.91	-0.51	-1.11	-0.77	-1	-1	-1	-1
Sulfolipid (m/z 856)	LC-MS	[Nikiforova 2005a]	97		Sulfur	0.07	0.20	0.03	0.07	-1.16	-0.69	-1.51	-1.13	-1	-1	-1	-1
Sulfolipid (m/z 858)	LC-MS	[Nikiforova 2005a]	102		Sulfur	0.05	0.15	0.04	0.07	-1.28	-0.82	-1.45	-1.16	-1	-1	-1	-1
S-adenosyl-methionine (SAM)	LC-MS	[Nikiforova 2005a]	84	C00019	Sulfur	0.16	0.29	0.03	0.03	-0.80	-0.54	-1.47	-1.56	-1	-1	-1	-1
Glutathione	HPLC-FD	[Nikiforova 2005a]	90	C00051	Sulfur	0.10	0.14	0.06	0.11	-0.98	-0.85	-1.19	-0.96	-1	-1	-1	-1
DL-Threonine	GC-MS	[Nikiforova 2005a]	41	C00820, C00188	amino	1.86	2.05	1.35	1.47	0.27	0.31	0.13	0.17	1	1	1	1
DL-Valine	GC-MS	[Nikiforova 2005a]	45	C06417, C00183	amino	1.96	2.01	1.41	1.76	0.29	0.30	0.15	0.25	1	1	1	1
DL-Asparagine	GC-MS	[Nikiforova 2005a]	77	C01905, C00152	amino	3.04	2.77	3.65	2.28	0.48	0.44	0.56	0.36	1	1	1	1
beta-Alanine	GC-MS	[Nikiforova 2005a]	66	C00099	amino	3.56	4.29	3.87	3.28	0.55	0.63	0.59	0.52	1	1	1	1
DL-Glutamine	GC-MS	[Nikiforova 2005a]	101	C00303	amino	5.20	3.10	6.75	4.18	0.72	0.49	0.83	0.62	1	1	1	1
4-Hydroxy-L-Proline	GC-MS	[Nikiforova 2005a]	58	C01015	amino	3.45	2.64	3.58	2.35	0.54	0.42	0.55	0.37	1	1	1	1
DL-Serine	GC-MS	[Nikiforova 2005a]	49	C00716	amino	2.60	2.51	1.99	2.25	0.41	0.40	0.30	0.35	1	1	1	1
DL-Tryptophan	GC-MS	[Nikiforova 2005a]	107	C00806	amino	28.31	16.60	7.06	5.67	1.45	1.22	0.85	0.75	1	1	1	1
Allantoin	GC-MS	[Nikiforova 2005a]	112	C01551	amino	6.12	15.09	4.77	5.57	0.79	1.18	0.68	0.75	1	1	1	1
DL-Arginine	GC-MS	[Nikiforova 2005a]	79	C00792, C00062	amino	0.19	0.63	0.15	0.23	-0.71	-0.20	-0.82	-0.64	-1	-1	-1	-1
DL-Ornithine	GC-MS	[Nikiforova 2005a]	244	C00515	amino	0.06	0.57	0.27	0.06	-1.20	-0.24	-0.57	-1.21	-1	-1	-1	-1
Melibiose	GC-MS	[Nikiforova 2005a]	59	C05402	carbon	1.36	2.40	2.06	2.23	0.13	0.38	0.31	0.35	1	1	1	1
4-Aminobutyric acid (GABA)	GC-MS	[Nikiforova 2005a]	57	C00334	carbon	3.03	2.34	1.80	1.94	0.48	0.37	0.26	0.29	1	1	1	1
Fructose-6-phosphate	GC-MS	[Nikiforova 2005a]	84	C00085	carbon	2.14	1.58	1.95	2.23	0.33	0.20	0.29	0.35	1	1	1	1
Raffinose	GC-MS	[Nikiforova 2005a]	141	C00492	carbon	42.80	24.25	3.00	2.65	1.63	1.38	0.48	0.42	1	1	1	1
Maltose	GC-MS	[Nikiforova 2005a]	52	C00208	carbon	0.72	0.70	0.46	0.52	-0.14	-0.16	-0.33	-0.28	-1	-1	-1	-1
Chlorophyll_a	LC-MS	[Nikiforova 2005a]	46	C05306	carbon	0.52	0.49	0.26	0.50	-0.29	-0.31	-0.59	-0.30	-1	-1	-1	-1
Chlorophyll_b	LC-MS	[Nikiforova 2005a]	33	C05307	carbon	0.64	0.57	0.43	0.58	-0.20	-0.25	-0.37	-0.23	-1	-1	-1	-1
Trehalose	GC-MS	[Nikiforova 2005a]	119	C01083	carbon	0.26	0.23	0.18	0.06	-0.59	-0.64	-0.74	-1.22	-1	-1	-1	-1
DL-Glycerol-1-phosphate	GC-MS	[Nikiforova 2005a]	63	C00623	lipid	3.04	3.81	4.26	2.57	0.48	0.58	0.63	0.41	1	1	1	1
Digalactosyldiglycerol (m/z 932)	LC-MS	[Nikiforova 2005a]	54		lipid	0.40	0.71	0.33	0.44	-0.39	-0.15	-0.49	-0.36	-1	-1	-1	-1
Monogalactosyldiglycerol (m/z 794)	LC-MS	[Nikiforova 2005a]	41		lipid	0.74	0.75	0.59	0.33	-0.13	-0.13	-0.23	-0.48	-1	-1	-1	-1
Digalactosyldiglycerol (m/z 956)	LC-MS	[Nikiforova 2005a]	73		lipid	0.31	0.69	0.35	0.42	-0.51	-0.16	-0.46	-0.38	-1	-1	-1	-1
Phosphoglycerate (m/z 742)	LC-MS	[Nikiforova 2005a]	75		lipid	0.36	0.42	0.16	0.28	-0.45	-0.38	-0.80	-0.55	-1	-1	-1	-1
Noradrenaline	GC-MS	[Nikiforova 2005a]	77	C00547	other	2.47	4.10	2.40	3.18	0.39	0.61	0.38	0.50	1	1	1	1
Anthocyanins	UV/VIS	[Nikiforova 2005a]	68	C08642	other	3.31	2.19	4.26	2.25	0.52	0.34	0.63	0.35	1	1	1	1
Hydroxypheotin_a	LC-MS	[Nikiforova 2005a]	61	(C05797)	other	0.45	0.56	0.52	0.68	-0.34	-0.25	-0.28	-0.16	-1	-1	-1	-1
1-Aminocyclopropane-1-carboxylic acid	GC-MRM		106	C01234	signal	0.69	0.52	0.35	0.53	-0.16	-0.28	-0.46	-0.27	-1	-1	-1	-1
Jasmonic acid	GC-MRM		68	C08491	signal	2.28	1.45	1.66	0.93	0.36	0.16	0.22	-0.03	1	1	1	0
Indole-3-acetic acid	GC-MRM		83	C00954	signal	0.46	0.58	0.18	0.38	-0.34	-0.23	-0.75	-0.41	-1	-1	-1	-1
Spermidine	HPLC-FD		33	C00315	signal	0.69	0.72	0.59	0.53	-0.16	-0.14	-0.23	-0.27	-1	-1	-1	-1
Putrescine	HPLC-FD		47	C00134	signal	2.567	1.913	1.876	1.148	0.41	0.28	0.27	0.06	1	1	1	0
Putrescine	GC-MS	[Nikiforova 2005a]	80	C00134	signal	5.15	8.87	8.75	2.61	0.71	0.95	0.94	0.42	1	1	1	1

Table IV-3:
Responsive genes
Locus and assigned EC Numbers for the genes that responded to applied conditions. Response values of five replicates (AVG) were used to calculate average ratios to the sample control: 'Constitutive' starvation for 10 days (AVG2) versus 10 days control (AVG1); 'induced' starvation for 6 days on –S medium after 8 days sufficient S supply (AVG4) versus 14 days control (AVG3); 'constitutive' starvation for 13 days (AVG6) versus 13 days control (AVG5); 'induced' starvation for 10 days on –S medium after 8 days sufficient S supply (AVG8) versus 18 days control (AVG7). 'Rule' columns show the results of the rule-based classification procedure for responsiveness to the applied conditions in application of the thresholds of +/-0.13 lg(AVGratio) (refer to experimental section: Genes and metabolites responsive to signalling molecules and sulfur).

						AVG ratio sample/control				lg (AVG ratio)				Rule: -0.16>Lg(AVG ratio)>+0.16			
Gene	Reference	RSD in %	EC No.	Enzyme Name	Group	AVG2/ AVG1	AVG4/ AVG3	AVG6/ AVG5	AVG8/ AVG7	AVG2/ AVG1	AVG4/ AVG3	AVG6/ AVG5	AVG8/ AVG7	AVG2/ AVG1	AVG4/ AVG3	AVG6/ AVG5	AVG8/ AVG7
AT5G63910	(Nikiforova 2003)	70	EC 1.8.3.5	prenylcysteine oxidase	sulfur	1.5	1.4	1.6	1.4	0.18	0.15	0.20	0.15	1	1	1	1
AT3G17390	(Nikiforova 2003)	62	EC 2.5.1.6	methionine adenosyltransferase	sulfur	3.1	1.6	2.3	2.7	0.50	0.21	0.36	0.43	1	1	1	1
AT3G02470	(Nikiforova 2003)	53	EC 4.1.1.50	adenosylmethionine decarboxylase	sulfur	2.0	2.4	1.5	1.8	0.30	0.37	0.19	0.25	1	1	1	1
AT2G02930	(Nikiforova 2003)	67	EC 2.5.1.18	glutathione transferase	sulfur	0.4	0.5	0.6	0.2	-0.37	-0.28	-0.21	-0.60	-1	-1	-1	-1
AT3G03630	(Nikiforova 2003)	64	EC 2.5.1.47	cysteine synthase	sulfur	0.4	0.7	0.5	0.3	-0.45	-0.19	-0.26	-0.47	-1	-1	-1	-1
AT5G05730	(Nikiforova 2003)	58	EC 4.1.3.27	anthranilate synthase	amino	3.2	2.6	2.6	1.4	0.50	0.42	0.41	0.16	1	1	1	1
AT1G07920	(Nikiforova 2003)	48	EC 3.6.5.3	protein-synthesizing GTPase	amino	1.5	1.6	1.4	1.5	0.18	0.20	0.16	0.16	1	1	1	1
AT2G05840	(Nikiforova 2003)	42	EC 3.4.25.1	proteasome endopeptidase complex	amino	1.3	1.6	1.4	1.7	0.13	0.20	0.15	0.23	1	1	1	1
AT5G42790	(Nikiforova 2003)	46	EC 3.4.25.1	proteasome endopeptidase complex	amino	1.4	1.5	2.0	1.5	0.16	0.18	0.31	0.18	1	1	1	1
AT1G78870	(Nikiforova 2003)	46	EC 6.3.2.19	ubiquitin-protein ligase	amino	1.7	1.7	1.9	1.4	0.22	0.23	0.28	0.14	1	1	1	1
AT2G23070	(Nikiforova 2003)	44	EC 2.7.1.37	protein kinase	amino	1.5	1.5	1.5	2.0	0.17	0.17	0.18	0.30	1	1	1	1
AT5G17330	(Nikiforova 2003)	48	EC 4.1.1.15	glutamate decarboxylase	amino	1.6	1.7	2.3	1.5	0.21	0.22	0.36	0.17	1	1	1	1
AT1G10060	(Nikiforova 2003)	52	EC 2.6.1.42	branched-chain-amino-acid transaminase	amino	1.5	2.3	2.2	3.4	0.17	0.35	0.33	0.53	1	1	1	1
AT5G57050	(Nikiforova 2003)	53	EC 3.1.3.16	phosphoprotein phosphatase	amino	1.4	1.9	1.5	2.0	0.16	0.27	0.19	0.29	1	1	1	1
AT5G64350	(Nikiforova 2003)	49	EC 5.2.1.8	peptidylprolyl isomerase	amino	1.8	1.8	1.9	1.7	0.26	0.27	0.29	0.24	1	1	1	1
AT1G65960	(Nikiforova 2003)	46	EC 4.1.1.15	glutamate decarboxylase	amino	1.8	1.5	1.5	1.7	0.25	0.19	0.19	0.23	1	1	1	1
AT5G37600	(Nikiforova 2003)	57	EC 6.3.1.2	glutamate-ammonia ligase	amino	2.9	2.4	2.0	2.8	0.46	0.38	0.30	0.44	1	1	1	1
AT3G17820	(Nikiforova 2003)	56	EC 6.3.1.2	glutamate-ammonia ligase	amino	2.1	1.8	1.6	3.3	0.31	0.26	0.21	0.52	1	1	1	1
AT1G20630	(Nikiforova 2003)	56	EC 1.11.1.6	catalase	amino	2.6	2.3	1.8	2.0	0.42	0.36	0.25	0.31	1	1	1	1
AT4G20850	(Nikiforova 2003)	57	EC 3.4.14.10	tripeptidyl-peptidase II	amino	2.1	1.5	1.6	2.1	0.33	0.19	0.22	0.33	1	1	1	1
AT1G53850	(Nikiforova 2003)	50	EC 3.4.25.1	proteasome endopeptidase complex	amino	1.6	1.8	1.6	1.6	0.19	0.26	0.19	0.20	1	1	1	1
AT1G56450	(Nikiforova 2003)	49	EC 3.4.25.1	proteasome endopeptidase complex	amino	1.6	2.1	2.0	2.7	0.19	0.33	0.30	0.44	1	1	1	1
AT2G39800	(Nikiforova 2003)	74	EC 2.7.2.11	glutamate 5-kinase	amino	2.0	1.8	1.7	1.6	0.31	0.27	0.24	0.20	1	1	1	1
AT1G25350	(Nikiforova 2003)	58	EC 6.1.1.18	glutamine-tRNA ligase	amino	2.0	2.1	1.9	1.7	0.30	0.32	0.27	0.22	1	1	1	1
AT5G57020	(Nikiforova 2003)	61	EC 2.3.1.97	N-myristoyl transferase	amino	0.5	0.7	0.4	0.6	-0.33	-0.15	-0.39	-0.21	-1	-1	-1	-1
AT1G11860	(Nikiforova 2003)	63	EC 2.1.2.10	aminomethyltransferase	amino	0.5	0.5	0.4	0.7	-0.30	-0.29	-0.44	-0.16	-1	-1	-1	-1
AT4G17300	(Nikiforova 2003)	65	EC 6.1.1.22	asparagine-tRNA ligase	amino	0.7	0.6	0.5	0.5	-0.17	-0.20	-0.33	-0.26	-1	-1	-1	-1
AT1G79530	(Nikiforova 2003)	39	EC 1.2.1.12	glyceraldehyde-3-phosphate dehydrogenase	carbon/energy	1.7	1.6	1.3	1.9	0.22	0.20	0.13	0.27	1	1	1	1
AT1G24180	(Nikiforova 2003)	49	EC 1.2.4.1	pyruvate dehydrogenase (acetyl-transferring)	carbon/energy	1.5	2.1	1.4	3.4	0.18	0.31	0.13	0.53	1	1	1	1
AT5G56350	(Nikiforova 2003)	46	EC 2.7.1.40	pyruvate kinase	carbon/energy	1.5	1.8	1.7	1.4	0.18	0.25	0.22	0.15	1	1	1	1
AT1G59820	(Nikiforova 2003)	40	EC 3.6.3.1	phospholipid-translocating ATPase	carbon/energy	2.0	1.4	1.4	1.7	0.29	0.15	0.15	0.23	1	1	1	1
AT3G10850	(Nikiforova 2003)	74	EC 3.1.2.6	hydroxyacylglutathione hydrolase	carbon/energy	1.8	2.0	1.5	1.4	0.25	0.31	0.17	0.13	1	1	1	1
AT1G15120	(Nikiforova 2003)	51	EC 1.10.2.2	ubiquinol-cytochrome-c reductase	carbon/energy	1.6	1.7	1.4	1.7	0.20	0.24	0.14	0.23	1	1	1	1
AT5G36880	(Nikiforova 2003)	38	EC 6.2.1.1	acetate-CoA ligase	carbon/energy	1.4	1.7	1.7	1.5	0.14	0.23	0.22	0.19	1	1	1	1
AT3G16950	(Nikiforova 2003)	73	EC 1.8.1.4	dihydrolipoyl dehydrogenase	carbon/energy	2.5	1.4	1.5	1.5	0.40	0.14	0.19	0.18	1	1	1	1
AT4G05390	(Nikiforova 2003)	54	EC 1.18.1.2	ferredoxin-NADP+ reductase	carbon/energy	1.5	1.8	1.9	1.5	0.19	0.24	0.27	0.16	1	1	1	1
AT1G75630	(Nikiforova 2003)	68	EC 3.6.3.14	H+-transporting two-sector ATPase	carbon/energy	2.8	1.9	2.0	2.5	0.45	0.28	0.30	0.41	1	1	1	1
AT4G34720	(Nikiforova 2003)	64	EC 3.6.3.14	H+-transporting two-sector ATPase	carbon/energy	2.6	1.6	2.3	1.8	0.41	0.22	0.37	0.26	1	1	1	1
AT5G25880	(Nikiforova 2003)	72	EC 1.1.1.40	malate dehydrogenase (NADP+)	carbon/energy	2.8	1.8	1.6	3.5	0.45	0.26	0.19	0.54	1	1	1	1
AT2G33220	(Nikiforova 2003)	58	EC 1.6.5.3	NADH dehydrogenase (ubiquinone)	carbon/energy	1.7	2.2	2.8	1.5	0.23	0.34	0.45	0.18	1	1	1	1
AT2G05710	(Nikiforova 2003)	54	EC 4.2.1.3	aconitate hydratase	carbon/energy	1.9	2.3	1.6	2.1	0.29	0.37	0.21	0.32	1	1	1	1
AT4G24620	(Nikiforova 2003)	50	EC 5.3.1.9	glucose-6-phosphate isomerase	carbon/energy	2.1	1.9	2.1	2.2	0.32	0.27	0.33	0.35	1	1	1	1
AT5G66760	(Nikiforova 2003)	54	EC 1.3.5.1	succinate dehydrogenase (ubiquinone)	carbon/energy	1.8	1.6	2.2	2.3	0.25	0.21	0.35	0.36	1	1	1	1
AT3G15640	(Nikiforova 2003)	42	EC 1.9.3.1	cytochrome-c oxidase	carbon/energy	1.5	2.1	1.5	2.3	0.18	0.32	0.18	0.36	1	1	1	1
AT3G54050	(Nikiforova 2003)	73	EC 3.1.3.11	fructose-bisphosphatase	carbon/energy	0.3	0.6	0.3	0.7	-0.48	-0.25	-0.58	-0.16	-1	-1	-1	-1
AT2G27190	(Nikiforova 2003)	65	EC 3.1.3.2	acid phosphatase	carbon/energy	0.5	0.6	0.4	0.4	-0.27	-0.21	-0.41	-0.37	-1	-1	-1	-1
AT1G20020	(Nikiforova 2003)	74	EC 1.18.1.2	ferredoxin-NADP+ reductase	carbon/energy	0.4	0.6	0.5	0.6	-0.42	-0.23	-0.27	-0.23	-1	-1	-1	-1
AT1G42970	(Nikiforova 2003)	57	EC 1.2.1.12	glyceraldehyde-3-phosphate dehydrogenase	carbon/energy	0.4	0.6	0.4	0.5	-0.38	-0.23	-0.42	-0.29	-1	-1	-1	-1
AT1G56190	(Nikiforova 2003)	57	EC 2.7.2.3	phosphoglycerate kinase	carbon/energy	0.4	0.5	0.4	0.6	-0.40	-0.27	-0.37	-0.23	-1	-1	-1	-1
AT3G04790	(Nikiforova 2003)	68	EC 5.3.1.6	ribose-5-phosphate isomerase	carbon/energy	0.6	0.7	0.6	0.5	-0.20	-0.19	-0.22	-0.26	-1	-1	-1	-1
AT1G77590	(Nikiforova 2003)	49	EC 6.2.1.3	long-chain-fatty-acid-CoA ligase	lipid	1.5	1.4	1.4	1.4	0.16	0.16	0.15	0.16	1	1	1	1
AT1G80460	(Nikiforova 2003)	62	EC 2.7.1.30	glycerol kinase	lipid	1.7	1.7	2.1	1.4	0.24	0.24	0.32	0.15	1	1	1	1
AT4G13180	(Nikiforova 2003)	38	EC 1.1.1.100	3-oxoacyl-[acyl-carrier-protein] reductase	lipid	1.8	1.5	1.6	1.5	0.26	0.18	0.20	0.17	1	1	1	1
AT5G06600	(Nikiforova 2003)	61	EC 3.1.2.15	ubiquitin thiolesterase	other	1.5	1.6	2.2	1.4	0.19	0.21	0.33	0.15	1	1	1	1
AT2G35380	(Nikiforova 2003)	54	EC 1.11.1.7	peroxidase	other	1.4	1.6	1.4	1.5	0.14	0.21	0.15	0.17	1	1	1	1
AT4G33420	(Nikiforova 2003)	38	EC 1.11.1.7	peroxidase	other	1.4	1.7	1.8	1.5	0.13	0.24	0.27	0.17	1	1	1	1
AT5G64100	(Nikiforova 2003)	63	EC 1.11.1.7	peroxidase	other	2.4	2.2	1.5	2.3	0.38	0.35	0.18	0.36	1	1	1	1
AT3G52090	(Nikiforova 2003)	63	EC 2.7.7.6	DNA-directed RNA polymerase	other	2.8	1.7	2.6	1.6	0.45	0.22	0.41	0.20	1	1	1	1
AT1G76680	(Nikiforova 2003)	75	EC 1.3.1.42	12-oxophytodienoate reductase	signal	3.7	3.4	4.2	2.3	0.56	0.53	0.62	0.37	1	1	1	1
AT3G44320	(Nikiforova 2003)	76	EC 3.5.5.1	nitrilase	signal	5.3	4.7	4.1	4.4	0.73	0.67	0.61	0.64	1	1	1	1

The whole subset was reformed as a template (all metabolites were assigned to KEGG Compound IDs, genes to EC numbers respectively) and searched for the pathways these compounds are annotated to. Database queries were designed and applied in order to count (i) the numbers of measured and (ii) the numbers of co-regulated metabolites or genes assigned to EC numbers that are present in a particular pathway, and (iii) the total number of all annotated compounds describing that particular KEGG reference pathway. The intention was to emphasise highly affected and best described pathways by the present dataset. The output of this procedure (accomplished following the workflow depicted on Figure IV-1) was a 'pathway ranking' (Table IV-4). Here, the extraction of 'focus pathways' is illustrated: in a particular KEGG reference pathway the number of compounds annotated in KEGG is counted and how many of those were measured and co-regulated with the set of sulfur, ACC, IAA and Put. Those pathways were selected as focus pathways, where co-regulated compounds represent more than 30% of the measured compounds. The higher the proportion of measured to annotated compounds the better a pathway is described.

Reference comparison. The interplay of hormones during sulfur limitation was analysed and discussed comprehensively using publicly available experimental data from SMD, TAIR, and Genevestigator. Array data were downloaded and summarised in order to overview the responsiveness of transcripts of genes that are involved in biosynthesis of signal molecules to phytohormone treatment and sulfur deficiency. Thresholds for regarding a gene as responding to the particular treatment were set to a sample/control ratio of either ≤ 0.5 or ≤ 2.0 (Table IV-5).

In addition, a motif search in the promoter regions of those genes that were responsive to sulfur limitation (Table IV-3) was carried out. Motifs of binding sites known to be responsive to ethylene, auxin, and jasmonate signalling were tested for significant enrichment in promoter regions of the responsive genes in comparison to promoters of the whole set of analysed genes (Nikiforova et al. 2003) by Fisher's exact test (Fisher 1936). The motif search was performed using the software tool PlantPag searching the databases PlantCARE (Lescot et al. 2002), PLACE (Higo et al. 1999), and AtcisDB (Davuluri et al. 2003). Fisher's exact test was carried out using the statistical software package R (R Development Core Team 2004).

Results

Effects of Sulfur Starvation on Signal Compounds

Five signal compunds were determined simultaneously in the plant system under investigation: 1-amino-cyclopropanoic acid (ACC, ethylene precursor), salicylic acid (SA), jasmonic acid (JA, jasmonate), indole-3-acetic acid (IAA, auxin), and zeatin (Z, cytokinin). To our knowledge, this is the first time that levels of these phytohormones and the three polyamines putrescine (Put), spermidine (Spd), and spermine (Spm) have been measured simultaneously under conditions of sulfate starvation.

Under conditions of late response to sulfate deficiency, these compounds form two groups: those were the levels of free Put and Spd, IAA and ACC that responded (by application of average thresholds) to the minus sulfate stimulus compared to untreated plants (Figure IV-2), and those were Z and SA (data not shown), and Spm. Whilst Z concentration level was not altered, SA showed a rather inconsistent pattern. In fact, SA was increased by a factor of two after 13 days of 'constitutive' starvation and decreased by a factor of 0.5 after six days of 'induced' starvation, no change was observed for the other two data points, namely 10 days 'constitutive' and 10 days 'induced' starvation (data not shown).

IAA decreased by factors ranging from 0.2 to 0.5, ACC from 0.4 to 0.7, and Spd by 0.5 to 0.7. Whereas earlier Put concentration was found increased by factors between 2.6 to 8.9 (Nikiforova et al. 2005a), in our measurements it was increased from 1.1 to 2.5 only (Figure IV-2). Polyamine concentrations were between 17 and 140 nmol/g Put, between 50 and 250 nmol/g Spd, and between 5 and 16 nmol/g Spm in individual samples. An increase of JA concentration by a factor of 2.3 could be observed after ten days, and of 1.7 after 13 days of 'constitutive' sulfur starvation. Although JA and Spm did not respond strictly consistent to sulfate starvation after application of our thresholds (refer to experimental section), average response values followed an up-regulation pattern for this phytohormone throughout the experiment (Figure IV-2), except for ten days of 'induced' sulfur starvation.

Figure IV-2 illustrates the groups of responding elements: ACC, IAA, and Spd were decreased and are compared to also decreased marker compounds of sulfur metabolism, cysteine and sulfur response values, in the system under investigation as published by Nikiforova et al. (2003). In contrast, the increased response values of Put, Spm and JA are compared to methionine and likewise increased OAS levels (Nikiforova et al. 2003). All compounds are presented using the average response values of five replicates.

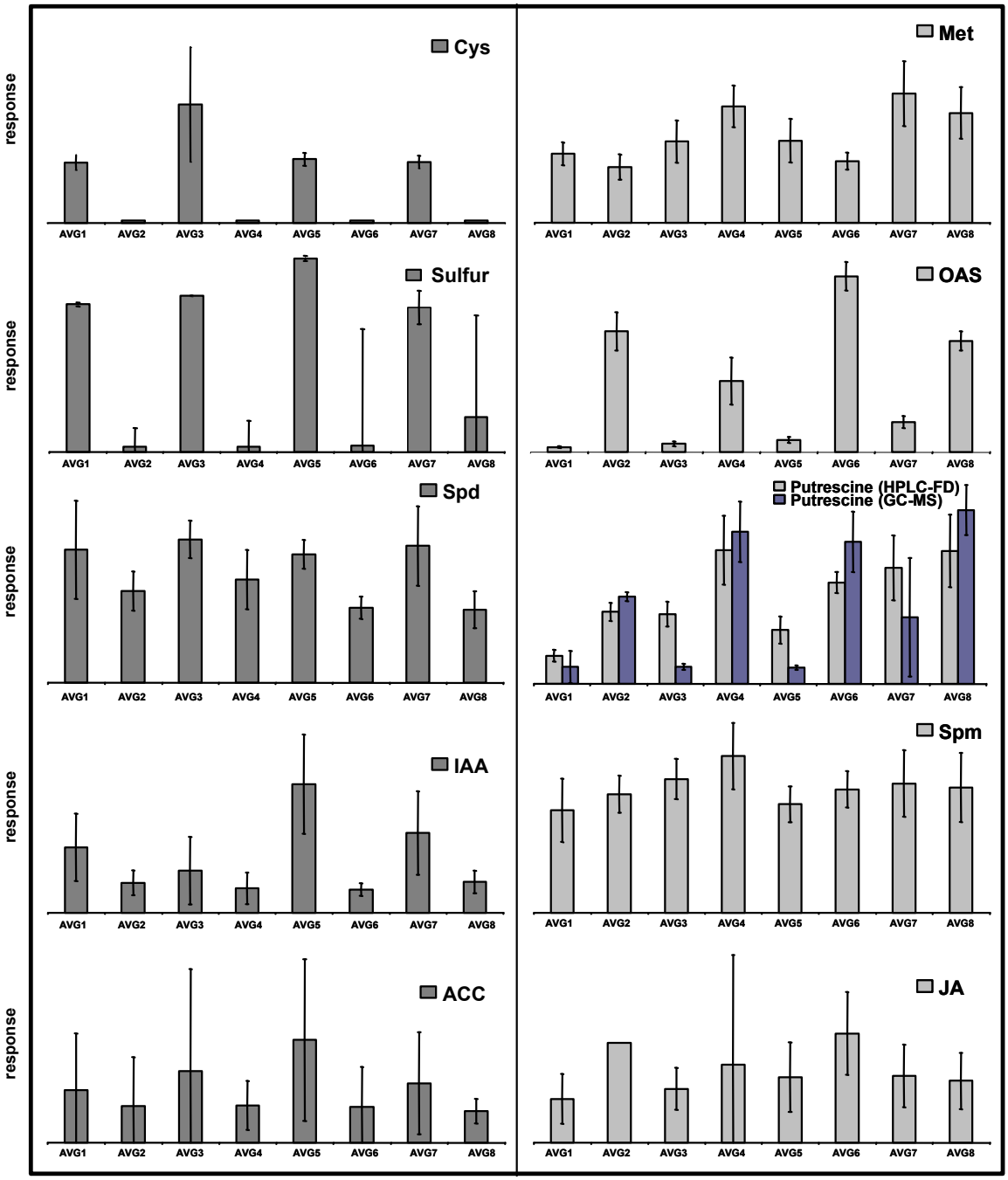

Figure IV-2:
Concentration patterns of decreased (left) and increased (right) signalling and sulfur-related compounds (response ratio to the appropriate internal standard).
'Constitutive' starvation for 10 days (AVG2) versus 10 days control (AVG1); 'induced' starvation for 6 days on –S medium after 8 days sufficient S supply (AVG4) versus 14 days control (AVG3); 'constitutive' starvation for 13 days (AVG6) versus 13 days control (AVG5); 'induced' starvation for 10 days on –S medium after 8 days sufficient S supply (AVG8) versus 18 days control (AVG7).

KEGG Reference Pathways Dissected with PaVESy

A rule-based approach was used to classify responsiveness to the experimental conditions (vide supra). The responding metabolites including sulfur and the set of signal compounds, ACC, IAA and Put, (Table IV-2) and responding genes (Table IV-3) were used to dissect

pathways affected by sulfur starvation. Results of this pathway search and 'weighing' were summarised and grouped (Table IV-4). KEGG reference pathways that met the applied thresholds as described in the experimental section are bold-typed (Table IV-4). In the following, we describe the three major groups of affected pathways.

Sulfur metabolism. Not surprisingly, in the results from the PaVESy pathway search (Table IV-4) a high impact on S-metabolism was stated. The sulfur metabolism KEGG reference pathway (RN00920) describes oxidation states between sulfate and sulfide and their incorporation into cysteine and methionine. Therefore, it exhibits links with methionine and serine metabolism; the latter also involves O-acetyl-serine (OAS). Glutathione (RN00480), cysteine (RN00272), and methionine (RN00271) metabolism pathways, with their close connection to sulfur metabolism, were grouped together with sulfur metabolism as sulfur-related pathways (Table IV-4).

The sulfur-related pathways contain decreased substances like glutathione, cysteine, and S-adenosyl-methionine (SAM, precursor for ethylene and polyamines) as well as glucosinolates and sulfolipids. In contrast, other substrates and compounds present in the sulfur-related pathways accumulated in consequence to low sulfur levels - a well-known phenomenon, for example, is the accumulation of O-acetyl-serine and its precursor serine (Table IV-2).

Nitrogen metabolism. In addition to sulfur metabolism, nitrogen metabolism was listed as a focus pathway as dissected by PaVESy analysis (Table IV-4). The nitrogen metabolism KEGG reference pathway (RN00910) describes levels of redox reactions from nitrate and nitrogen to ammonia and cyanate, and the incorporation of nitrogen into amino acids and amines/amides. It exhibits links to urea cycle (RN00220), glutamate (RNRN00471), glyoxalate (RN00630), and methane metabolism (RN00680) (Table IV-4).

The flow of nitrogen after nitrate assimilation and reduction is mainly mediated via amino acids. Compounds that are responsible for the effect on nitrogen metabolism, in the present case, are the amino compounds, in particular amino acids and Put (Table IV-2); further the levels of transcripts of genes encoding enzymes responsible for the metabolism of amino groups, such as aminomethyltransferase EC 2.1.2.10, glutamate-ammonia ligase EC 6.3.1.2, and nitrilase EC 3.5.5.1 (Table IV-3). Thus, pathways involving amino acids being affected, e.g. tryptophan (RN00380) or glycine, serine and threonine metabolism (RN00260) as well as arginine/ornithine (RN00472) pathway and others are highlighted in Table IV-4.

Table IV-4:
List of focus pathways sorted by best described responsive pathways
KEGG reference pathway name (Pathway description: see footnote for abbreviations); KEGG Pathway ID (ID); number of annotated compounds in KEGG (annot); number of measured metabolites/ genes assigned to EC numbers with available transcript levels (anal); number of responsive metabolites/ genes assigned to EC numbers (res) among the measured compounds; ratio of measured to annotated metabolites/ genes assigned to EC numbers (anal/annot); ratio of responsive to measured metabolites or genes assigned to EC numbers (res/anal); bold values meet the thresholds of focus pathways (refer to experimental section: Pathway search with PaVESy and data reduction).

KEGG Database				Metabolites				Genes assigned to EC No				Sum of both Compounds			
Pathway Description	Group	ID	annot	anal	res	anal/ annot	res/ anal	anal	res	anal/ annot	res/ anal	anal	res	anal/ annot	res/ anal
Cysteine M	sulfur	RN00272	49	9	7	18%	**78%**	5	1	10%	20%	14	8	**29%**	**57%**
Glutathione M	sulfur	RN00480	55	10	4	18%	**40%**	7	1	13%	14%	17	5	**31%**	29%
Methionine M	sulfur	RN00271	56	9	4	16%	**44%**	5	1	9%	20%	14	5	25%	**36%**
Sulfur M	sulfur	RN00920	58	6	5	10%	**83%**	4	1	7%	25%	10	6	17%	**60%**
Nitrogen M	amino	RN00910	199	25	12	13%	**48%**	10	3	5%	**30%**	35	15	18%	**43%**
D-Arginine and D-ornithine M	amino	RN00472	24	6	4	**25%**	**67%**	1	1	4%	**100%**	7	5	**29%**	**71%**
Glutamate M	amino	RN00251	52	13	4	**25%**	**31%**	14	3	**27%**	21%	27	7	**52%**	26%
D-Glutamine and D-glutamate M	amino	RN00471	26	6	2	23%	**33%**	1		4%	0%	7	2	**27%**	29%
Urea cycle; M of amino groups	amino	RN00220	56	12	4	21%	**33%**	10	1	18%	10%	22	5	**39%**	23%
Alanine and aspartate M	amino	RN00252	59	12	4	20%	**33%**	11	2	19%	18%	23	6	**39%**	26%
Glycine, serine and threonine M	amino	RN00260	98	19	8	19%	**42%**	10	2	10%	20%	29	10	**30%**	**34%**
Cyanoamino acid M	amino	RN00460	53	10	4	19%	**40%**	3	1	6%	**33%**	13	5	25%	**38%**
Arginine and proline M	amino	RN00330	107	19	5	18%	26%	15	1	14%	7%	34	6	**32%**	18%
Phenylalanine, tyrosine, tryptophan B	amino	RN00400	56	8	3	14%	**38%**	5	1	9%	20%	13	4	23%	**31%**
beta-Alanine M	amino	RN00410	59	8	2	14%	25%	7	1	12%	14%	15	3	**25%**	20%
Valine, leucine and isoleucine B	amino	RN00290	42	5	1	12%	20%	5	2	12%	**40%**	10	3	24%	**30%**
Taurine and hypotaurine M	amino	RN00430	36	4	1	11%	25%	1	1	3%	**100%**	5	2	14%	**40%**
Methane M	amino	RN00680	49	5	3	10%	**60%**	4	2	8%	**50%**	9	5	18%	**56%**
Valine, leucine and isoleucine D	amino	RN00280	60	6	1	10%	17%	10	1	17%	10%	16	2	**27%**	13%
Histidine M	amino	RN00340	72	7	3	10%	**43%**	3		4%	0%	10	3	14%	**30%**
Selenoamino acid M	amino	RN00450	63	6	3	10%	**50%**	7	2	11%	29%	13	5	21%	**38%**
Tryptophan M	amino	RN00380	112	9	5	8%	**56%**	11	2	10%	18%	20	7	18%	**35%**
Glycolysis / Gluconeogenesis	C/ energy	RN00010	54	6	2	11%	**33%**	20	8	**37%**	**40%**	26	10	**48%**	**38%**
Carbon fixation	C/ energy	RN00710	40	6	1	15%	17%	13	6	**33%**	**46%**	19	7	**48%**	**37%**
Citrate cycle (TCA cycle)	C/ energy	RN00020	42	6		14%	0%	12	3	**29%**	25%	18	3	**43%**	17%
Oxidative phosphorylation	C/ energy	RN00190	17	2		12%	0%	5	4	**29%**	**80%**	7	4	**41%**	**57%**
Pentose phosphate pathway	C/ energy	RN00030	53	8	2	15%	25%	12	3	23%	25%	20	5	**38%**	25%
Reductive carboxylate cycle	C/ energy	RN00720	32	7		22%	0%	5	2	16%	**40%**	12	2	**38%**	17%
Galactose M	C/ energy	RN00052	64	16	2	**25%**	13%	8		13%	0%	24	2	**38%**	8%
Starch and sucrose M	C/ energy	RN00500	88	16	4	18%	25%	8	1	9%	13%	24	5	**27%**	21%
Inositol phosphate M	C/ energy	RN00562	37	7	1	19%	14%	3		8%	0%	10	1	**27%**	10%
Pyruvate M	C/ energy	RN00620	73	4	1	5%	25%	15	6	21%	**40%**	19	7	**26%**	**37%**
Propanoate M	C/ energy	RN00640	69	8	4	12%	**50%**	9	1	13%	11%	17	5	25%	29%
Aminosugars M	C/ energy	RN00530	58	3	2	5%	**67%**	5		9%	0%	8	2	14%	25%
Glycerolipid M	lipid	RN00561	107	6	4	6%	**67%**	10	1	9%	10%	16	5	15%	**31%**
B of steroids	lipid	RN00100	96	3	2	3%	**67%**	10		10%	0%	13	2	14%	15%
Sphingoglycolipid M	lipid	RN00600	45	4	2	9%	**50%**	3		7%	0%	7	2	16%	29%
Aminoacyl-tRNA B	other	RN00970	66	17	8	**26%**	**47%**	17	2	**26%**	12%	34	10	**52%**	29%
Alkaloid B II	other	RN00960	38	10	3	**26%**	**30%**	1		3%	0%	11	3	**29%**	27%
Streptomycin B.	other	RN00521	39	8	3	21%	**38%**	5		13%	0%	13	3	**33%**	23%
Biotin M	other	RN00780	25	4	2	16%	**50%**			0%	0%	4	2	16%	**50%**
Pantothenate and CoA B.	other	RN00770	48	5	3	10%	**60%**	4	1	8%	25%	9	4	19%	**44%**
Purine M	other	RN00230	117	11	4	9%	**36%**	25	5	21%	20%	36	9	**31%**	25%
Flavonoids, stilbene and lignin B	other	RN00940	88	7	1	8%	14%	3	1	3%	**33%**	10	2	11%	20%
Porphyrin and chlorophyll M	other	RN00860	139	11	4	8%	**36%**	3		2%	0%	14	4	10%	29%
Indole and ipecac alkaloid B	other	RN00901	84	6	2	7%	**33%**			0%	0%	6	2	7%	**33%**
Ubiquinone B	other	RN00130	60	4	2	7%	**50%**	2	1	3%	**50%**	6	3	10%	**50%**
Pyrimidine M	other	RN00240	98	6	2	6%	**33%**	9	1	9%	11%	15	3	15%	20%
Terpenoid B	other	RN00900	67	4	1	6%	25%	4		6%	0%	8	1	12%	13%
Folate B	other	RN00790	70	4	2	6%	**50%**	2		3%	0%	6	2	9%	**33%**

M…metabolism; B…biosynthesis; D…degradation; C…carbon

The amino acid arginine holds a central position in plant metabolism; arginine metabolism exhibits links to different KEGG reference pathways including the TCA cycle (RN00020). Arginine by itself is a member of the urea cycle (RN00220) with an important function for the nitrogen balance in plants. The KEGG reference pathway of arginine and proline metabolism (RN00330) also describes metabolic reactions leading to the biosynthesis of polyamines.

A detailed reference list for the description of metabolic effects under sulfur starvation in various plant species can be found in Nikiforova et al. (2005a).

Carbohydrate and energy metabolism. In contrast to the metabolite data set (Table IV-2) the automated pathway search using transcript data for the set of responsive genes (Table IV-3) focussed energy-related pathways such as carbon fixation (RN00710), pyruvate metabolism (RN00620), and oxidative phosphorylation (RN00190), as well as glycolysis/ gluconeogenesis (RN00010) (Table IV-4). This difference was caused by the fact, that not enough annotated metabolites were measured for a proper description of these pathways. Details of particular genes involved in carbohydrate and energy metabolism are given and discussed by Nikiforova et al. (2003).

Other. Among the pathways outside the three groups named above, other pathways also met the applied thresholds of our evaluation procedure (Table IV-4). Some of them are not present in *Arabidopsis*, for example KEGG Alkaloid Biosynthesis II (RN00960). Alkaloid biosynthesis II was addressed regarding altered levels of arginine, ornithine, and Put (Table IV-2) and its close connection to biosynthesis of amino acids such as tryptophan and lysine or the urea cycle. Although alkaloids do not occur in *Arabidopsis*, these results may be a source of valuable hints on general coherences important for other species.

Database Search

Transcript levels of genes encoding enzymes of ethylene, auxin, JA and polyamine biosynthesis, and genes of sulfur-related pathways were searched in SMD, TAIR and Genevestigator databases concerning their induction by ACC, auxin, methyl jasmonate (MeJa), and sulfur limitation treatment. The number of available experiments in each category is presented in comparison to the number of response events among these experiments (Table IV-5).

Among the treatments with signal compounds, auxin treatment seemed to exhibit the greatest impact on biosynthesis of the investigated signalling compounds. In contrast, genes of sulfur metabolism were widely unaffected by phytohormone and ACC treatment. Only a small number of them responded to MeJa treatment.

Table IV-5:

Expression history of genes related to ACC, IAA, JA and Put biosynthesis and sulfur metabolism (SMD, TAIR and Genevestigator databases, 2005).

Transcript levels of genes after treatment with the signal compounds ACC, Auxin and MeJa, or sulfate limitation. Total number of experiments (exp) compared to the number of events, when the particular gene responded to the tested stimulus (resp). For thresholds refer to experimental section: Reference comparison.

			ACC treatment		Auxin treatment		MeJa treatment		Sulfur limitation		Nikiforova 2003	
Locus	Function	Pathway	exp	resp	exp	resp	exp	resp	exp	resp	exp	resp
AT3G61510	ACC synthase 1	ethylene	2	0	2	0	1	0	1	0		
AT1G01480	ACC synthase 2	ethylene	2	1	2	1	1	0	6	2		
AT4G37770	ACC synthase, IAA-inducible	ethylene	2	1	2	1	1	0	1	0		
AT2G22810	ACC synthase 4, induced by IAA	ethylene	2	0	2	1	1	0	1	0		
AT5G65800	ACC synthase 5, cytokinin responsive	ethylene	2	0	2	2	1	0	1	0		
AT4G11280	ACC synthase 6	ethylene	3	0	4	3	1	0	1	0		
AT1G05010	ACC synthase	ethylene	18	3	24	2	5	1	10	0		
AT3G49700	ACC synthase	ethylene	2	0	2	2	1	0	1	0		
AT4G26200	ACC synthase, IAA-insensitive	ethylene	16	1	22	4	2	0	5	2	4	2
AT1G62960	ACC synthase 10	ethylene	5	0	5	0	4	0	6	1		
AT5G28360	ACC synthase, pseudogene	ethylene	3	1	1	0	2	0	5	0		
AT4G08040	ACC synthase, putative	ethylene	2	0	2	1	1	0	6	0		
AT5G51690	ACC synthase, putative	ethylene	5	0	24	6	4	0	10	1		
AT5G51691	ACC synthase, putative	ethylene	0	0	0	0	0	0	0	0		
AT4G23340	ACC oxidase	ethylene	0	0	0	0	0	0	5	0		
AT3G47190	ACC oxidase	ethylene	0	0	0	0	0	0	0	0		
AT1G62380	ACC oxidase 2	ethylene	18	0	24	2	5	0	10	1	4	0
AT1G12010	ACC oxidase, putative	ethylene	4	0	4	0	3	0	6	0		
AT1G77330	ACC oxidase, putative	ethylene	5	0	24	5	4	0	10	1		
AT2G19590	ACC oxidase, putative	ethylene	4	1	23	1	2	0	5	0		
AT1G09400	OPR	jasmonate	2	0	2	1	1	0	6	0		
AT1G17990	OPR	jasmonate	5	0	15	4	4	1	5	1		
AT1G76680	OPR1	jasmonate	1	0	19	0	1	0	9	4	4	4
AT1G76690	OPR2	jasmonate	5	0	5	1	4	0	6	0		
AT2G06050	OPR3	jasmonate	5	0	5	0	4	1	1	0		
AT1G18020	OPR, putative	jasmonate	0	0	0	0	0	0	0	0		
AT5G54206	OPR, putative	jasmonate	0	0	0	0	0	0	0	0	4	1
AT5G42650	allene oxid cyclase AOC4	jasmonate	5	0	24	3	4	1	5	0	4	0
AT1G13280	allene-oxide cyclase AOC	jasmonate	5	0	24	2	4	1	5	1	4	1
AT3G25770	AOC, putative	jasmonate	0	0	0	0	0	0	5	0		
AT3G25780	AOC, putative	jasmonate	5	0	5	0	4	1	6	0		
AT1G19640	JA carboxylmethyltransferase JMT	jasmonate	2	0	2	1	1	1	1	0		
AT4G28680	tryptophan decarboxylase	auxin	15	3	21	4	2	0	10	0	4	1
AT3G26830	cytochrome P450 71B15, putative	auxin	5	1	5	0	4	1	6	0		
AT4G31500	cytochrome P450 CYP79B1	auxin	18	0	24	0	5	1	10	1		
AT4G39950	cytochrome P450 CYP79B2	auxin	5	0	24	3	4	1	10	1	4	2
AT2G22330	cytochrome P450 CYP79B3	auxin	5	0	5	1	4	1	1	1		
AT5G20960	indole-3-acetaldehyde oxidase AO1	auxin	5	0	24	3	4	0	5	0	4	0
AT3G25780	Nitrilase	auxin	2	0	2	0	1	1	6	0		
AT3G44310	Nitrilase 1	auxin	0	0	0	0	0	0	5	0		
AT3G44300	Nitrilase 2	auxin	18	4	24	5	5	0	1	0		
AT3G44320	nitrilase 3	auxin	18	0	24	5	5	2	5	4	4	4
AT2G16500	Arginine decarboxylase 1	Polyamines	5	0	5	0	4	0	6	0		
AT4G34710	Arginine decarboxylase 2	Polyamines	5	0	2	0	4	2	4	1		
AT3G02470	SAMDC 1	Polyamines	5	0	5	0	4	0	6	0		
AT5G15950	SAMDC 2	Polyamines	5	0	5	0	4	0	6	0		
AT3G17715	pseudogene, similar to SAMDC 3	Polyamines	0	0	0	0	0	0	0	0		
AT1G31410	putrescine receptor	Polyamines	5	0	5	0	3	0	6	0		
AT1G23820	spermidine synthase 1	Polyamines	5	0	24	3	4	0	5	0		
AT1G70310	spermidine synthase 2	Polyamines	5	0	24	4	4	0	10	0		
AT5G53120	spermidine synthase, putative	Polyamines	5	0	5	0	4	0	1	0		
AT5G04610	spermidine synthase-related	Polyamines	2	0	21	0	1	0	5	0		
AT5G19530	spermine/spermidine synthase family	Polyamines	4	1	24	9	3	0	10	1		
AT5G13700	Polyamine oxidase precursor	Polyamines	2	0	2	0	1	0	1	0		
AT3G03780	methionine synthase, putative	sulfur	5	0	5	0	4	0	6	0	4	1
AT5G20980	methionine synthase, putative	sulfur	5	0	5	0	4	0	6	0		
AT5G28020	amino-acid biosynthesis OASTL-D2	sulfur	5	0	5	0	4	0	6	0	4	0
AT1G55920	response to sulfate starvation SAT1	sulfur	5	0	5	0	4	0	6	4	4	0
AT3G03630	chloroplast OASTL-CS26	sulfur	5	0	5	0	4	0	6	0	4	2
AT3G57050	cystathionine beta-lyase, chloroplast	sulfur	5	0	5	0	4	0	6	0	4	0
AT1G33320	cystathionine γ-synthase, putative	sulfur	2	0	2	0	1	0	6	2		
AT3G01120	cystathionine γ-synthase, putative	sulfur	5	0	5	0	4	0	1	0	4	0
AT4G14880	cysteine biosynthesis OASTL-A1	sulfur	5	0	5	0	4	0	6	5	4	0
AT2G17640	cysteine biosynthesis SAT106	sulfur	5	0	5	0	4	0	6	5	4	0
AT5G56760	cysteine biosynthesis SAT52	sulfur	5	0	5	0	4	1	1	0		
AT3G13110	cysteine biosynthesis SAT-A	sulfur	5	0	5	0	4	1	6	0	4	1
AT3G59760	cysteine biosynthesis OASTL-C	sulfur	5	0	5	0	4	1	6	0	4	1
AT3G61440	cysteine synthase activity OASTL-C1	sulfur	5	0	5	0	4	0	6	0	4	0
AT3G04940	cysteine synthase activity OASTL-D1	sulfur	5	0	5	0	4	0	6	4	4	0
AT3G22460	N/S metabolism, OASTL-like	sulfur	4	0	4	0	3	0	6	0		
AT2G43750	plastid OASTL-B	sulfur	5	0	5	0	4	0	6	0	4	0
AT1G02500	S-adenosylmethionine synthetase 1	sulfur	5	0	5	0	4	0	6	0	4	1
AT4G01850	S-adenosylmethionine synthetase 2	sulfur	5	0	5	0	4	0	1	0	4	0
AT2G36880	SAM synthetase, putative	sulfur	5	0	5	0	4	0	6	0		
AT3G17390	SAM synthetase, putative	sulfur	5	0	5	0	4	0	6	0	4	3
AT3G22890	sulfate adenylyltransferase ATPS1	sulfur	5	0	5	0	4	0	6	0		
AT4G14680	sulfate adenylyltransferase ATPS3	sulfur	5	0	5	0	4	4	6	0		
AT5G43780	sulfate adenylyltransferase ATPS4	sulfur	5	0	5	0	4	0	6	0	4	0
AT4G21990	sulfate assimilation APR3	sulfur	5	1	5	0	4	0	6	1		
AT1G19920	sulfate assimilation ATPS2	sulfur	5	0	5	0	4	0	6	0	4	0
AT1G62180	sulfate assimilation ATPS2	sulfur	5	0	5	0	4	0	4	0	4	0
AT4G04610	sulfate reduction, APR1	sulfur	5	0	5	0	4	0	6	3		
AT5G04590	sulfite reductase activity SIR	sulfur	5	0	5	0	4	0	6	2	4	0
AT4G08620	sulfate transporter activity Sultr1;1	sulfur	4	0	4	0	3	0	1	1		
AT1G78000	sulfate transporter activity Sultr1;2	sulfur	5	0	5	0	4	0	6	0		
AT5G10180	sulfate transporter activity Sultr2;1	sulfur	5	0	5	0	4	0	6	1	4	4
AT1G77990	sulfate transporter activity Sultr2;2	sulfur	4	0	4	0	3	0	6	2	4	1
AT4G02700	sulfate transporter activity Sultr3;2	sulfur	5	0	5	0	4	0	6	3	4	1
AT1G23090	sulfate transporter activity Sultr3;3	sulfur	5	0	5	0	4	0	1	0	4	2
AT3G15990	sulfate transporter activity Sultr3;4	sulfur	2	0	2	0	1	0	6	0		
AT5G19600	sulfate transporter activity Sultr3;5	sulfur	5	0	5	0	4	0	6	0		
AT5G13550	sulfate transporter activity Sultr4;1	sulfur	5	0	5	0	4	3	6	0	4	3
AT3G12520	sulfate transporter activity Sultr4;2	sulfur	5	0	5	0	4	0	6	1		
AT1G80310	sulfate transporter activity Sultr5;1	sulfur	5	0	5	0	4	0	6	0	4	0
AT2G25680	sulfate transporter activity Sultr5;2	sulfur	2	1	2	0	1	1	6	1		
AT3G51895	sulfate transporter activity Sultr3;1	sulfur	5	1	5	0	4	0	6	0		

Sulfur limitation treatments showed the highest impact on genes of sulfur metabolism. Genes of auxin, jasmonate and ethylene pathways were affected by sulfate limitation in nearly 10% of the available experiments, whereas genes of polyamine biosynthesis showed the lowest sensitivity in the accessible experiments (Table IV-5).

In addition, we found four out of 36 auxin-responsive motifs (AuxRR-core, ARF binding site motif, CATATGGMSAUR and SGBFGMGMAUX28), and two out of nine JA-associated binding sites (JASE1ATOPR1 and JASE2ATOPR1) significantly concentrated in the set of sulfate-limitation responsive genes (Table IV-3) as determined by Fisher's exact test (Fisher 1936). For ethylene, there was no enrichment of binding motifs observed among eight specific binding sites searched.

Discussion

Pathway Search - Identifying Spheres of Influence for Signal Compounds

IAA, Put and the ethylene precursor ACC were signalling compounds that responded to sulfate limitation in the plant system under investigation. Dissected 'focus' pathways, which likely interact with sulfur metabolism, included nitrogen metabolism and carbon-associated pathways, such as glycolysis, carbon fixation, and pyruvate metabolism. Here, we discuss the possible regulation of metabolic effects in response to sulfate limitation in the light of current knowledge.

Nitrogen. Nitrogen metabolism is well known to be influenced by the sulfur nutrition status of plants (DeBoer and Duke 1982; Prosser et al. 1997; Crawford 2000; Hesse et al. 2004; Kopriva et al. 2004). The content of reduced sulfur and nitrogen in plants is strictly maintained at a ratio of 1:20 (Crawford 2000). For example, DeBoer and Duke (1982) found nitrogenase activity depressed during early stages of sulfur deficiency, suggesting the use of joint signalling channels for sulfur and nitrogen assimilation (Marschner 1995).

Harmonisation of nitrogen and sulfur metabolism is likely to be subject to regulation by signalling compounds (Hesse et al. 2003; Kopriva et al. 2004). Interactions between levels of phytohormones and the regulation of amino acids as carriers of the nitrogen flow through an organism or the regulation of protein biosynthesis have been widely studied. Ethylene, for one, was found to act on amine synthesis in plants (Berger 2002). The concentration of amino acids such as cyanoalanine, cysteine and O-acetyl-serine were elevated after ethylene treatment (Maruyama et al. 2001). Ethylene treatment further

influenced the activity of enzymes such as nitrate reductase, arginine decarboxylase, cyanoalanine synthase, and phenylalanine ammonia-lyase in various plant species (Abeles et al. 1992).

Putrescine biosynthesis, for another example, is evidently linked to the assimilation of NH_3 in stressed or unstressed plant tissues, possibly with the function of NH_3-detoxification in alternative to proline accumulation, which may be essential for stress tolerance or sensitivity in plants (Slocum and Weinstein 1990). Put accumulation also was shown to be accompanied by de novo protein biosynthesis (Saftner and Mehta 1990).

Yi-Fang Tsay (TAIR 2005) found further several elements of auxin biosynthesis down-regulated in the transcript profile of an *Arabidopsis* mutant, which was defective in nitrogen assimilation. Among these elements of auxin biosynthesis were the transcripts of genes encoding tryptophan synthase, tryptophan decarboxylase and indole-3-acetaldehyde oxidase; nitrilase 2 and 3 were enhanced.

Carbon-related pathways. The involvement of phytohormones in concerted inter-regulation of nitrogen, sulfur and carbon metabolism has beeen discussed by various authors, for example DeBoer and Duke (1982), Marschner (1995), Crawford (2000), Hawkesford (2000), Nikiforova et al. (2003), Hesse et al. (2003), and Kopriva et al. (2004).

Trehalose and maltose, for instance, were decreased after sulfur starvation (Table IV-3). These two sugars are interconvertible, and trehalose is suggested to play a role in sugar sensing. Sugar sensing was associated with seed germination and development, and, among other signalling compounds, ethylene was discussed as an interacting signal pathway for sugar sensing (Smeekens 2000). Berger (2002) discussed the action of ethylene on sugar and organic acid synthesis in plants. Environmental stresses (drought and cold) led to alterations in carbohydrate concentration and partitioning, and most likely sugar sensing interlaces with stress signalling pathways, such as the jasmonate signalling paths, to modulate metabolism (Smeekens 2000).

The levels of sucrose, trehalose and myo-inositol and other sugar-related compounds, including some sugar acids, were enhanced further after auxin incubation in liquid culture of *A. thaliana* (Stephan Krueger, data not shown). The concentration ratio of signal elements to auxin is important for many developmental processes in plants (Taiz and Zeiger 1998). Although mechanisms of how auxin would mediate a regulation of carbon metabolism are not clear (and may in fact be multiple), an involvement of this phytohormone seems likely.

In conclusion, the metabolism of three macronutrients was shown to be affected under the experimental conditions, as reflected by alterations in a high number of specific pathways (Table IV-4). The presented data evaluation strategy successfully filtered 'focus' pathways known to be relevant to the sulfate starvation response in plants on one hand, but also known to be regulated by the addressed signalling compounds on the other hand. The results clearly indicate, that the extent of metabolic effects reflects an already advanced general affection of the whole metabolism rather than resolving the succession of events leading from the nutritional stimulus via signal transmission to the first observation of metabolic changes. Consequently, a future challenge would be the analysis of early sulfur starvation stages in order to achieve a temporal resolution of the response events.

Finally, we want to emphasise that the pre-processing of the analytical data is crucial for the outcome of the automated pathway search. Pre-processing strategies will decide about extent and complexity of data in the template, which contains the compounds that are responsive to the applied experimental conditions and is used for the automated pathway search; consequently it determines the selection of focus pathways. These strategies need to be developed with caution and awareness; appropriate to answer the question under investigation.

Data Base Search – the Predictive Potential of Transcript Data

For evaluation and confirmation of the analytical results, publicly available experimental data from SMD, TAIR and Genevestigator were used (2005). We focussed on available transcript data for *A. thaliana* exposed to phytohormones and sulfate limitation; particular attention was paid to genes involved in sulfur assimilation and in biosynthesis of the signalling molecules auxin, ACC and polyamines (Table IV-5). JA was included with respect to the observed trend in its response to the experimental conditions (Figure IV-2) and to the fact that several authors suggest a role for this phytohormone in signalling under sulfate starvation (Nikiforova et al. 2003; Maruyama-Nakashita et al. 2003; Hirai et al. 2004).

We found auxin and ACC concentration decreased under conditions of sulfate starvation. Considering that phytohormone treatments are carried out using an excess of the particular hormone, a comparison of phytohormone treatments to the situation where endogenous plant hormone concentration is lowered appears rather difficult. For an example where phytohormone treatment had an opposite effect than the decreased phytohormone

concentrations found in our investigation, the ribulose bisphosphate subunits At1g67090, At5g38430 and At5g38420, being synergistically correlated to ACC and IAA under long-term sulfur starvation (all three were decreased in Nikiforova et al. 2003), stayed mainly unaffected by ethylene treatment and were enhanced under auxin treatment (TAIR 2005), but suppressed by application of MeJa (Schenk et al. 2000). Another example is the induction of the sulfate transporter At5g10180 (Nikiforova et al. 2003), which we found in co-regulation with decreased auxin levels, in contrast to suppression after auxin treatment in an experiment from TAIR (2005).

To assess the probability of involvement of phytohormones in signalling in the system under investigation on transcript level, the enrichment of binding sites that were associated with the action of auxin, ethylene and jasmonate was evaluated in the set of genes that responded to the applied condtions (Table IV-3). The motif search indicated a potential sensitivity of sulfate-limitation responsive genes to auxin and jasmonate. Considering the general trends of phytohormone interactions (Table IV-5), the analysis of the signalling compounds that indicated an interaction of auxin, ethylene and potentially JA under given conditions is not contradictory. Apart from JA biosynthesis, which appeared not affected by ACC treatment, several response events were observed in transcript levels of genes of phytohormone biosynthesis after treatments with auxin, ACC and JA (Table IV-5).

Genes of sulfur metabolism showed almost no induction by auxin, ACC and JA exposure, indicating that these signal compounds might be responsible for the induction of metabolic responses rather than for the regulation of sulfur metabolism itself.

In general, a comparison of transcript data from external resources with the transcript profile from Nikiforova et al. (2003) reflected the complexity of regulation processes in plants (Table IV-5). For the transcript induction of particular genes in response to a certain treatment only trends, but no general causal connections could be observed. In conclusion, the use of transcript data for the prediction of metabolic effects, the expected impact of signal elements, or in return the impact of a nutritional impulse on phytohormone biosynthesis seems rather ambiguous. The complexity of the observed effects, their wide distribution across many pathways indicate that effects in response to a stimulus might be seen more widely across the whole plant's metabolism than initially expected and difficult to predict. Therefore, we consider the introduced methodology highly useful in terms of overviewing genuine, unexpected effects in response to applied (environmental) stimuli.

Acknowledgements

We thank Dr. Axel Mueller, formerly Prof. Elmar Weiler's laboratory (Ruhr-University Bochum, Germany), for kind donation of stable-isotope labelled standards of JA, SA and IAA. Dr. Peter Buchner and Andrew Humphrey (Rothamsted Research, Harpenden, UK) proof read the manuscript. Dirk Steinhauser (Max-Planck-Institute of Molecular Plant Physiology, Potsdam, Germany) assisted with correlation analysis; Stephan Krueger provided access to data from his diploma thesis (Max-Planck-Institute of Molecular Plant Physiology Potsdam / University of Potsdam, Germany).

This work was supported by the European Union (QLRT-2000-00103) and the Max-Planck Society. All work was carried out in the Max-Planck-Institute of Molecular Plant Physiology, Potsdam, Germany.

Chapter V – Summary, Discussion and Outlook

Signal-Metabolome Interactions in Plants

Claudia Birkemeyer

The main focus of this work was the establishment and first application of an analytical methodology that would provide the opportunity to explore interactions of signal compounds and metabolites. In an initial approach, the relationship of the mere endogenous concentrations of five signal compounds, measured as described in Chapter I, and major metabolites covered by the metabolite profiling technique (Roessner et al. 2000; Fiehn et al. 2000b) was investigated. For that, a data set of metabolite and transcript profiling data that was obtained from *Arabidopsis thaliana* seedlings exposed to long-term sulfur starvation (Nikiforova et al. 2003 and 2005a) was complemented by acquisition of phytohormone profiling data and analysed further.

The section begins with a brief list of achieved results and continues with a discussion addressing the state of development in phytohormone and metabolite analysis and data evaluation. Concluding from that, further potential of method development is deduced and aimes in order to meet future challenges are finalised.

Summary – List of Outcomes

Chapter I

A derivatisation procedure was developed and optimised for quantification of the main representatives of six phytohormone classes (indicated in brackets) in order to make these analytes accessible for analysis with GC-MS:

1-amino-cyclopropanoic acid, ACC (precursor of ethylene)

Salicylic acid, SA (salicylates)

Jasmonic acid, JA (jasmonates)

Abscisic acid, ABA (abscisates)

Indole-3-acetic acid, IAA (auxins)

Zeatin, Z (cytokinins, free bases)

Chapter II

The influence of pH, polarity and temperature on robustness of the recovery of metabolites was investigated in plant extracts that were prepared using different extraction protocols. An unsupervised, retention time index-based library of mass spectra was created using the acquired GC-MS chromatograms. The polarity of the extraction solvent was found to modulate the recovery of metabolites more than variation of the other two parameters, temperature and pH. The variation of temperature affected the recovery of lipid metabolites to a greater extent than polar metabolites.

Chapter III

Stable-isotope *in vivo* labelling is discussed by way of a robust method of accurate quantification in the analysis of metabolites in extracts from plant tissues. A 99 atom % labelling efficiency was achieved by breeding yeast on ^{13}C glucose; the suitability of the obtained reference material for application in metabolomic studies is illustrated.

Chapter IV

Phytohormone, metabolite, and gene expression data that were obtained from *Arabidopsis thaliana* seedlings exposed to long-term sulfur starvation were evaluated. The case study presents a data-processing concept that meets the challenge to combine the evaluation of data of different-type compounds. The interplay of ethylene, jasmonate and auxin signalling appeared to determine the coordinated adaptation of sulfur, nitrogen and carbon metabolism to the nutritional status of the plant. Publicly available transcript data have been found to be unappropriate for prediction of metabolic effects under the applied conditions.

Discussion

This section comments the general context of the individual projects that are presented in the four chapters. The discussion about the detailed outcomes of the particular projects is implemented into the corresponding chapters.

The distinctive response of a metabolic system to signalling processes, which was addressed in this work, predefines the need of observing both, metabolic changes and behaviour of signal compounds, simultaneously under the same specific circumstances; ideally in one and the same experiment and plant material.

Therefore, initial efforts focussed on development of an analytical method for the simultaneous analysis of signal compounds and metabolites. A screening for suitable derivatisation methods for phytohormones prior to GC-MS analysis resulted in the selection of tert.-butyldimethylsilylation using N-methyl-N-(tert.-butyldimethylsilyl) trifluoroacetamide (MTBSTFA) as reagent. This procedure comprised several advantages; the high reactivity of the chosen reagent combined with a low specificity, for example, provided potential for extending the coverage of the analytical method to further signal compounds and their conjugates. Also, the reaction with MTBSTFA showed a lower tendency to form multiple derivatives than other tested reagents. The intense $[M-57]^+$ fragments, which are usually generated during Electron Impact ionisation of MTBSTFA-derivatives in the ion source of the mass spectrometer, were highly suitable as quantitative mass traces (Chapter I). The linear behaviour of calibration curves that were obtained for the main derivatives of several phytohormones indicates potential for the application of an external calibration procedure for phytohormone quantification. External calibration could eventually replace the current internal calibration with expensive stable-isotope labelled standards, but would require a highly reproducible sample purification protocol.

Various methods for phytohormone extraction, purification and concentration can be used in combination with the suggested derivatisation procedure, for example the methods of Dobrev and Kaminek (2002), Mueller et al. (2002) and Schmelz et al. (2003).

A possible overestimation by dissociation of phytohormone conjugates that could be caused by residual enzyme action or hydrolysis during sample preparation and the interaction of the internal phytohormone standards with plant matrix must be avoided. Bieleski (1964) suggested a solvent mixture for the extraction of cytokinins in order to suppress the activity of relevant enzymes. Phytohormones are commonly extracted from plant material in pure methanol (Mueller et al. 2002); acidified methanol-water (Liu et al.

2002; Kowalczyk and Sandberg 2001); or acidified methanol-water-chloroform (Schmelz et al. 2003) - all which were applied for acid phytohormones. Bieleski solvent, a particular, acidified methanol-chloroform-water mix was used for cytokinins (Bieleski 1964; Novak et al. 2003).

In order to make an effort in combining the profiling methods of metabolites and phytohormones these typical extraction protocols that are usually applied in phytohormone analysis were tested on robustness for application in metabolite profile analysis. Experimental design was extended further to check also the efficiency of extraction solvents that were used earlier for extraction of other semipolar compounds (Specht and Tillkes 1980). The experiment was carried out using a multi-factorial design in order to test interferences of different extraction conditions (Chapter II).

Solvent mixtures of methanol-chloroform-water as a two-phase system were found to increase the number of metabolites that were extracted from plant matrix and detected with GC-MS in comparison to the application of methanol-water mixes (Gullberg et al. 2004). In contrast, Marhajan and Ferenci (2003) found cold extraction in methanol-water beneficial in comparison to extraction in methanol-chloroform in order to recover the highest possible number of analytes from E. coli cells that were simultaneously determined after separation with thin-layer chromatography (TLC) by scintillation counting of the ^{14}C-labelled metabolites.

However, within this context it shall be noted that, besides on particular extraction conditions, the number and recovery of detected analytes, as a matter of course, also depend on preference or discrimination of particular compounds caused by the characteristics of the used separation system (i.e. GC or TLC) and detection technique (i.e. MS or scintillation counting of labelled compounds). In this sense, a real ‘non-selective’ approach for metabolite profiling seems not feasible at least at the current state of development. The commitment to a particular protocol used in a metabolomic study ideally compromises an optimum for most, but never for all of the compounds (Gullberg et al. 2004).

For an envisaged combination of phytohormone and metabolite extraction two-phase solvent mixtures of different polarity and pH values as well as one-phase extractions with methanol and acetone have been studied (Chapter II). Hierarchical cluster analysis and principal component analysis showed that for a majority of extracted metabolites the original extraction protocol for plant metabolites (Roessner et al. 2000; Wagner et al. 2003) and a protocol incorporating the extraction solvent-mixture according to Bieleski

(1964) resulted in similar recovery behaviour (Chapter II). In conclusion, a combination of the extraction of phytohormones and metabolites from the same plant sample and in one step seems feasible.

Phytohormone analysis is still reliant on the use of stable-isotope labelled standards (Mueller et al. 2002; Schmelz et al. 2003; Dobrev and Kaminek 2002). Labelled standards of signal compounds often are either not commercially available or very expensive. By comparison, normalisation of metabolite profiling is still in its infancy. Applicants of this method often are confronted with artefacts, molecular interaction and decomposition, and suboptimal analytical performance, all which complicate subsequent quantification (Chapter II). Stable-isotope *in vivo* labelling now provides the opportunity to access labelled standards for every compound present in the *in vivo*-labelled organism. By application of this strategy to higher plants, labelled phytohormone standards also should be principally accessible as reference standards for quantitative analysis.

The elements that are most abundant in living beings, and thus would be favoured for *in vivo* substitution with stable mass-isotopomers to produce appropriate internal standards, are carbon, hydrogen, nitrogen and oxygen. Hydrogen, for instance, could be replaced with deuterium by growing plants in heavy water. Complete *in vivo* deuterium-labelling of plant metabolites was not achieved so far; the seeds do not swell in heavy water, and so the plants need to be transferred from normal to heavy water after an initial sprout growth period. The monitoring of the multiple-labelled compounds for quantification is rather tedious; caused by different size and intramolecular polarisation, deuterium-labelled standards show slightly different separation behaviour than the respective endogenous compound (data not shown), depending on the grade of substitution. In addition, acidic, labelled hydrogen atoms tend to be replaced during dissociation processes in protic extraction solvents.

^{18}O *in vivo* labelling using ^{18}O-Acetate was applied for investigation on biosynthesis of lipids by Pollard and Ohlrogge (1999). In order to replace all ^{16}O hetero atoms by ^{18}O in plant metabolites not only the exposure to $H_2{}^{18}O$ but also to $C^{18}O_2$, $^{18}O_2$ respectively, would be required. Thus, complete labelling using the stable isotope of oxygen seems relatively cumbersome.

In comparison, Harada et al. (2004) reported a method for the *in vivo* enrichment of ^{15}N in *Arabidopsis* cell cultures. Nitrogen atoms were rapidly and completely replaced, but a range of important metabolites that do not contain nitrogen atoms, e.g. sugars, consequently would have to be excluded from application of this technology. In

conclusion, *in vivo* labelling with ^{13}C-isotopes appears to be the favourable strategy for future investigations.

The last section of this work focussed on development of an evaluation procedure for phytohormone in conjunction with metabolite data that would be able to meet the challenges of the '-omics' era. Phytohormone analysis in biological systems revealed dramatic changes of phytohormone levels in genetically modified plants (Werner et al. 2001; Skirycz et al., submitted 2005). In contrast, the plants exposed to long-term sulfate starvation, as used for investigations in Chapter IV, did not show a particularly high magnitude of changes in hormonal concentrations and balance.

The biological importance of thresholds of alterations in concentration levels of phytohormones leading to changes in metabolism is yet to be elucidated. References addressing how much the concentration of a signal molecule in response to particular circumstances must change to elicit a certain signalling pathway are not available. Elicitor compounds exhibit concentration thresholds demanded for induction of signalling processes. These concentration thresholds depend on specific parameters such as plant species and tissues, tight environmental conditions, and developmental stage of the given plant individual (Trewavas 2000).

Signalling can be induced not only by changing the mere concentration of a signal compound. Moreover, receptor concentrations and the presence of a down-stream signal transmission apparatus are crucial factors for activation of signalling paths (Trewavas 2000). Also, signalling processes may be triggered by relative changes, i.e. the ratio, of the endogenous concentration of one phytohormone to other signal compounds, possibly without observing altered levels of the particular phytohormone under investigation (Trewavas 1982; Taiz and Zeiger 1998). This fact further indicates the use of common signal transduction pathways, as emphasised by Trewavas (2000) and by other authors investigating signalling processes under particular circumstances, for instance Sano et al. (1996) and Sasaki et al. (2001).

Lichtenthaler (1996) introduced a response/ time-course diagramm of plants that are exposed to stress (Figure V-1). According to that, the sampling of the sulfate-starved *Arabidopsis* seedlings for phytohormone analysis carried out as described in Chapter IV was placed in a state of resistance or even exhaustion of the plant after exposure to sulfate limitation. In conclusion, the observed rather small changes in phytohormone concentration may reflect the tendency of plants to restore hormonal homeostasis after signalling (Eklof et al. 1996; Motyka et al. 1996).

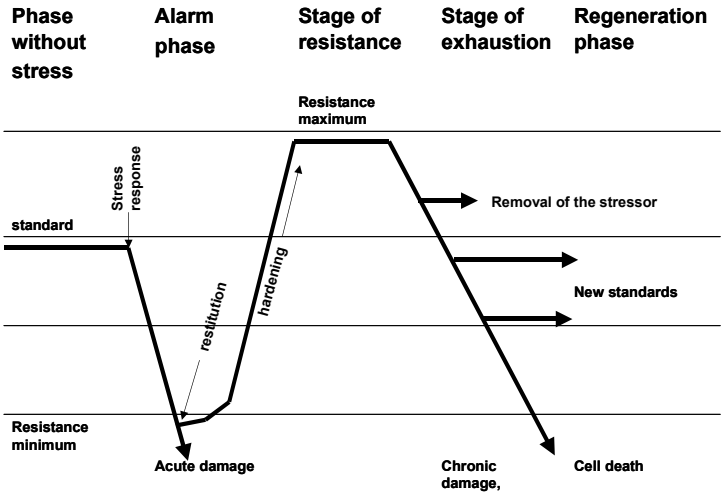

Figure V-1:

Stages of response to stress stimuli in plants

(Lichtenthaler 1996)

With respect to late sampling, a direct causal connection of phytohormone and metabolite levels was difficult to extract from the results of Chapter IV. The results indicated an already proceeded general affection of the whole metabolism rather than they resolved the progression of events leading from signal rise to metabolic changes in detail. The concrete measure of time between signal rise and particular metabolic appearance under the applied conditions would be, however, subject to future investigations.

Effects of signalling on plant metabolism appear in a detailed determined chronological manner (Chapter IV); they are dependent on nature, magnitude and length of the impulse (Trewavas 1982; Taiz and Zeiger 1998; Bray et al. 2000). Precise spatial and temporal activation of signal transduction is crucial to avoid any 'uncontrolled' induction of processes that could be fatal in further plant development (Trewavas 2000).

Different profiling techniques document regulating events on different physiological levels, at different points of time during the course of response in the plant organism. All stimulus responses are further overlapped by developmental processes and vice versa (Trewavas 2000). Multiple secondary effects, such as unrecognised but interfering environmental conditions or diurnal rhythm, do also affect metabolic levels (Urbanczyk et al. 2003) and may mask coherences. Such factors are especially important for an investigation aiming on understanding the regulation mechanisms of long-term responses to a particular stimulus (Chapter IV). The temporal distance between phytohormone signalling and secondary metabolic effects implies interwoven regulation mechanisms, disguising and disrupting direct causal connections between these both events.

Hypotheses about involvement of signal compounds in signal transmission that resulted from data analysis introduced in Chapter IV thus require further investigation. To confirm the involvement of a signal in response to an environmental stimulus a time series with narrow intervals directly after the exposure to the stimulus would be an appropriate strategy also for studies on long-term responses. The advantage of using the methodology that was introduced in Chapter IV is to narrow down the effort to analysis of the most promising pathways.

Further, in contrast to the problems discussed so far, other great advantages became obvious from the simultaneous application of both profiling methods. Many experiments in investigations about plant hormone action and effects are done applying phytohormones in excess, or studying the behaviour of specific mutants (Chapter IV). Therefore, the effect of a reduced phytohormone level, for instance, would be difficult to predict. Application of the presented methods provides the opportunity also to investigate the effects of 'physiologically' decreased hormone concentration, which means naturally provoked decrease of phytohormone concentrations in response to a stimulus. Effectively, the *native* balancing of a plant's metabolic system in reaction to changes in the level of signal molecules can be monitored, in contrast to 'artificially' raised physiological disturbances by hormone exposure. In conclusion, a methodolgy is propsosed that enables further investigation on signal-metabolome interactions as formulated as the initial task of this project.

Outlook

In order to provide an optimised platform for phytohormone analysis, the extraction and pre-purification of the analytes from plant tissue requires further method development. Often, the sample preparation procedure is laborious or the method uses advanced equipment (Moritz and Olsen 1995; Prinsen et al. 1998; Schmelz et al. 2003). A promising approach seems to be solid phase extraction, which enables fast and efficient purification of trace compounds (Chen et al. 1988; Dobrev and Kaminek 2002; Mueller et al. 2002).

State-of-the-art phytohormone analysis with GC-MS uses stable-isotope labelled standards for quantification. The strategy of stable-isotope *in vivo* labelling (Chapter III) offers a potential source of labelled standards for all kinds of signal compounds. This method has to be tested for its usability in phytohormone analysis. Establishment of a reproducible

technology for the production of labelled reference material would be the first step toward establishment of quantitative routine analysis.

For joint analysis of different profiling data, as introduced for phytohormones and metabolites in Chapter IV, future efforts aim at optimisation of experimental designs. The co-evaluation of data acquired with metabolite- and phytohormone-profiling methods faces the difficulty to combine time-resolved events to each other. The major changes in hormonal balance are expected to precede changes on metabolic level; the latter will probably also appear in a particular order. Experimental setups intending to apply both analytical methods, must take into account these expected inconsistencies; their design should be appropriate to follow the time course from signal rise to metabolic appearance. On that score, the gradual exposure of plants to short-term environmental stimuli seems to be promising rather than, for example, investigations of plants with constitutively expressed transformations; in these, direct effects in response to signalling are difficult to dissect from pleiotropic effects.

The concrete extent, or the scale, of chronological resolution from signal rise to metabolic appearance in a particular physiological situation will be, however, subject to empirical studies in the first instance. The collection of a sufficient number of data would be a first step leading to refined hypotheses about conditional or, alternatively, general coherences between signals and metabolites. The use of statistical methods and models may assist in formulation of these hypotheses. In particular, that kind of data might enable the potential of deducting from the ‘metabolic phenotype’ to phytohormone action, or even indicate specific markers for particular phytohormones.

The transmission of ‘information’ from stimulus to (metabolic) phenotype includes various types of regulating events. Other profiling platforms could be applied to study the intermediate regulation levels between signalling and metabolic effects. Alternatively, interesting paths could be first selected and subjected subsequently to detailed studies on linking elements in transcriptome and proteome. That way, the proposed strategy (Chapter IV) may assist to focus analytical efforts efficiently only on genuine effects occurring in response to the signal under investigation.

References

Abeles FB, Morgan PW, Saltveit ME (1992) The Mechanisms of Ethylene Action. In: Ethylene in plant biology. *Academic Press inc.; San Diego, California*, 246.

Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. *Nature* **422** (6928), 198-207.

Agelopoulos NG, Hooper AM, Maniar SP, Pickett JA, Wadhams LJ (1999) A Novel Approach for Isolation of Volatile Chemicals Released by Individual Leaves of a Plant in situ. *Journal of Chemical Ecology* **25** (6), 1411-1425.

Arigita L, Tamés RS, González A (2004) Ethylene biosynthesis and endogenous polyamines in relation to development of in vitro cultured kiwifruit explants. *Functional Plant Biology* **31** (6), 603-609.

Ausloos P, Clifton CL, Lias SG, Mikaya AI, Stein SE, Tchekhovskoi DV, Sparkman OD, Zaikin V, Zhu D (1999) The critical evaluation of a comprehensive mass spectral library. *Journal of the American Society for Mass Spectrometry* **10** (4), 287-299.

Badenoch-Jones J, Summons RE, Rolfe BG, Letham DS (1984) Phytohormones, rhizobium mutants, and nodulation in legumes: 4. Auxin metabolites in pea root-nodules. *Journal of Plant Growth Regulators* **3** (1), 23-39.

Bagni N, Serafini-Fracassini D (1974) The role of polyamines as growth factors in higher plants and their mechanism of action. In: Plant Growth Substances 1973. *Hirokawa Publishing Co., Tokyo*

Barabasi AL, Oltvai ZN (2004) Network biology: understanding the cell´s functional organisation. *Nature Reviews Genetics* **5** (2), 101-113.

Bennet-Clark TA, Kefford NP (1953) Chromatography of the growth substances in plant extracts. *Nature* **171**, 645-648.

Berger S, Weichert H, Porzel A, Wasternack C, Kuehn H, Feussner I (2001) Enzymatic and non-enzymatic lipid peroxidation in leaf development. *Biochimica et Biophysica Acta* **1533** (2), 266-276.

Berger S (2002) Jasmonate-related mutants of *Arabidopsis* as tools for studying stress signalling. *Planta* **214** (2), 497-504.

Bieleski RL (1964) The problem of halting enzyme action when extracting plant tissues. *Analytical Biochemistry* **9** (4), 431-442.

Birkemeyer C, Kolasa A, Kopka J (2003) Comprehensive chemical derivatization for gas chromatography–mass spectrometry-based multi-targeted profiling of the major phytohormones. *Journal of Chromatography* A **993** (1-2), 89-102.

Birkemeyer C, Luedemann A, Wagner C, Erban A, Kopka J (2005) Metabolome analysis: the potential of *in vivo* labeling with stable isotopes for metabolite profiling. *Trends in Biotechnology* **23** (1), 28-33.

Blau K, JM Halket (Eds.) (1993) Handbook of Derivatives for Chromatography, John Wiley & Sons, Chichester.

Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology* **31**, 911-917.

Bray EA, Bailey-Serres J, Weretilnyk E (2000) Responses to Abiotic Stress. IN: Buchanan BB, Gruissem W, Jones RL, (eds.) Biochemistry and molecular biology of plants. *American Society of Plant Physiologists, Rockville, Maryland*, 1158.

Brian PW, Elson GW, Hemming HG, Radley M (1954) The plant-growth promoting properties of gibberellic acid, a metabolic product of the fungus *Gibberella fujikuroi. Journal of the Science of Food and Agriculture* **5** (12), 602-612.

Buchanan B, Gruissem W, Jones R (eds.) (2000) Biochemistry and molecular biology of plants. *American Society of Plant Physiologists, Rockville, Maryland*

Buchholz A, Hurlebaus J, Wandrey C, Takors R (2002) Metabolomics: quantification of intracellular metabolite dynamics. *Biomolecular Engineering* **19** (1), 5-15.

Castrillo JI, Hayes A, Mohammed S, Gaskell SJ, Oliver SG (2003) An optimized protocol for metabolome analysis in yeast using direct infusion electrospray mass spectrometry. *Phytochemistry* **62** (6), 929-937.

Chen KH, Miller AN, Patterson GW, Cohen JD (1988) A rapid and simple procedure for purification of indole-3-acetic acid prior to GC-SIM-MS analysis. *Plant Physiology* **86**, 822-825.

Christensen B, Nielsen J (2000) Metabolic network analysis. *Advances in Biochemical Engineering and Biotechnology* **66**, 209-231.

Colebatch G, Desbrosses G, Ott T, Krusell L, Kloska S, Kopka J, Udvardi MK (2004) Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in Lotus japonicus. *The Plant Journal* **39** (4), 487-512.

Corey EJ, Venkateswarlu A (1972) Protection of hydroxyl groups as tert-butyldimethylsilyl derivatives. *Journal of the American Chemical Society* **94** (17), 6190-6193.

Crawford NM, Kahn ML, Leustek T, Long SR (2000) Nitrogen and sulphur. IN: Buchanan BB, Gruissem W, Jones RL, (eds.) Biochemistry and molecular biology of plants. *American Society of Plant Physiologists, Rockville, Maryland*, 786-849.

Croker SJ, Gaskin P, Hedden P, MacMillan J, MacNail KAG (1994) Quantitative analysis of gibberellins by isotope-dilution mass-spectrometry – a comparison of the use of calibration curves, an isotope-dilution fit programm and arithmetical correction of isotope ratios. *Phytochemical Analysis* **5** (2), 74-80.

Croker SJ, Hedden P (2000) Analysis of gibberellins. *Methods in Molecular Biology* **141**, 93-99.

Davies PJ (Ed.) (1995) Plant Hormones. *Kluwer Academic Publishers, Dordrecht, Netherlands*

Davuluri RV, Sun H, Palaniswamy SK, Matthews N, Molina C, Kurtz M, Grotewold E (2003) AGRIS: *Arabidopsis* Gene Regulatory Information Server, an information resource of *Arabidopsis* cis-regulatory elements and transcription factors. *BMC Bioinformatics*. **4** (1), 25-35.

De Koning W, Van Dam K (1992) A method for the determination of changes of glycolytic metabolites in yeast on a sub-second time scale using extraction at neutral pH. *Analytical Biochemistry* **204** (1), 118-123.

DeBoer DL, Duke SH (1982) Effects of sulfur nutrition on nitrogen and carbon metabolism in lucerne (*Medicago sativa l.*). *Physiologia Plantarum* **54**, 343-350.

Dobrev PI, Kaminek M (2002) Fast and efficient separation of cytokinins from auxin and abscisic acid and their purification using mixed-mode solid-phase. *Journal of Chromatography A* **950** (1-2), 21-29.

Edlund A, Ekloef S, Sundberg B, Moritz T, Sandberg G (1995) A microscale technique for gas-chromatography mass-spectrometry measurements of picogram amounts of indole-3-acetic-acid in plant tissues. *Plant Physiology* **108** (3), 1043-1047.

Ekloef S, Åstot C, Moritz T, Blackwell J, Olsson O, Sandberg G (1996) Cytokinin metabolites and gradients in wild type and transgenic tobacco with moderate cytokinin overproduction. *Physiologia Plantarum* **98** (2), 333-344.

Epstein E, Cohen J (1981) Microscale Preparation of Pentafluorobenzyl Esters. Electron-Capture Gas Chromatographic Detection of Indole-3-Acetic Acid from Plants. *Journal of Chromatography* **209** (3), 413-420.

Fiehn O, Kopka J, Doermann P, Altmann T, Trethewey RN, Willmitzer L (2000b) Metabolite profiling for plant functional genomics. *Nature Biotechnology* **18** (11), 1157-1161.

Fiehn O, Kopka J, Trethewey RN, Willmitzer L (2000a) Identification of uncommon plant metabolites based on calculation of elemental compositions using gas

chromatography and quadrupole mass spectrometry. *Analytical Chemistry* **72** (15), 3573-3580.

Fiehn O (2002) Metabolomics – the link between genotypes and phenotypes. *Plant Molecular Biololgy* **48** (1-2), 155-171.

Fischer E, Sauer U (2003) Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC-MS. *European Journal of Biochemistry* **270** (5), 880-891.

Fischer RA (1936) The use of multiple measurement in taxonomic problems. *Annals of Eugenics* **7**, 179-188.

Forster J, Famili I, Fu P, Palsson BO, Nielsen J (2003) Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. *Genome Research* **13** (2), 244-253.

Gaskin P, MacMillan J (1991) GC-MS of Gibberellins and Related Compounds: Methodology and a Library of Spectra. *Cantocks Enterprises, Bristol, UK*.

Gibon Y, Blaesing OE, Hannemann J, Carillo P, Hohne M, Hendriks JHM, Palacios N, Cross J, Selbig J, Stitt M (2004) A robot-based platform to measure multiple enzyme activities in *Arabidopsis* using a set of cycling assays: Comparison of changes of enzyme activities and transcript levels during diurnal cycles and in prolonged darkness. *The Plant Cell* **16** (12), 3304-3325.

Glassbrook N, Beecher C, Ryals J (2000) Metabolic profiling on the right path. *Nature Biotechnology* **18** (11), 1142-1143.

Gollub J, Ball CA, Binkley G, Demeter J, Finkelstein DB, Hebert JM, Hernandez-Boussard T, Jin H, Kaloper M, Matese JC, Schroeder M, Brown PO, Botstein D, Sherlock G (2003). The Stanford Microarray Database: data access and quality assessment tools. *Nucleic Acids Research* **31** (1), 94-96.

Goodacre R, Vaidyanathan S, Bianchi G, Kell DB (2002) Metabolic profiling using direct infusion electrospray ionisation mass spectrometry for the characterisation of olive oils. *Analyst* **127** (11), 1457-1462.

Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB (2004) Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in Biotechnology* **22** (5), 245-252.

Grove MD, Spencer GF, Rohwedder WK, Mandava N, Worley JF, Warthen JD, Steffens GL, Flippenanderson JL, Cook JC (1979) Brassinolide, a plant growth-promoting steroid isolated from *Brassica napus* pollen. *Nature* **281** (5728), 216-217.

Gullberg J, Jonsson P, Nordstrom A, Sjostrom M, Moritz T (2004) Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of *Arabidopsis* thaliana samples in metabolomic studies with gas chromatography/mass spectrometry. *Analytical Biochemistry* **331** (2), 283-295.

Guo Z, Zhang QC, Zou HF, Guo BC, Ni JY (2002) A method for the analysis of low-mass molecules by MALDI-TOF mass spectrometry. *Analytical Chemistry* **74** (7), 1637-1641.

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology* **17** (10), 994-999.

Hammond-Kosack K, Jones JDG (2000) Responses to Plant Pathogens. IN: Buchanan BB, Gruissem W, Jones RL, (eds.) Biochemistry and molecular biology of plants. *American Society of Plant Physiologists, Rockville, Maryland*, 1102.

Harada K, Fukusaki E, Bamba T, Kobayashi A (2004) In-vivo 15N-enrichment of metabolites in *Arabidopsis* cultured cell T87 and its application for metabolomics. Third international congress on plant metabolomics, #23.

Harrigan GG, Goodacre R (edts). (2003) Metabolic Profiling: Its role in Biomarker discovery and gene function analysis. *Kluwer Academic Publishers, Dordrecht, Netherlands*

Hawkesford MJ (2000) Plant responses to sulphur deficiency and the genetic manipulation of sulphate transporters to improve S-utilization efficiency. *Journal of Experimental Botany* **51** (342), 131-138 MP Special Issue.

Hellerstein MK (2003) *In vivo* measurement of fluxes through metabolic pathways: The missing link in functional genomics and pharmaceutical research. *Annual Review Nutrition* **23**, 379-402.

Hesse H, Nikiforova VJ, Gakière B, Hoefgen R (2004) Molecular analysis and control of cysteine biosynthesis: integration of nitrogen and sulfur metabolism. *Journal of Experimental Botany* **55** (401), 1283-1292.

Hesse H, Trachsel N, Suter M, Kopriva S, von Ballmoos P, Rennenberg H, Brunold C (2003) Effect of glucose on assimilatory sulphate reduction in *Arabidopsis thaliana* roots. *Journal of Experimental Botany* **54** (388), 1701-1709.

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database:1999. *Nucleic Acids Research* **27** (10), 297-300.

Hirai MY, Yano M, Goodenowe DB, Kanaya S, Kimura T, Awazuhara M, Arita M, Fujiwara T, Saito K (2004) Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **101** (27), 10205-10210.

Hocart CH, Wong OC, Letham DS, Tay SAB., MacLeod JK (1986) Mass-spectrometry and chromatography of tert-butyldimethylsilyl derivatives of cytokinins bases. *Analytical Biochemistry* **153** (1), 85-96.

Huhman DV, Sumner LW (2002) Metabolic profiling of saponins in Medicago sativa and Medicago truncatula using HPLC coupled to an electrospray ion-trap mass spectrometer. *Phytochemistry* **59** (3), 347-360.

Ilyin SE, Belkowski SM, Plata-Salaman CR (2004) Biomarker discovery and validation: technologies and integrative approaches. *Trends in Biotechnology* **22** (8), 411-416.

Kanehisa M (1997) A database for post-genome analysis. *Trends in Genetics* **13** (9), 375-376.

Kang MJ, Tholey A, Heinzle E (2001) Application of automated matrix-assisted laser desorption/ ionisation time-of-flight mass spectrometry for the measurement of enzyme activities. *Rapid Communications in Mass Spectrometry* **15** (15), 1327-1333.

Kaplan F, Kopka J, Haskell DW, Zhao W, Schiller KC, Gatzke N, Sung DY, Guy CL (2004) Exploring the Temperature-Stress Metabolome of *Arabidopsis*. *Plant Physiology* **136** (4), 4159-4168.

Katona ZF, Sass P, Molnar-Perl I (1999) Simultaneous determination of sugars, sugar alcohols, acids, and amoni acids in apricots by gas chromatography-mass spectrometry. *Journal of Chromatogrphy A* **847** (1-2), 91-102.

Knapp D.R., Handbook of Analytical Derivatization Reactions, John Wiley & Sons, New York, 1979.

Kopka J, Ohlrogge JB, Jaworski JG (1995) Analysis of in Vivo Levels of Acyl-Thioesters with Gas Chromatography/Mass Spectrometry of the Butylamide Derivative. *Analytical Biochemistry* **224** (1), 52-60.

Kopka J, Fernie A, Weckwerth W, Gibon Y, Stitt M (2004) Metabolite profiling in Plant Biology: Platforms and Destinations. *Genome Biology* **5** (6), 109-117.

Kopka J, Schauer N, Krueger S, Birkemeyer C, Usadel B, Bergmuller E, Dormann P, Weckwerth W, Gibon Y, Stitt M, Willmitzer L, Fernie AR, Steinhauser D (2005)

GMD@CSB.DB: the Golm Metabolome Database. *Bioinformatics* **21** (8), 1635-1638.

Kopriva S, Rennenberg H (2004) Control of sulphate assimilation and glutathione synthesis: interaction with N and C metabolism. *Journal of Experimental Botany* **55** (404), Sulphur Metabolism in Plants Special Issue, 1831-1842.

Kowalczyk M, Sandberg G (2001) Quantitative Analysis of Indole-3-Acetic Acid Metabolites in *Arabidopsis*. *Plant Physiology* **127** (4), 1845-1853.

Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research* **30** (1), 325-327.

Letham DS (1963) Purification of factors inducing cell division extracted from plum fruitlets. *Life Sciences* **3**, 152-157.

Letham DS, Singh S, Wong O.C (1991) Mass-spectrometric analysis of cytokinins in plant-tissue: 7. Quantification of cytokinin bases by negative-ion mass-spectrometry. *Journal of Plant Growth Regulation* **10** (2), 107-113.

Lichtenthaler HK (1996) Vegetation stress: An introduction to the stress concept in plants. *Journal of Plant Physiology* **148** (1-2), 4-14.

Little JL (1999) Artifacts in trimethylsilyl derivatization reactions and ways to avoid them. *Journal of Chromatography A* **844** (1-2), 1-22.

Liu BF, Zhong XH, Lu YT (2002) Analysis of plant hormones in tobacco flowers by micellar electrokinetic capillary chromatography coupled with on-line large volume sample stacking. *Journal of Chromatography A* **945** (1-2), 257-265.

Lockhart DJ, Dong HL, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang CW, Kobayashi M, Horton H, Brown EL (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nature Biotechnology* **14** (13), 1675-1680.

Ludemann A, Weicht D, Selbig J, Kopka J (2004) PaVESy: Pathway visualization and editing system. *Bioinformatics* **20** (16), 2841-2844.

Ludewig M, Doerffling K, Koenig WA (1982) Electron-capture capillary gas chromatography and mass spectrometry of trifluoroacetylated cytokinins. *Journal of Chromatography* **243** (1), 93-98.

Maharjan RP, Ferenci T (2003) Global metabolite analysis: the influence of extraction methodology on metabolome profiles of Escherichia coli. *Analytical Biochemistry* **313** (1), 145-154.

Mamer OA (1994) Metabolic profiling – a dilemma for mass spectrometry. *Biological Mass Spectrometry* **23** (9), 535-539.

Marschner H (1995) Mineral nutrition in higher plants. *Academic press limited; London.*

Maruyama A, Saito K, Ishizawa K (2001) ß-Cyanoalanine synthase and cysteine synthase from potato: molecular cloning, biochemical characterization, and spatial and hormonal regulation. *Plant Molecular Biology* **46** (6), 749-760.

Maruyama-Nakashita A, Nakamura Y, Yamaya T, Takahashi H (2004) A novel regulatory pathway of sulfate uptake in *Arabidopsis* roots: implication of CRE1/WOL/AHK4-mediated cytokinin dependent regulation. *The Plant Journal* **38** (4), 779-789.

Mashego MR, Wu L, Van Dam JC, Ras C, Vinke JL, Van Winden WA, Van Gulik WM, Heijnen JJ (2004) MIRACLE: mass isotopomer ratio analysis of U-13C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites. *Biotechnology and Bioengineering* **85** (6), 620-628.

Matuszewski BK, Constanzer ML, Chavez-Eng CM (2003) Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS. *Analytical Chemistry* **75** (13), 3019-3030.

Miersch O, Bohlmann H, Wasternack C (1999) Jasmonates and related compounds from Fusarium oxysporum. *Phytochemistry* **50** (4) 517-523.

Moritz T, Olsen JE (1995) Comparison between high-resolution selected-ion monitoring, selected reaction monitoring, and 4-sector tandem mass-spectrometry in quantitative analysis of gibberellins in milligram amounts of plant tissue. *Analytical Chemistry* **67** (10), 1711-1716.

Motyka V, Faiss M, Strnad M, Kamínek M, Schmülling T (1996) Changes in cytokinin content and cytokinin oxidase activity in response to derepression of ipt gene transcription in transgenic tobacco calli and plants. *Plant Physiology* **112** (3), 1035-1043.

Mueller A, Duechting P, Weiler EW (2002) A multiplex GC-MS/MS technique for the sensitive and quantitative single-run analysis of acidic phytohormones and related compounds, and its application to *Arabidopsis thaliana*. *Planta* **216** (1), 44-56.

Mueller LA, Zhang PF, Rhee SY (2003) AraCyc: a biochemical pathway database for *Arabidopsis*. *Plant Physiology* **132** (2), 453-460.

Nicholson JK, Lindon JC, Holmes E (1999) "Metabonomics": Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **29** (11), 1181-1189.

Nielsen J (2003) It is all about metabolic fluxes. *Journal of Bacteriology* **185** (24), 7031-7035.

Nikiforova VJ, Freitag J, Kempa S, Adamik M, Hesse H, Hoefgen R (2003) Transcriptome analysis of sulfur depletion in *Arabidopsis thaliana*: interlacing of biosynthetic pathways provides response specificity. *The Plant Journal* **33** (4), 633-650.

Nikiforova VJ, Kopka J, Tolstikov V, Fiehn O, Hopkins L, Hawkesford MJ, Hesse H, Hoefgen R (2005*a*) Systems rebalancing of metabolism in response to sulfur deprivation, as revealed by metabolome analysis of *Arabidopsis* plants. *Plant Physiology* **138** (1), 304-318.

Nikiforova VJ, Daub CO, Hesse H, Willmitzer L, Hoefgen R (2005*b*) Integrative gene-metabolite network with implemented causality deciphers informational fluxes of sulphur stress response. *Journal of Experimental Botany* **56** (417), 1887-1896.

Novák O, Tarkowski P, Tarkowská D, Doležal K, Lenobel R, Strnad M (2003) Quantitative analysis of cytokinins in plants by liquid chromatography–single-quadrupole mass spectrometry. *Analytica Chimica Acta* **480** (2), 207–218.

Oliver SG, Winson MK, Kell DB, Baganz F (1998) Systematic functional analysis of the yeast genome. *Trends in Biotechnology* **16** (9), 373-378.

Palni LMS, Summons RE, Letham DS (1983) Mass-spectrometric analysis of cytokinins in plant-tissues: 5. Identification of the cytokinins complex of datura-innoxia crown gall tissue. *Plant Physiology* **72** (3), 858-863.

Pearce G, Strydom D, Johnson S, Ryan CA (1991) A polypeptide from tomato leaves induces wound-inducible proteinase-inhibitor proteins. *Science* **253** (5022), 895-898.

Pengelly WL, Meins F (1977) A specific radioimmunoassay for nanogram quantities of the auxin, indole-3-acetic acid. *Planta* **136** (2), 173-180.

Pickett JA, Rasmussen HB, Woodcock CM, Matthes M, Napier JA (2003) Plant stress signalling: understanding and exploiting plant–plant interactions. *Biochemical Society Transactions* **31** (1), 123-127.

Platzner IT (1997) Modern isotope ratio mass spectrometry. *John Wiley Sons Inc., New York, USA*

Pollard M, Ohlrogge J (1999) Testing Models of Fatty Acid Transfer and Lipid Synthesis in Spinach Leaf Using *in Vivo* Oxygen-18 Labeling. *Plant Physiology* 121, 1217-1226.

Prinsen E, Redig P, Strnad M, Galis I, Van Dongen W, Van Onckelen H (1995) Quantifying Phytohormones in Transformed Plants. *Methods in Molecular Biology* **44**, 245-262.

Prinsen E, van Dongen W, Esmans EL, van Onckelen H (1998) Micro and capillary liquid chromatography tandem mass spectrometry: a new dimension in phytohormone research. *Journal of Chromatography A* **826** (1), 25-37.

Prinsen E, Van Laer S, Oeden S, Van Onckelen H (2000) Auxin analysis. *Methods in Molecular Biology* **141**, 49-65.

Prosser IM, Schneider A, Hawkesford MJ, Clarkson DT (1997) Changes in nutrient composition, metabolite concentrations and enzyme activities in spinach in the early stages of S-deprivation. IN: Cram WJ, Kok LJ, Stulen I, Brunold C, Rennenberg H (Eds) (1997) Sulphur metabolism in higher plants. Molecular, ecophysiological and nutritional aspects. *Backhuys Publishers, Leiden, The Netherlands*, 339-342.

R Development Core Team (2004). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*, ISBN 3-900051-07-0.

Raamsdonk LM, Teusink B, Broadhurst D, Zhang NS, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnology* **19** (1), 45-50.

Rademacher W, Graebe JE (1984) Isolation and analysis by gas-liquid-chromatography of auxins, gibberellins, cytokinins, and abscisic acid from a single sample of plant material. *Berichte der Deutschen Botanischen Gesellschaft* **97** (1-2), 75-85.

Rhee SY, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A, Garcia-Hernandez M, Huala E, Lander G, Montoya M, Miller N, Mueller LA, Mundodi S, Reiser L, Tacklind J, Weems DC, Wu Y, Xu I, Yoo D, Yoon J, Zhang P (2003) The *Arabidopsis* Information Resource (TAIR): a model organism database providing a centralized, curated gateway to *Arabidopsis* biology, research materials and community. *Nucleic Acids Research* **31** (1), 224-228.

Rivier L, Crozier A (Eds.) (1987) Principles and Practice of Plant Hormone Analysis. *Acad. Press, London*.

Roessner U, Luedemann A, Brust D, Fiehn O, Linke T, Willmitzer L, Fernie AR (2001) Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *The Plant Cell* **13** (1), 11-29.

Roessner U, Wagner C, Kopka J, Trethewey RN, and Willmitzer L (2000) Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *The Plant Journal* **23** (1), 131-142.

Roessner-Tunali U, Hegemann B, Lytovchenko A, Carrari F, Bruedigam C, Granot D, Fernie AR (2003) Metabolic profiling of transgenic tomato plants overexpressing hexokinase reveals that the influence of hexose phosphorylation diminishes during fruit development. *Plant Physiology* **133** (1), 84-99.

Saftner RA, Mehta AM (1990) 1-aminocyclopropane-1-carboxylic acid. Transport, ethylene production, and polyamine interactions. IN: Flores HE, Arteca RN, and Shannon JC (Eds.) (1990) Polyamines and ethylene: biochemistry, Physiology and interactions. *American Society of Plant Physiologists, Rockville, Maryland,* 267-277.

Sano H, Seo S, Koizumi N, Niki T, Iwamura H, and Ohashi Y (1996) Regulation by cytokinins of endogenous levels of jasmonic and salicylic acids in mechanically wounded tobacco plants. *Plant and Cell Physiology* **37** (6), 762-769.

Sasaki Y, Asamizu E, Shibata D, Nakamura Y, Kaneko T, Awai K, Amagai M, Kuwata C, Tsugane T, Masuda T, Shimada H, Takamiya X, Ohta H, and Tabata S (2001) Monitoring of methyl jasmonate-responsive genes in *Arabidopsis* by cDNA macroarray: Self-activation of jasmonic acid biosynthesis and crosstalk with other phytohormone signaling pathways. *DNA Research* **8** (4), 153-161.

Sauer U, DR, Fiaux J, Hochuli M, Glaser R, Szyperski T, Wüthrich K, Bailey JE (1999) Metabolic flux ratio analysis of genetic and environmental modulations of Escherichia coli central carbon metabolism. *Journal of Bacteriology* **181** (21), 6679-6688.

Sauer U (2004) High-throughput phenomics: experimental methods for mapping fluxomes. *Current Opinion in Biotechnology* **15** (1), 58-63.

Schauer N, Steinhauser D, Strelkov S, Schomburg D, Allison G, Moritz T, Lundgren K, Roessner-Tunali U, Forbes MG, Willmitzer L, Fernie AR, Kopka J (2005) GC-MS libraries for the rapid identification of metabolites in complex biological samples. *FEBS Letters* **579** (6), 1332-1337.

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene-expression patterns with a complementary-DNA microarray. *Science* **270** (5235), 467-470.

Schlenk H, Gellerman JL (1960) Esterification of fatty acids with diazomethane on a small scale. *Analytical Chemistry* **32** (11), 1412-1414.

Schmelz EA, Engelberth J, Alborn HT, O'Donnell P, Sammons M, Toshima H, Tumlinson JH (2003) Simultaneous analysis of phytohormones, phytotoxins, and volatile organic compounds in plants. *Proceedings of the National Academy of Sciences of the United States of America* **100** (18), 10552-10557.

Schneider G, Schmidt J (1996) Liquid chromatography electrospray ionization mass spectrometry for analysing plant hormone conjugates. *Journal of Chromatography A* **728** (1-2), 371-375.

Sembdner G, Schneider G, Schreiber K (1987) Methoden zur Pflanzenhormonanalyse, *Springer-Verlag Berlin*.

Skirycz A, Reichelt M, Birkemeyer C, Rolcik J, Kopka J, Zanor MI, Gershenzon J, Strnad M, Szopa J, Mueller-Roeber B, Witt I (2005) Dof Transcription Factor AtDof1.1 (OBP2) exhibits vascular tissue specific expression and modulates glucosinolate accumulation in *Arabidopsis thaliana*. Submitted to *The Plant Journal*

Slocum RD, Weinstein LH (1990) Stress-induced Put accumulation as a mechanism of ammonia detoxification in cereal leaves. IN: Flores HE, Arteca RN, and Shannon JC (eds.) (1990) Polyamines and ethylene: biochemistry, Physiology and interactions. *American Society of Plant Physiologists, Rockville, Maryland,* 157-165.

Smeekens S (2000) Sugar-induced signal transduction in plants. *Annual Reviews in Plant Physiology and Plant Molecular Biology* **51**, 49-81.

Smith MA, Davis PJ (1985) Separation and quantitation of polyamines in plant tissue by high performance liquid chromatography of their dansyl derivates. *Plant Physiology* **78** (1), 89-91.

Soga T, Ueno Y, Naraoka H, Ohashi Y, Tomita M, Nishioka T (2002) Simultaneous determination of anionic intermediates for Bacillus subtilis metabolic pathways by capillary electrophoresis electrospray ionization mass spectrometry. *Analytical Chemistry* **74** (10), 2233-2239.

Specht W, Tillkes M (1980) Gas-Chromatographic Determination of Pesticide-Residues after Cleanup by Gel-Permeation Chromatography and Mini-Silica Gel-Column Chromatography. 3. Communication - Cleanup of Foods and Feeds of Vegetable and

Animal Origin for Multiresidue Analysis of Fat-Soluble and Water-Soluble Pesticides. *Fresenius Journal of Analytical Chemistry* **301** (4), 300-307.

Stein SE (1999) An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. *Journal of the American Society for Mass Spectrometry* **10** (9), 770-781.

Steinhauser D, Usadel B, Luedemann A, Thimm O, Kopka J (2004) CSB.DB: a comprehensive systems-biology database. *Bioinformatics* **20** (18), 3647-3651.

Sumner LW, Mendes P, Dixon RA (2003) Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry* **62** (6), 817-836.

Szyperski T (1998) 13C-NMR, MS and metabolic flux balancing in biotechnology research. *Quaterly Reviews of Biophysics* **31**, 41–106.

Taiz L, Zeiger E (1998) Chapter 22: Ethylene. IN: Plant Physiology <dt.>. *Spektrum Akademischer Verlag Heidelberg, Berlin* (2000), 649-667. Original edition: *Sinauer Associates, Inc., USA*.

Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal* **37** (6), 914-939.

Tokimatsu T, Sakurai N, Suzuki H, Ohta H, Nishitani K, Koyama T, Umezawa T, Misawa N, Saito K, Shibata D (2005) KaPPA-View. A Web-Based Analysis Tool for Integration of Transcript and Metabolite Data on Plant Metabolic Pathway Maps. *Plant Physiology* **138** (3)**,** 1289-1300.

Trethewey RN, Krotzky AJ, Willmitzer L (1999) Metabolic profiling: A rosetta stone for genomics? *Current Opinion in Plant Biology* **2** (2), 83-85.

Trewavas AJ (1982) Growth substance sensitivity: the limiting factor in plant development. *Physiologia Plantarum* **55** (1), 60-72.

Trewavas AJ (2000) Signal perception and transduction. IN: Buchanan BB, Gruissem W, Jones RL, (eds.) Biochemistry and molecular biology of plants. *American Society of Plant Physiologists, Rockville, Maryland*, 930-987.

Unlu M, Morgan ME, Minden JS (1997) Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. *Electrophoresis* **18** (11), 2071-2077.

Urbanczyk-Wochniak E, Luedemann A, Kopka J, Selbig J, Roessner-Tunali U, Willmitzer L, Fernie AR (2003) Parallel analysis of transcript and metabolic profiles: a new approach in systems biology. *EMBO Reports* **4** (10), 989-993.

Van den Bergh G, Arckens L (2004) Fluorescent two-dimensional difference gel electrophoresis unveils the potential of gel based proteomics. *Current Opinion in Biotechnology* **15** (1), 38-43.

Wagner C, Sefkow M, Kopka J (2003) Construction and application of a mass spectral and retention time index database generated from plant GC/EI-TOF-MS metabolite profiles. *Phytochemistry* **62** (6), 887-900.

Weckwerth W, Tolstikov V, Fiehn O. (2001) *Proceedings of 49th ASMS Conference on Mass Spectrometry and Allied Topics*, ASMS, Chicago, 1-2.

Weckwerth W, Wenzel K, Fiehn O (2004) Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks. *Proteomics* **4** (1), 78-83.

Weiler EW (1979) Radioimmunoassay for the determination of free and conjugated abscisic acid. *Planta* **144** (3), 255-263.

Weiler EW (1980) Radioimmunoassay for trans-zeatin and related cytokinins. *Planta* **149** (2), 155-162.

Weiler EW, Wieczorek U (1981) Determination of femtomol quantities of gibberellic acid by radioimmunoassay. *Planta* **152** (2), 159-167.

Weiler EW (1983) Immunoassay of Plant Constituents. *Biochemical Society Transactions* **11** (3), 485-495.

Weiler EW (1984) Immunoassay of Plant Growth Regulators. *Annual Reviews in Plant Physiology* **35** (1), 85-95.

Went FW (1928) Wuchstoff und Wachstum. *Recueil des Travaux Botaniques Néerlandais* **25**, 1-116.

Werner T, Motyka V, Strnad M, Schmuelling T (2001) Regulation of plant growth by cytokinin. *Proceedings of the National Academy of Sciences of the United States of America* **98** (17), 10487-10492.

Weyers JDB, Paterson NW (2001) Plant hormones and the control of physiological processes. *New Phytologist* **152** (3), 375-407.

Wiechert W (2001) 13C metabolic flux analysis. *Metabolic Engineering* **3** (3), 195-206.

Wittmann C (2002) Metabolic flux analysis using mass spectrometry. *Advances in Biochemical Engineering and Biotechnology* **74**, 39-64.

Wittmann C, Heinzle . (2001) Application of MALDI-TOF MS to lysine-producing Corynebacterium glutamicum. A novel approach for metabolic flux analysis. *European Journal of Biochemistry* **268** (8), 2441-2455.

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. *Arabidopsis* Microarray Database and Analysis Toolbox. *Plant Physiology* **136** (1), 2621-2632.

Links to databases

AtcisDB: http://arabidopsis.med.ohio-state.edu/

CSB.DB: http://csbdb.mpimp-golm.mpg.de/

GENEVESTIGATOR: https://www.genevestigator.ethz.ch/

KEGG: http://www.genome.ad.jp/kegg/

PaVESy: http://pavesy.mpimp-golm.mpg.de/PaVESy.htm

PLACE: http://www.dna.affrc.go.jp/PLACE/

PlantCARE: http://intra.psb.ugent.be:8080/PlantCARE/

PlantPAG: http://plantpag.mpimp-golm.mpg.de/index.html

R: http://www.R-project.org

SMD: http://genome-www5.stanford.edu//

TAIR: http://arabidopsis.org//

Supplementary data

Supplementary file 1

The data file, phytoh.msp [a], contains mass spectra of all phytohormone derivatives in Tables I-2 – I-4 including Quadrupole and ion trap electron impact ionisation mass spectra.

The spectrum name was designed to allow sorting according to the reference substance. For example, the name ABA TBS1#EI#Q#MTBSTFA[b] codes for the name of the reference substance and type of derivative, mode of ionisation, mode of mass spectral detection and reagent. The spectrum ID allows sorting according to reagent, mode of ionisation, type of mass spectral detection device, and the file name of the source chromatogram, for example MTBSTFA#EI#Q#1235DW21_37.655.

[a] The file format *.msp can be imported into NIST98 or AMDIS software (AMDIS software can be downloaded from http://chemdata.nist.gov/mass-spc/Srch_v1.7/index.html and http://chemdata.nist.gov/mass-spc/amdis/, respectively).

[b] EI, electron impact; Q, quadrupole technology; T, ion trap technology; MTBSTFA, N-methyl-N-(tert.-butyldimethylsilyl)-trifluoroacetamide reagent; MSTFA, N-methyl-N-(trimethylsilyl) trifluoroacetamide

Zusammenfassung

Die instrumentelle Analytik stellt mit ihrem unschätzbaren Methodenreichtum Analysenwerkzeuge zur Verfügung, die seit ihrem Einzug in die Biologie die Aufzeichnung immer komplexerer ‚Momentaufnahmen' biologischer Systeme ermöglichen. Konkret hervorzuheben wären dabei vor allem multiselektive Methoden wie die Genexpressionsprofilanalyse, die Proteinprofilanalyse und die Metabolitprofilanalyse, die jeweils sehr verschiedene Plattformen der instrumentellen Analytik nutzen können.

Für die Auswertung derart komplexer Daten müssen nun auch entsprechende Prozessierungsmethoden zur Verfügung gestellt werden. Das neu entstandene Fachgebiet der Systembiologie erarbeitet deshalb Strategien zum Sammeln, Auswerten und Archivieren komplexer Daten, um dieses gesammelte Wissen in Form von Datenbanken, Modellen und Netzwerken der allgemeinen Nutzung zugänglich zu machen.

In diesem Kontext wurde den vorhandenen Profilanalysen eine multiselektive Methode zur Erfassung von Pflanzenhormonen hinzugefügt (Kapitel I). Im Hinblick auf eine mögliche Kopplung von Hormon- und Metabolitprofilanalyse wurde die analytische Robustheit der Metabolitprofilanalyse untersucht; als Ergebnis erscheint eine Verbindung beider Methoden durchaus möglich (Kapitel II). In Kapitel III wird das Potential einer innovativen Normalisierungsstrategie für die quantitative Profilanalyse biologischer Proben mit Hilfe der Massenspektrometrie erläutert (*in vivo* labelling mit stabilen Isotopen).

Hormon- und Metabolitprofilanalyse wurden dann parallel angewandt, um den Zusammenhang zwischen der Konzentration von Signalkomponenten und der metabolischen Ausprägung anhand von Keimlingen der Modellpflanze *Arabidopsis thaliana* unter Sulfatmangel zu untersuchen (Kapitel IV). Es wurde eine Prozessierungsmethode entwickelt, die es auf einfache Art und Weise erlaubt, Daten oder Komponenten verschiedenen Ursprungs wie Signalelemente, Gene und Metabolite, die in biologischen Systemen zeitlich versetzt aktiv oder verändert erscheinen, im gegenseitigen Kontext zu betrachten. Die abschließende Analyse aller Daten fokussiert auf Konditionalität von Signal-Metabolismus Interaktionen.

Acknowledgements

I wish to thank all people who supported me:

Prof. Willmitzer and Dr. Joachim Kopka for giving me the opportunity to carry out this PhD project at the MPI.

Dr Joachim Kopka for long-standing and patient support; for his willing advice and (hopefully) not losing trust in me.

Those who kindly agreed to asses my written work: Prof. Ivo Feussner (Göttingen), Prof. John Pickett (Harpenden), Prof. Elmar Weiler (Bochum).

Members of AG Kopka for accommodating me in great work atmosphere and providing lots of help: Alexander, Alexander, Ines, Dirk, Katrin, Takeshi.

MPI infrastructure groups: IT team (Eckhard, Wilfried and Carsten), Green team (Karin, Helga and Britta), Haustechnik (thanks Toralf, Micha and Harald); Verwaltung and the canteen.

My direct collaborators on different projects: Ania Kolasa, Viktoria Nikiforova, Joachim Fisahn, Heidi Fuss, Ulrike Schmidt, Isabell Witt, Aleksandra Skirycz, Petra Birth, Ellen Zuther, Prof. Birgit Dräger, Malgosia Malgorzata, Dr. Agnieszka Sirko, Cornelia Wagner, Thomas Schmülling, Tomas Werner.

All students and trainees who helped me: Anna Happe-Kramer, Luise von Malotky, Tatjana Schwarz, Julia Broska, Christian Bockisch, Andrea Neuber, Leonore.

All those who also did not let me down all the time: Karl-Stephan Neufeldt, Aimee Galster, Andy Humphrey, Ruth Gordon-Weeks.

And: Änne Eckhardt, Stephan Krueger, Ilse Balbo, Ute Rössner-Tunali, Carsten Müssig, Daniel Weicht, Rainer Hoefgen, Holger Hesse, Nicole Gatzke, Daniela Zoeller, Rafal Kartanowicz, Delia Corol, Prof. Mike Beale, Peter Buchner, ...

Last but surely not least, I wish to thank my family for being patient with me and my fading energy - although my kids for one did not really had another choice -, in particular during the course of writing up, …I love you.