University of Natural Resources and Life Sciences (BOKU), Vienna

Department of Chemistry

Institute of Analytical Chemistry

Head of Institute: Univ.Prof. Dipl.-Ing. Dr.techn. Gerhard Stingeder

Selectivity Considerations for Analytical Method Development in Metabolomics

Submitted by

LE Si Hung, MSc.

Supervisor: Assoc. Prof. Dr. Stephan Hann

This PhD thesis is presented for the fulfillment of the requirements for the degree of Dr.-Studium der Bodenkultur; Lebensmittel- und Biotechnologie at the University of Natural Resources and Life Sciences, Vienna

Acknowledgments

First of all, I would like to thank my supervisor Assoc. Prof. Dr. Stephan Hann for guiding my PhD project during 3 years at the Division of Analytical Chemistry at University of Natural Resources and Life Science, BOKU, Vienna.

I would like to thank Ass. Prof. B.Sc. Ph.D. Tim Causon and Dr. Christina Troyer for the co-supervision and all helpful advice.

I want to thank all members of working group for the supporting and teaching during my time in BOKU. I would like to send thanks to my working group in Hanoi University of Science in Vietnam for all their help and encouragement.

I particularly thank to Kaan for his effort to convert me to vegetarianism, and Michi for translating my abstract to German.

I would like to thank my family for all their unconditional support.

This work was funded by the ASEA-UNINET on behalf of the Austrian Federal Ministry of Science, Vienna Business Agency and EQ BOKU VIBT GmbH are acknowledged for providing mass spectrometry instrumentation.

I hereby declare that I am the sole author of this work; no assistance other than that permitted has been used and all quotes and concepts taken from unpublished sources, published literature or the internet in wording or in basic content have been identified by footnotes or with precise source citations.

Abstract

Metabolomics aims to provide knowledge about metabolism occurring in biological systems based on comprehensive analysis of the metabolomes in various type of samples using appropriate analytical tools. In this work, current challenges in metabolomics including maximizing metabolome coverage and sensitive quantitative analysis of polar metabolite classes are highlighted. Several efforts to overcome these challenges are proposed, tested, validated, and further discussed. The results include a study on the impact of column choice on metabolomics data, and the development of a selective analysis method for important intracellular phosphorylated metabolites.

The first part of the work presents a comparison of several fully-wettable reversed-phase (RP) liquid chromatography (LC) stationary phases. Its purpose is to address the difficulties and potential of currently available separation technologies for extending metabolome coverage in metabolomics. All columns were tested under the same analytical conditions with LC-Time-of-Flight-mass spectrometry (LC-TOFMS), using a wide range of authentic metabolite standards and biotechnologically-relevant yeast cell extracts. Data on total workflow performance considering for both non-targeted screening and differential metabolomics workflows include retention behavior, peak capacity, coverage, and molecular feature extraction repeatability from these columns are presented for both authentic standards and *Pichia pastoris* (yeast) cell extract samples.

The second part of the work addresses challenges in cleanup and enrichment for targeted analysis of metabolite classes; specifically, the phosphorylated primary metabolites related to pathways of glycolysis and pentose phosphate. Following screening of several commercially available metal oxide-based solid phase extractions (SPE) materials, all steps of selective enrichment processes were optimized for 12 phosphorylated compounds, and the simultaneous removal of highly abundant matrix components like organic acids and sugars. The full analytical workflow was then validated with the best performing material and included a procedure satisfying absolute quantification demands using isotopologue dilution strategies for determination of phosphorylated metabolites present in wild type yeast cell extracts.

Zusammenfassung

Ziel der Erforschung des Metabolismus (Metabolomik) ist die Generierung von neuen Erkenntnissen über biologische Systeme. Anhand umfangreicher metabolischer Analysen, mithilfe von geeigneten analytischen Techniken, können in Zellen phänotypische Variationen durch Unterschiede im Metabolom aufgedeckt werden. Das Ziel dieser Dissertation ist es bestimmte analytische Probleme innerhalb der Metabolomik zu verstehen und Herausforderungen in der Metabolismusforschung zu bewältigen.

Der erste Teil der Arbeit umfasst eine Vergleichsstudie von verschiedenen voll benetzbaren Umkehrphasen (RP) in der Flüssigkeitschromatographie (LC). Dabei werden die Problematiken und das Potential der aktuell vorhandenen Säulentechnologien diskutiert, um das erfasste Substanzspektrum in der Metabolomik zu erweitern. Jede chronographische Phase wurde unter denselben analytischen Bedingungen mittels Flüssigkeitschromatographie gekoppelter Flugzeitmassenspektrometrie (LC-TOFMS) getestet. Dafür wurde ein breites Spektrum von authentischen Metabolitenstandards und biotechnologisch relevante Hefezellenextrakte injiziert. Die Kenndaten der Trennmethoden werden anhand von Retentionszeitstabilität, Peakkapazität, Komponentenerfassung und Wiederholbarkeit der Datenextraktion dargestellt. Hierfür müssen die unterschiedlichen Arbeitsweisen von non-targeted Screening und spezifischer metabolomischer Methoden bei authentischen Standards sowie in Zellextraktproben von *Pichia pastoris* (Hefe) berücksichtigt werden.

Im zweiten Teil werden Herausforderungen in der Aufreinigung und Anreichung von spezifischen Metabolitklassen besprochen. Im Fokus stehen dabei phosphorylierte Primärmetaboliten, welche in Stoffwechselwegen wie dem Pentosephosphateweg und Glykolyse vorkommen. Als Ausgangsbasis für die Optimierungen der Reinigungs- und Anreicherungsprozesse wurde zunächst ein Screening von kommerziell erhältlichen metalloxidbasierten Festphasenextraktionsmaterialien durchgeführt. Darauffolgend wurden 12 phosphorylierte Komponenten selektiv angereichert und häufig vorkommende Matrixinhaltsstoffe, wie organische Säuren oder Zucker, simultan

abgetrennt. Der gesamte analytische Arbeitsprozess wurde für das geeignetste Material und jene Methode validiert, welche die geforderte Bestimmungsgrenze der Analyten nach der Probenpräparation erreichte und beinhaltete die Isotoplogenverdünnungsanalyse zu der Bestimmung von phosphorylierten Metaboliten in natürlich vorkommenden Hefezellenextrakten..

TABLE OF CONTENT

1. INTRODUCTION

1.1 Metabolome and Metabolomics

Metabolism is defined as a sum of life-sustaining cellular regulatory processes in living organisms, included biochemical reactions (more than 11,000 reactions with 18,000 annotated metabolites are currently listed in Kyoto Encyclopedia of Genes and Genomes (https://www.genome.jp/kegg/docs/statistics.html)) that convert food to energy for cellular processes, or to build blocks for proteins, or lipids, or to eliminate nitrogenous wastes in organisms [1]. Metabolism modulates protein trafficking, localization and enzyme activity, and can affect cell signaling. Two major categories of metabolism include catabolism and anabolism. Catabolism is the breaking down of organic matter to archive the energy, while anabolism uses the energy to build up of components of proteins and nucleic acids. When one chemical compound follows a series of biochemical reactions of metabolism to become another compound using enzymes, the series can be categorized as a metabolic pathway. For example, glycolysis, pentose phosphate pathway, citric acid cycle, urea cycle, carbon fixation are some important metabolic pathways in the cell. Metabolites are therefore simply the end products of metabolism in a biological system such as in the cell, body fluids, or tissue [2–4], whereby their level can indicate the ultimate response of biological systems to genetic or environmental changes (or phenotypes) [2,5,6]. In another way, the metabolite level in an organism can reflect its character and give us understanding of phenotype. Fiehn *et al.* used term ‘metabolome’ in their yeast metabolism work, in which metabolome can be defined as the full complement of small molecular weights metabolites (< 2 kDa) [7,8]. For example, glutathione is one major component of antioxidant defense, xenobiotic and eicosanoid metabolism, related to thiol metabolism [9]. S-adenosylmethionine is an intermediate formed enzymatically from methionine and adenosinetriphosphate (ATP) [10].

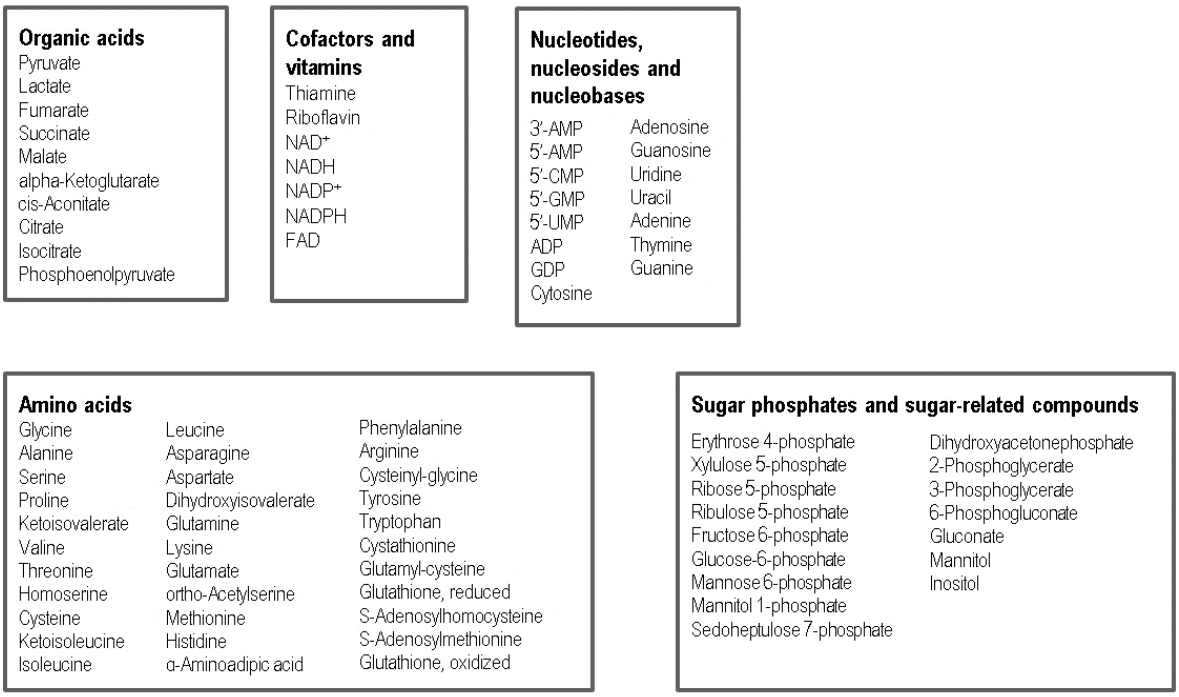

Figure 1. List and categories of some metabolites

Metabolomics was initially defined as the comprehensive (qualitative and quantitative) analysis of the metabolome: the intermediates and products of biological processes small molecule in a wide range of samples including tissue, biofluids or cell extracts [2,4,11–13]. Together with proteomics and genomics, metabolomics studies can reveal the functional of gene, address/diagnostic biological problems in many living organisms [1,14–18] by understanding connections between metabolite variations and changes in gene expression, protein expression and enzyme activity [19–21]. One key issue in metabolomics is to distinguish the phenotypic difference of the cells from metabolome differences [3], and to do this it requires sensitive, quantitative, and rugged analysis methods. However, due to the wide range of metabolite concentrations, compound classes of interest in biological samples (e.g. Fig. 1), it is challenging to develop a "one-ring-to-rule-them-all" analysis method. In other words, none of the individual methods can cover whole range of metabolites present in cells [3].

Some of the main applications of metabolomics focus on diagnostic, functional genomics, systems biology, which rely heavily on analytical data [18,22]. One of first examples reported was the use of metabolite profiling to estimate successfully modes of actions of various herbicides on plant seeds using gas chromatography (GC) [23]. Araníbar *et al.* [24] combined nuclear magnetic resonance (NMR) spectroscopy with bioinformatics to successfully classify herbicides with unknown modes of action. Jones and colleagues also successfully developed metabolite models of cardiac disease through metabolic profiling using NMR [14]. Lutz *et al.* [25] reported a method for determination of specific ^{13}C enrichment phosphorylated glucose metabolites by using ^{13}C-coupled, ^{1}H-decoupled ^{31}P-NMR spectroscopy in tissue. Andreas *et al.* [26] used a multiplatform strategy including ^{1}H-NMR, liquid chromatography mass spectrometry (LC-MS), and capillary electrophoresis mass spectrometry (CE-MS) for comprehensive characterization of metabolic profiles from seventy breast milk samples. Abdersson *et al.* [27] proposed an ion pair reversed-phase (RP)LC method for the quantitative determination of the redox status of homocysteine in plasma of healthy subjects and patients. Guozhu and Yonezawa also reported the use of metabolite profiling of tissue and body fluid samples from head and neck squamous cell carcinoma patients, and effectively discovered biomarkers potentially for disease detection [28,29]. The results from Dang *et al.* [30] using LC-MS showed that the changes in levels of tricarboxylic acid (TCA) cycle intermediates were not related to the presence of mutated cytosolic isocitrate dehydrogenase 1 enzyme. Metabolomics applications also extend to microorganisms such as yeast, which plays an important role in modern biotechnology for the production of foods, enzymes, chemicals, and pharmaceutical reagents (i.e. arming yeast) [31]. Metabolic profiling study from Kondo's group using gas chromatography time of flight mass spectrometry (GC-TOFMS) combined with principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA) was used for the discovery of several potential biomarkers to distinguish several groups of commercial yeast culture samples [32]. Seifar *et al.* [33] reported a method to analyze metabolites from TCA cycle for yeast and bacteria using GC-MS/MS, applying isotope dilution strategies for potential error correction from sample preparation, matrix effects, derivatization and liner contamination.

Indeed, metabolomics is both a highly interesting and a challenging field. Therefore, several analytical platforms have been utilized, and suitable methods have been developed for these challenges. These methods are often targeted toward particular metabolite classes enabling us to obtain better understanding about different metabolomic systems in various types of sample. Additionally, to increase the total coverage of metabolomics studies, so-called “non-targeted” metabolomics strategies have emerged over the last 15-20 years.

1.1.1 Nuclear magnetic resonance spectroscopy- and high-resolution mass spectrometry- based metabolomics

Due to the great diversity in physical and chemical properties and the huge number of cellular metabolites, it is impossible to cover the whole range of compounds even with a combination of many analytical tools. The two most common analytical techniques in metabolomics are NMR and MS due to their high selectivity and ability to analyze hundreds of metabolites in a single measurement [18,34,35].

1.1.2.1 Nuclear magnetic resonance

NMR spectroscopy is the most useful tool to determine the absolute structure for both organic and inorganic compounds. It is a non-destructive technique allowing the analysis of biofluids and tissues without sample separation or preparation. Therefore, NMR is able to trace metabolic pathways and can even measure metabolic fluxes utilizing isotope labeled substrates [36–38]. NMR detects atomic nuclei that spin in presence of a strong external electromagnetic field [7]. If the number of nuclear particles in atom is even, then the spin of atom is zero (no net magnetic moment). Other atoms (e.g. ^{1}H, ^{13}C, ^{19}F, and ^{31}P) have a net spin, which results in a magnetic moment [39]. In the presence of the magnetic field (B_0), 99.99% of nuclei tend to align with the field (parallel spin state). The B_0 causes a small energy difference between parallel (spin-up) and antiparallel (spin-down) spin states (ΔE).

The stronger the B_0, the larger the ΔE. Moreover, the spin-up nuclei also precess around the axis of B_0. By applying a proper radio-frequency (RF) pulse (B_1) perpendicular to B_0 direction, nuclei precessing with same frequency as RF pulse can be excited to the higher energy state (spin-down). Also the net magnetic vectors are shifted away from the axis of B_0 an angle (flip angle θ) [39]. After B_1 is turned off, the nutated nuclei continues to precess around B_0, and generates a small induced current in the receiver coil. Natural abundance and resonance frequency affect the sensitivity of NMR observation. For example, proton (^1H) has the highest sensitivity due to its high natural abundance (~100%) and nearly 4-fold higher Larmor frequency compared to low natural abundance ^{13}C (1.1%). NMR requires a large amount of sample, and has limitations in sensitivity and dynamic range for metabolomics due to the presence of water (inside living tissues) signal. The method also has lower sensitivity for isotopes such as ^{31}P compared to ^1H [34,40,41], which can create a bias toward higher abundance metabolites. One additional drawback of NMR is that the coupling with chromatographic separation is not easy. For on-line coupling of LC with NMR, several attempts were carried out during early 1980s. However, the coupling methods suffered from low sensitivity and high cost due to the using of expensive deuterated solvents (e.g. CCl_4). Recently, a new LC-NMR probe design has become available, and the use of D_2O and acetonitrile (for which the signal could be easily suppressed) as solvents reduced the cost per measurement [42]. The complexity of biological samples, limited resolution, and sensitivity of NMR have restricted both the number and the quantitative accuracy of detected metabolites [34]. In order to overcome current limitations of NMR methods for comprehensive metabolomics measurements, many efforts have been reported. For example, to deal with presence of several high concentration species (e.g. glucose, lactose), which can overlap with a large number of metabolites signals, a method termed “Add-to-Subtract” was proposed [34]. Basically, the subtraction is based on spectra before and after concentrated glucose addition into sample. The combining of 1D/2D NMR experiments, database matching, and authentic compounds spiking allows approximately 70 metabolites in blood to be identified in the NMR spectra of blood [43]. Despite of many efforts to overcome bottlenecks, current NMR application in metabolomics is still out weighted especially for

non-targeted analysis. Therefore, mass spectrometry, which has better metabolome coverage, has become more and more dominant in metabolomics recently.

1.1.2.2 Mass spectrometry

MS is currently the most popular analytical instrumental approach used in metabolomics. Compared to NMR, MS is more flexible and can be coupled with a range of separation platforms and ionization sources relevant for analytical metabolomics. A mass spectrometer is composed of three main parts: ionization source, mass analyzer, and detector. Several ionization techniques can be employed for metabolomics studies depending on the separation mode and analytical goals at hand. However soft ionization techniques, which can generate molecular or pseudomolecular ions (i.e. $[M+H]^+$ or $[M-H]^-$), are very valuable for metabolomics as they allow the determination of accurate mass information for a wider range of metabolites. This is indeed the case for LC-MS approaches, where the most popular ionization techniques are electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), and atmospheric pressure photoionization (APPI). ESI is the most commonly employed soft ionization process for metabolomics studies and involves the following steps. First, sample solution is sprayed into the source chamber to from droplets with help of a drying gas. Then droplets go through a capillary, which has a potential difference of approximately 3000V, and get charged after exit. As soon as the solvent vaporizes, the droplets shrink until they disappear (undergo Coulombic explosion) leaving highly charged analyte molecules. ESI is suitable for wide-range of both charged, polar and moderately non-polar metabolites [13,44,45]. APCI and APPI also create less in-source fragmentation due to the ion activation/cooling under atmospheric conditions allowing production of less energetically charged molecular species [44]. APCI is a gas phase chemical ionization, which uses HPLC mobile phase (e.g. H_2O, methanol) to form chemical ionization reagents. The vaporized mobile phase gas then reacts with electrons from corona discharge to from adduct ions [46]. These ions will transfer protons to the analyte molecules (positive mode). In APPI, the sample solution is vaporized prior entering ionization chamber

at atmospheric pressure, then mixture of solvent and sample molecules is exposed to ultraviolet light (e.g. from krypton lamp which emits photons with energy of 10 eV). These photons have enough energy to ionize the target molecules, but not enough to ionize air and other unwanted molecules. APCI and APPI are used mainly for analysis of uncharged, non-polar, and thermally stable compounds. Therefore, they are less suited for broad metabolomics applications. For gas chromatography (GC)-MS approaches, electron ionization (EI, hard ionization) and chemical ionization (CI, soft ionization) are used. EI is the easiest to perform and is most suitable for routine GC analysis. EI involves a process, in which gaseous analyte molecules entering the source are bombarded with electrons at 70 eV under high vacuum. In this environment, the molecules have excess energy and form many fragments (thus different and more selective fragment ions may be formed). EI has its own drawbacks such as dominance of highly trimethylsilylated phosphate ions in the case of trimethylsilylated phosphorylated metabolites (e.g. at *m/z* 299 and 315) while normally offering low abundance of a molecular ion, as well as similar ionization fragmentation patterns for many metabolite isomers. In comparison, CI is relatively soft and exhibits more selective ionization compared to EI [47]. In positive CI (PCI), the ionization involves the collision of analyte molecule with a reagent gas (e.g. hydrogen, methane, isobutane, ammonia). Methane is suitable as a reagent gas for analytes that have proton affinity (PA) greater than 552 kJ mol^{-1}, and therefore is more popular in CI (isobutene and ammonia are suitable for analytes with PA > 820 kJ mol^{-1} and 854 kJ mol^{-1}, hydrogen requires special care to prevent explosion). While in negative CI (NCI), the ionization first involves bombardment of reagent gas with high energy electrons to produce thermal electrons (with lower energy levels). Then thermal electrons react with sample molecules to create negative charged analyte ions.

Following ionization, the mass analyzer separates the ions according to their *m/z* values. Based on detected ion signals, determination of molecular masses is performed with help of instrument calibration, and can also allow metabolite identity confirmation usually relying on tandem or two-stage mass analyzer systems (i.e. combinations of two or more mass analyzers). In addition to this, signal intensities at the detector can be used for

determining the metabolite concentrations with (absolute quantification) or without (relative quantification) application of external calibration with authentic standards. Currently, four major mass analyzers are used that are relevant to metabolomics: quadrupole, time-of-flight (TOF), orbital ion trap (marketed as "Orbitrap"), and Fourier-transform ion cyclotron resonance (FT-ICR). For each of these types of mass analyzer, the principle of separation, detection approach and performance characteristics vary. One of the most critical parameters for measurement of unknown molecules is mass resolving power. Mass resolving power is a parameter used for broadly categorizing mass analyzers into low-resolution and high-resolution. Specifically, the mass resolving power, which reflects the capacity to separate ions with a small difference in *m/z*, is equals to the ratio of the measured mass (m) to the full width of the peak at half maximum height (Δm) (although more strict models may also be used). For example, low-resolution analyzers such as quadrupole instruments have mass resolving power less than 500 at *m/z* of 250. This level of resolution is suitable for quantitative metabolomics using tandem mass spectrometry (e.g. triple quadrupole MS), whereby analytical selectivity is achieved by implementation of selective mass transitions across the first and third quadrupole with an optimized collision induced fragmentation energy applied in the second quadrupole. While offering low resolution, such instruments offer excellent sensitivity and low limits of detection can be achieved for targeted metabolomics studies. In the case of non-targeted metabolomics, high-resolution mass analyzers (HRMS) are beneficial. For example, TOF instruments have a resolving power ranging from 5,000 to 60,000 (even $R \geq 100{,}000$ at *m/z* of 400 on a multireflecting TOF [48,49]). Orbital ion traps, in which trap ions in stable trajectories between an inner spindle electrode and an outer cylindrical electrode, have mass resolution related to scan time. For example, the resolving power at *m/z* of 400 is 7500, 60,000, and 100,000 with 0.1 s, 1 s, and 1.9 s scan cycle time respectively and can be even higher with longer dwell times [48,50,51]. This means to reach the higher mass resolution, it is need to have a compromise with acquisition speed (number of spectra per second), which becomes one limit of this technology [52]. Moreover, to avoid overfilling the trap (due to space charge effect), Orbitrap measurements may be biased against low abundance ions, thus may affect the apparent isotope pattern of some compounds [23]. Finally, FT-ICR instruments, which can trap ions

in a strong magnetic field combined with a weak electric field, provide very high mass resolving power greater than 500,000 [53]. While FT-ICR-MS also has very high mass accuracy (<1 ppm error in terms of difference between measured *m/z* and true *m/z* of ion), the inability to resolve isomers via chromatographic separation due to the long scan times required, and high cost (due to expensive superconducting magnet and liquid helium for cooling [54]) are the main hindrances for routine non-targeted metabolomics studies.

The advantage of HRMS instruments is not only to provide accurate masses in high-resolution at high mass accuracy, but also enable the analyst to assess metabolite isotopic distribution (or pattern) based on their natural distribution ratios [55,56]. Additional isotopologue ions (M+1, M+2, M+3) will be found together with each monoisotopic molecular ion (denoted as M+0). This set of peaks create particular pattern associated with the relative abundance of the isotopes in the chemical formula [44,57]. Moreover, the charge state can be easily determined based on the distance of isotopologue ions on *m/z* scale. However, interference and saturation effects caused by the matrix and the experimental conditions may also influence the accuracy of isotope ratio measurements. Böcker *et al.* [55] reported using isotope patterns obtained from HR-MS to correctly identify sum formula for >90% of the small molecules with mass less than 1000 Da. Kind and Fiehn [56] showed that accurate interpretation of isotopic abundance patterns helps to remove >95% of false candidate molecule formulas (in range of 0-500 Da) for database annotations. Also the combination of isotopologues and a reasonable mass accuracy (3 ppm) is more effective in annotation than relying only on high mass accuracy.

Recently, application of Orbitrap-MS application has increased in metabolomics studies relying on high resolution mass spectrometry [58]. For example, Liu *et al.* [59] applied an analysis method coupling an amide column with Orbitrap, simultaneously and robustly targeted 168 polar metabolites, and thousands of additional features. Weidt *et al.* [60] proposed a novel targeted/non-targeted GC-Orbitrap method for study of interaction between *Candida albicans* and *Staphylococcus aureus* microorganisms. One advantage of the Orbitrap is the low chemical background noise free due to the fact the background (which caused by stray, or incoherent ions) was broad, smooth, and easily be subtracted by software

[48]. Orbitrap resolution is inversely proportional to the square root of *m/z* (due to frequency of ion oscillation along z-axis is equal to square root of (k*(z/m)), which is favorable for metabolomics where most mass of interest are below 500 g/mol. TOF also remains one of the most popular mass analyzers with extensive application in various type of biological samples (e.g. plant tissue, serum, urine, tissue) [61–64]. In TOF-MS, when ion beam reaches the pulser, the ions are accelerated into a flight tube, which has a defined flight length. The time taken for an ion to reach the detector placed at the end of flight path (flight time), is proportional to the square root of the mass to charge. It means that ions arrive in order of increasing mass, light ions travel faster and have shorter flight time than heavy ions. Using ions of known *m/z* to calibrate the instrument, the flight time of every ion can be converted to a corresponding *m/z* value. Theoretically, ions with same *m/z* value will have the same kinetic energy in the flight tube. However, due to the difference in position of these ions in space at the beginning (broadening of the ion package), the kinetic energies are slightly different, causing the spreading of the ion package in flight time. Hence, to improve the repeatability of flight time and increase the resolution, a reflectron (ion mirror) is always employed to refocus the ion package. Some advantages of TOF are the good selectivity, ruggedness in terms of long-term effects and speed especially enabling the combination with high-throughput or high resolution LC, CE, or GC separations. Furthermore, TOF remains competitive in terms of mass accuracy in range of 1-2 ppm, at the same time gives good resolving power for the whole mass range needed for metabolomics studies [19,44]. While low resolution instruments such as triple quadrupole MS continue retain their place as a mainstay for targeted metabolomics, the field can clearly benefit from the use of HRMS making this technology increasingly popular. However, these approaches are not error-free and some problems still need to be considered such as isomeric or isobaric metabolite separation, in-source fragmentation, signal suppression, instrument noise, chemical background noise and matrix interference [65,66]. Furthermore, ESI or APCI sources are often used in LC-based metabolomics as this combination provides the most coverage with minimal sample preparation. However, when encountering a wide range of metabolites, a clear problem is that the ionization conditions cannot be optimized for every compound. Fragmentation occurring during ionization can create additional fragment ion species,

adducts (such as non-covalent species from alkali metals, or radicals) resulting in misinterpretation during data processing [64,65,67]. Also, in the case of complex biological samples that include a lot of salts, proteins, or for moderately and highly polar metabolites that elute in the void volume, ion suppression/enhancement is known to occur due to poor ionization efficiency, overlapping molecule signals, and extremely poor selectivity [68].

1.1.2 Separation techniques in MS-based metabolomics

As stated earlier, HRMS indeed has many advantages in metabolomics, however it cannot resolve all issues relevant for metabolomics alone. Perhaps the best way to extend its advantages at the same time overcome disadvantages is to combine HRMS with chromatographic or electrophoretic separation techniques. Certainly, applying separation techniques prior to MS detection brings a lot of benefit. First, taking advantage of the additional selectivity of chromatographic separation together with *m/z* selectivity, isomers and other compounds that have identical exact masses can be resolved. Second, multidimensional separation methods also enable the flexibility for dealing with the high diversity in physicochemical properties of the metabolome [26,69]. Lastly, chromatography also helps to remove sample matrix components, and other interfering compounds from complex biological samples due to the fact that retention of target analytes can be achieved while interferences from unwanted components can be avoided [70]. Three major separation techniques that have been widely used for metabolite separation in biological samples with hyphenation with HRMS are capillary electrophoresis (CE), gas chromatography (GC), and liquid chromatography (LC).

1.1.3.1 Capillary electrophoresis

CE is powerful and fast separation tool for highly polar, charged metabolites such as nucleic acids, amino acids, carboxylic acids, and sugar phosphates in various biological

samples [71–73]. In CE, analyte ions from a given sample will migrate through a narrow capillary when an electric field is applied across the capillary tube. The most frequently used capillary tubing material is bare fused silica (BFS). However various coating can be employed at the wall of the capillary to reduce the analyte adsorption and allow fine tuning of the electroendosmotic flow (EOF) such as amines, oligoamines, cationic surfactant, carboxymethyl chitosan, or methacrylate copolymer [74]. Two factors affect the velocity of a migrating ion: electrophoretic mobility, which is the solute's characteristic response to the applied electrical field, and the phenomenon of EOF, which results in the bulk flow of background electrolyte through the capillary in response to applied electrical field. The net migration velocity (i.e. including direction) of the ion is given by the sum of these two contributions. Taking advantage of the difference in intrinsic electrophoretic mobility of compounds, which is dependent on the charge and size of analytes in solution, CE is generally suitable for separation and analysis of polar and charged metabolites [75]. Despite the excellent separation performance, the general drawbacks of CE for metabolomics are the poor sensitivity, migration time variability, and low method ruggedness [75,76]. Despite of all disadvantages, CE is still considered as a quick and cheap method characterized by high separation efficiency without requiring complex sample preparation, and can be preferred when dealing with samples that are volume-limited [77].

1.1.3.2 Gas chromatography

GC-MS was used commonly in early metabolomics studies due to highly efficient separations, sensitivity and reproducibility. On the other hand, GC-MS often requires derivatization for non-volatile (or polar) compounds, and tailored sample preparation is required [78–80]. GC capillary columns were introduced in 1959, whereby the column is not filled with packing material, but rather a thin film of liquid (wax) stationary phase is coated to the inner wall. The fused silica capillary column is given a polyimide coating to provide better flexibility and column lengths of typically up to 60 m [81]. The narrower inner diameters (I.D) of the capillary column, the higher separation efficiency it offers (e.g.

theoretical value calculated for 30 m column, 0.1 mm I.D offers 7300 theoretical plates per meter, while 0.53 mm only offers 1300 plates per meter). Retention in GC is mainly dominated by the vapor pressure of analytes. Nevertheless, there are many variations of GC capillary column chemistry allowing important variation in retention and selectivity. Typical stationary phases used in GC capillary columns are 100% dimethylpolysiloxane, 5% phenylpolysiloxane with 95% dimethylpolysiloxane (for separation of non-polar compounds), 50% phenylpolysiloxane with 50% dimethylpolysiloxane (for mid-polar compounds separation), 14% cyanopropylphenylpolysiloxane with 86% dimethylpolysiloxane (for low/mid-polar compounds separation), and polyethylene glycol (for fatty acid methyl esters separation) [82]. Possible interactions between analytes and the stationary phases including dispersive interaction, dipole interaction, and hydrogen bonding. The mobile phase in capillary GC (carrier gas) normally is inert gas such as helium, nitrogen, or hydrogen. GC-MS can be used for measurement of a wide range of samples relevant for metabolomics such as plants, urine, plasma, and yeast cell extracts [28,32,33,83–91]. GC is favorable method for the analysis of volatile, thermally stable metabolites. Chemical derivatization can also be used to extend the coverage of GC-MS by attaching relative large, non-polar structures to the targeted molecules. The polarity of derivatized molecules is decreased while the volatility is improved [3,47]. For example, alkylsilylation is used to convert problematic functional groups (i.e. hydroxyl, amide, amine, phosphate, thiol groups) to alkylsilyl derivatives. Oximation prior to silylation can prevent cyclization of sugar compounds (less peaks per sugar), and protect keto groups in alpha-ketoacids [3]. However, adding more steps in derivatization will increase total analysis time [84]. Despite these practical challenges, with a suitable, validated derivatization strategy, GC-MS is an extremely powerful technique for metabolomics due to its broad metabolite coverage. GC-MS has the power to resolve many organic acids, amino acids, sugars, and sugar-related compounds with high efficiency and repeatability [92,93]. Furthermore, the combination of GC separation with EI spectra is unsurpassed in terms of identity confirmation potential due to the excellent coverage and reproducibility of GC-MS spectral libraries in metabolomics.

1.1.3.3 Liquid chromatography

In contrast to GC-MS, high-performance LC and ultrahigh-performance LC allow analysis of a wide range of metabolites without complicated sample preparation steps such as derivatization. While a large range of retention modes are available for LC, reversed-phase liquid chromatography (RPLC) remains the most popular mode for practical broad scale metabolome coverage as the retention characteristics of these materials are well-suited to a range of moderately polar and non-polar metabolites [71]. RPLC employs an organic-aqueous mobile phase and a non-polar stationary phase (i.e. silica particles derivatized with hydrophobic chains) in the form of spherical particles (1-5 μm) packed into analytical dimensions (1-5 mm internal diameter). The hydrophobic compounds in the polar mobile phase will be absorbed onto or partitioned into the non-polar stationary phase, while the hydrophilic compounds will experience weaker interactions. The more hydrophobic a compound, the more strongly it will interact with the non-polar stationary phase. By increasing organic solvent gradient, the polarity of mobile phase decreases and the hydrophobic compounds are eluted from the columns. The fact is that polar charged metabolites, which are not retained on reversed-phase column. These compounds often co-elute near to the column void volume and may be affected by signal suppression within the MS ionization and ion transmission processes. For this reason, ion-exchange chromatography (IC) [47,58,94–97] and hydrophilic interaction liquid chromatography (HILIC) have been introduced as alternative separation mechanisms to address this problem for metabolomics [98–102].

In IC, analyte molecules are separated based on the exchanging of analyte ions for species of the same charge (counterions) electrostatically held to chemical groups, which are immobilized on stationary phase. The ion exchanger is usually a base matrix (in a form of porous beads to provide enough adsorption surface), and charged ligands or ionizable functional groups (positively or negatively charged) immobilized onto the matrix [103]. Various chemical groups have been used for different types of ion exchanger such as quaternary ammonium, quaternary aminoethyl, diethylaminoethyl for anion exchange, and sulfopropyl, sulfoethyl, carboxymethyl for cation exchange. Since the neutrality in system

must be preserved, there must be an exchangeable counterion associated with the immobilized groups (e.g. Na^+, H^+ in case of cation exchangers, and OH^-, Cl^- in case of anion exchanger). Changing pH or increasing ionic strength of mobile phase is needed to elute the analytes that bind to stationary phase. A pH change may alter the charge on particular molecules, therefore changes the binding. While increasing the concentration of similar charge species in mobile will compete with analytes in order to bind to stationary phase [104]. As an additional complication for metabolomics, IC coupling with MS requires that a suppressor is placed between column and detector to reduce background conductivity of the eluent. This is desirable for conventional IC measurements and essential for IC-MS as it prevents nonvolatile salts entering the ionization source.

The term "HILIC" broadly covers a combination of hydrophilic stationary phases (e.g. silica, amino, cyano, diol, alkylamide…) with organic-rich eluents. In HILIC, analytes are eluted in order of increasing polarity [105]. HILIC stationary phases include classical bare silica, modified silica gels with polar functional groups, or polymer-based stationary phases with analogous polar chemistry. HILIC mobile phases comprise water-miscible polar organic solvents such as acetonitrile, alcohols, or dioxane with a low amount of water in isocratic mode, or with gradually increasing amount of water in gradient mode. The relative solvent strengths in HILIC are: acetone < isopropanol < acetonitrile < ethanol < dioxane < methanol < water. If the pH and ionic strength of mobile phase are not suitably adjusted, asymmetric peak shape, peak tailing or poor recovery from the stationary phase may be observed [105]. The widely acceptable mechanism of HILIC can be described as a liquid/liquid system. In which the analyte is distributed between a water-rich layer on the surface of polar stationary phase and water-deficient mobile phase [106]. The more hydrophilic the analyte, the more it interacts with immobilized water-layer on the surface of stationary phase, therefore it is more strongly retained. However, the efficiency of HILIC separations also depend on the choice of stationary phases, potential ion exchange effects, separation conditions including sample solvent, mobile phase pH and composition. The popularity of IC and HILIC for metabolomics applications continues to increase due to improvements in column technology and the overriding need for suitable methods covering the polar metabolome.

Despite of all remained obstacles, coupling chromatographic separation with HRMS clearly benefits metabolomics studies. Especially for non-targeted metabolomics, when high mass resolving power alone is not enough to deal with hundreds, thousands co-eluted unknown molecular features.

1.2. Non-targeted metabolomics: chromatography with high-resolution mass spectrometry

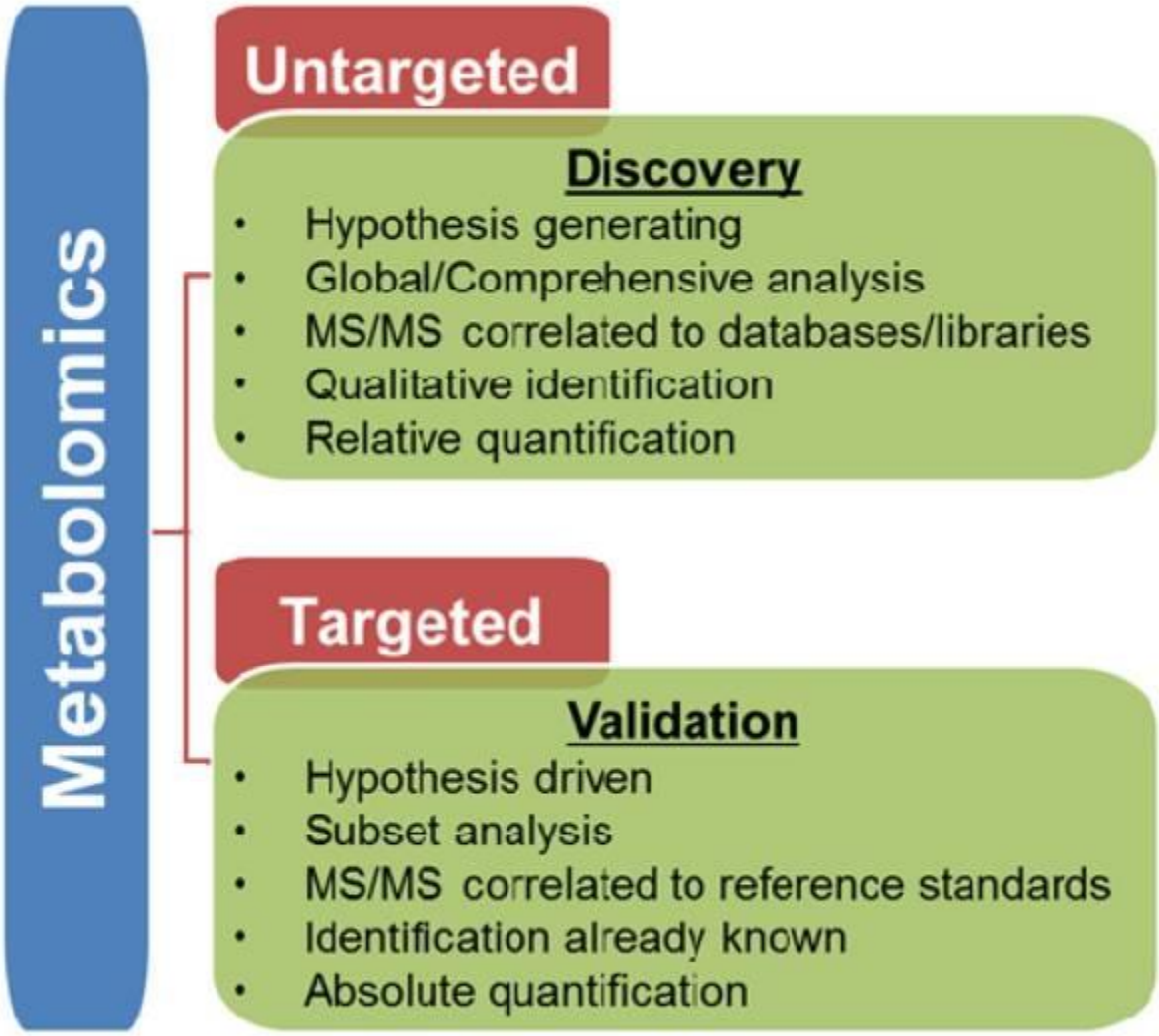

Figure 2. Goals of metabolomics (from Schirimpe-Rutledge, 2016) [107]

Analytical metabolomics using mass spectrometry as a key technology can be broadly divided into two approaches, which have their own advantages and disadvantages. The first is “targeted metabolomics”, which focus on the analysis of a set of known metabolites, is

highly favored for quantitative studies. The other approach aims toward the "unbiased" analysis of metabolome of interest under a given set of conditions and is known as "non-targeted metabolomics" (or untargeted metabolomics) [108]. Using HRMS in combination with chromatography, non-targeted metabolomics´ purpose is to analyze all measurable analytes in a sample, including unknowns [109]. Therefore, all parameters relevant to the final result (i.e. sample type, number of samples, sample preparation, chromatographic separation, and analytical system) need to be considered carefully before designing the analytical method so that the number of detected metabolite features can be maximized, and the obtained results are unbiased, and reproducible. This task is challenging and does not have a global solution. Each step leading to the final result is considered in the following section (Fig. 3).

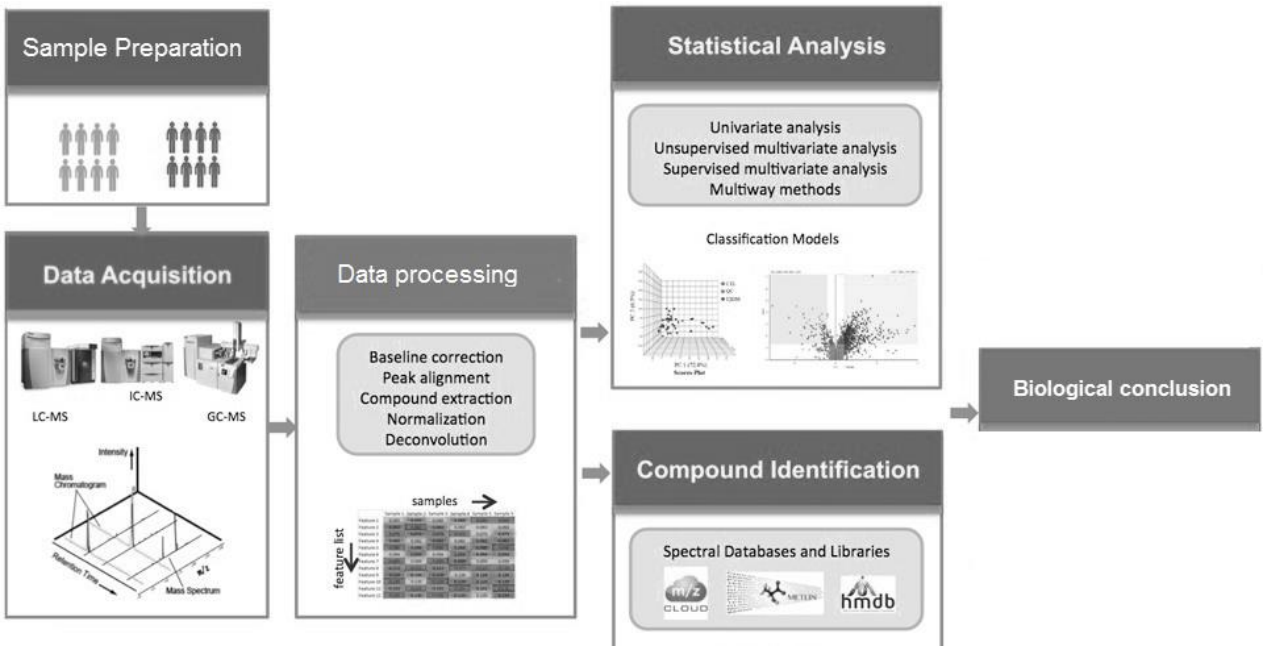

Figure 3. Non-targeted metabolomics workflow (graphic adapted from ThermoFisher Scientific untargeted metabolomics workflows)

1.2.1 Sample collection and sample preparation

Collection of samples, which involves biofluids (e.g. blood, urine, plasma), and cell or tissue extracts is the most critical and first step dictating the quality of following

metabolomics studies. Depending on the analytical goals and the sample types considered, different standard protocols have been proposed and must be applied strictly. Following sample collection is preparation of samples, the choice of which is dependent on sample type, the analytical method to be used, and overall cost. As a rule of thumb, the sample preparation strategy should be minimal but effective, and compatible with the selected analytical method without losing valuable information for metabolite classes of interest.

In the case of cellular metabolomics studies, sample collection requires that metabolic activity must be stopped, and cells must be separated from media, which involving several steps such as quenching, filtration, extraction, and concentration. Quenching, which stops rapid cell metabolism and extracellular enzymatic activity, can be achieved by rapid change in pH, solvent, and/or temperature. The quenching mixture is separated from the cells by centrifugation at low temperature to avoid further turnover of metabolites (or degradation). A common quenching method in microbiological studies, which was validated for reliable separate intra- and extracellular metabolites is using cold aqueous methanol solution as proposed by Koning *et al.* [110] in 1992. Later, other authors reported the leakage of some metabolites from yeast cells during the quenching using Koning's method and proposed a new quenching condition (with pure methanol) to prevent the leakage [111,112]. The next step is separation of the extracellular metabolites from biomass, then subsequent extraction of the intracellular metabolite with minimal losses. There have been considerable efforts to develop reliable and reproducible methods for extraction of metabolites. Various strategies were applied i.e. combining methanol-water with chloroform mixture, boiling ethanol, or cold methanol extraction, in which cold methanol extraction has been quite popular recently [40,111–114]. Finally, the extracted solutions are concentrated due to dilution of previous steps. Freeze-drying and lyophilization are common methods to concentrate samples without thermal degradation, and remove water from aqueous samples. Non-aqueous extracts can also be concentrated using evaporation, or selective enrichment steps such as solid phase extraction can also be applied for metabolomics, which are discussed in section 1.3.

1.2.2 Data acquisition

The combination of LC with MS can routinely detect of thousands of peaks representing metabolites from biological samples, making this approach the most popular for non-targeted metabolomics [107,115]. Common data acquisition approaches for non-targeted metabolomics analysis (NTA) include full-scan (full-spectrum), data-dependent acquisition (DDA), and data-independent acquisition (DIA). Full-scan is the general mode of data acquisition of HRMS, in which mass spectra are continually acquired in a certain range of *m/z* values (from low to high) in a certain period of time (≤1 s). The content of the spectrum varies depends on the ionization method, instrument resolution, and method parameters. DDA is mode of data collection in tandem or two-stage mass spectrometry, (e.g. MS and MS/MS). In this way, MS fragment data of unknown compounds of a sample can be acquired alongside with precursor data within one chromatographic run. In DDA, a certain number or precursor ions, whose *m/z* values were recorded in preliminary scans are chosen automatically by the acquisition software using a certain set of rules, isolated, and subjected to a second stage of mass selection (e.g. QTOF) analysis [116]. The switching between MS and MS/MS stage is triggered by a certain set of predefined rules based on (commonly) the intensity of precursor ion observed, and additional criteria e.g. isotope pattern, charge states, or preferred *m/z* values on a list. DIA is an alternative to DDA, in which all precursor ions within certain *m/z* ranges (windows) will be sequentially fragmented and analyzed in second stage [117]. For example, Tsugawa *et al.* [118] suggested DIA parameters for hydrophilic metabolites analysis using an AB Sciex Triple TOF 5600+ system as: MS1 accumulation time, 50 ms; MS2 accumulation time, 30 ms; collision energy, 45 V; collision energy spread, 15 V; cycle time, 640 ms; Q1 window, 25 Da; mass range, *m/z* 50-500. The acquisition is repeated in stepped selection windows, and generally results more complex spectrum with high number of product ions created from multiple precursor ions contained in a same chosen window. While fragment level information from DDA or DIA is extremely important for identity confirmation workflows, full-scan HRMS data remains the cornerstone of NTA and subsequent statistical analysis for metabolomics studies.

1.2.3 Data processing

After data acquisition, the first data processing step is peak picking and compound extraction (or feature extraction) [119]. Following these steps, each peak now reflects a group of ions with unique *m/z* properties and retention time (RT) value, and is considered as a metabolite feature (or "molecular feature" or "compound"). Dedicated software for interpretation is required to automatically perform the above steps, and to generate a simplified data matrix suitable for statistical analysis. The software used for these tasks may use their own algorithms or integrate several popular algorithms. For example, to align peaks in XCMS, ion features with same *m/z* values and RTs are placed in clusters, it then calculates the deviation from median RT of their respective cluster, warps the remaining of ion features. Another option, LCMSWARP sets a certain confidence for MS/MS identification as a criteria to warp samples, etc. [120]. Finally, the metabolite feature list, which includes information about the *m/z* value and intensity of all aligned isotopologue and adducts as well as the compounds RTis created and assessed based on several parameters such as: accuracy of mass measurement to group ions related to the charge-state envelope, isotopic distribution, and/or the presence of adducts and dimers, as well as potential neutral loss of molecules, peak alignment, and feature frequency [121]. For subsequent statistical analysis, additional processing steps such as scaling, normalization may also be used to help identify significantly changing of features between sample groups [64,115].

1.2.4 Statistical analysis and confirmation of metabolite identification

The metabolite feature list generated from initial data processing is further interpreted using statistical analysis or utilized for subsequent identity confirmation workflows. Statistical processing such as multivariate analysis, differential analysis, and cluster analysis is now commonplace for NTA [65]. These chemometric techniques are integrated in many metabolomics software such as MetAlign (omicX), MZMine (http://mzmine.github.io/), XCMS (The Scripps Research Institute), MetaboAnalyst 4.0

(https://www.metaboanalyst.ca), R package mixOmics (http://mixomics.org/), and Mass Profiler Professional (Agilent Technologies). These techniques allow interpretation of the datasets, reduction of dataset size without losing valuable MS information, and finally revealing valuable information about metabolites in specific biological samples [38]. Both unsupervised and supervised chemometric techniques can be integrated in the software [64]. Unsupervised techniques, which work without pre-knowledge about class information, include cluster analysis (CA) and principal component analysis (PCA) [122]. CA considers the similarity between sets of samples on the basis of their metabolite profiles, then classifies the samples into groups based on degree of association strength between groups involving hierarchically and non-hierarchically ordered clusters. To decide the similarity, Euclidean distance measure is often used, in which the distance is given by $d_{12} = [(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2]^{1/2}$. Others distance measures methods are Minkowski, city-block, and Mahalanobis distances [123]. In contrast, PCA sample grouping is based on the transformation of the original data to new coordinate system, which reflects the relationships among the data. In PCA, each sample is considered as a point in an *n*-dimensional space (*n* is number of monitoring parameters), then the coordinator of space is normally translated to the center of data set (mean centering). A line through coordinator (PC 1) is fitted among the points in space until the sum of square of distance between the projection of each point on PC 1 and coordinator is maximum. Then the second line (PC 2) which is perpendicular to PC 1, and PC 3 which is perpendicular to PC 1 and PC 2, and so on PC n^{th} are constructed. The set of coordinators of the points in new coordinate system is now score matrix (**T**), in which $\mathbf{X} = \mathbf{T}*\mathbf{L}^T$ (where **L** is loading matrix, which provides projection of the points onto the PCs, *T* is transpose of a matrix). The original matrix (**X**) is transformed into a new coordinate system constructed by PCs, which allows a maximization the variation between groups of sample while total variance of data is unchanged [108].

In contrast, supervised methods such as discriminant analysis (DA), partial least square (PLS), principle components regression (PCR), or artificial neural networks (ANNs), class information is needed to construct model, and estimate necessary parameters [124]. For example, in PLS, the data measure is put in X matrix, while Y is response matrix where the

class assignment is described by in term of zero and one. PLS tries to find the relations between X and Y matrices, by maximizing the covariance of their latent variables. Furthermore, orthogonal-PLS (OPLS), a modification of PLS even provides better interpretation of the relevant variables, as it decomposes the predictive information related to the response Y, and "orthogonal" structured information not related to the response, and residual variation [125]. These multivariate techniques may help to determine peaks with significant differences across sample groups, i.e., potential biomarkers. However, none of the techniques is currently considered as the optimum for all applications [19]. Thus instead of relying on only one technique, the combination of several techniques is more favorable, and specific for individual metabolomics challenges. For example, study of Arapitsas *et al.* [126] focused on the effect of storage conditions on the metabolite content of wine, using PCA for data quality control, and orthogonal partial least-squares discriminant analysis (OPLS-DA) was used to find features markers. Rombouts *et al.* [22] established a validated Orbitrap method in combination with OPLS-DA models for non-targeted metabolomics and lipidomics analysis of colon tissue and cell lines. The author successfully reported 15,000 components with CV < 30% (in term of peak area), and discriminated the non-transformed and transformed state in human colon tissues and cell lines. Bade *et al.* [127] reported a method for suspect screening of larger number contaminants in environmental waters using artificial neural networks for RT prediction and HR-MS data analysis. Using training (n = 344) and verification (n = 100) datasets, the optimized 4-layer back-propagation multi-layer perceptron model predicted successfully 85% of all compounds to within 2 min of their measured RT, revealed 95% prediction accuracy of 100 randomly selected compounds and within the 2-minute elution interval. Recently, Samaraweera *et al.* [124] proposed an ANNs retention index (RI) model for chemical structure identification in NTA. In which the model was able to filter off 58% fail candidates from PubChem candidate sets, resulted approximately 2-fold improvement in average rankings in comparisons with other software.

In addition to statistical analysis, identity confirmation of metabolites is finally essential to draw biological conclusions from NTA data. To address the question, Schirimpe-Rutledge and Schymanski [107,128] proposed a 5-level workflow for metabolite

identification confidence using multidimensional MS including exact mass measurement (level 5), generating molecular formula (level 4), matching MS data with databases (level 3), matching MS/MS data with databases (level 2), and validated identification using reference standards (Level 1) (Fig. 4). At level 5, exact mass measurement for a unique feature is obtained. At level 4, information of isotope abundance distribution, charge state, adduct ion, etc. are used to generate molecular formula. To generate molecular formulas, Kind and Fiehn [129] proposed seven heuristic rules including: number element restriction, LEWIS and SENIOR chemical rules, isotopic patterns, hydrogen/carbon ratios, element ratio of N, O, P, S versus C, element ratio probabilities, and trimethylsilylated compounds. At level 3, experimental mass measurement of feature (MS1) was compared to a database of known metabolites (e.g. HMDB, METLIN) within a mass tolerance window (e.g. 10 ppm) with support of diagnostic fragment ions. At level 2, experimental MS/MS spectra are matched against reference fragmentation spectra. Across level 2 up to level 5, matching scores can be generated to represent how good the matching between experimental data from unknown feature and experimental data from the standard is. At the last level, level 1, the validated identification, reference standard is used to confirm structure, which is a gold standard for analytical chemistry. However, it is important to note that to differentiate structural and stereo-isomers, MS/MS data is still not enough, additional separation (i.e. GC, LC, or ion mobility), and sample information is required. The lack of experimental spectral data sets for metabolomics is also a challenge for this task, as less than 2% of human metabolites are possible to be identified via non-targeted MS-based metabolomics techniques, and 90% of known human metabolites don't have authentic standards, or experimentally collected NMR, GC-MS or MS/MS spectra [130]. Fortunately, many commercial software packages for *in silico* metabolite and spectral prediction have been developed to overcome this problem, in which the software could predict physiologically practical metabolites and to accurately predict NMR, GC-MS or MS/MS spectra of both known and predicted metabolite structure. For example, the developers of the human metabolome database (HMDB) claim that using *in silico* approaches in comparison with traditional manual approach enables an increase of nearly three times the number of metabolite confirmations (known, expected, and predicted) in HMDB 4.0 [130]. Finally, at all levels described here, it is important to remember that

chromatographic retention time (or electrophoretic migration time) is also an essential part of the identity confirmation workflow. To this end, the retention properties as well as the repeatability of the separation method are of key importance in development of rugged workflows suitable for metabolomics applications.

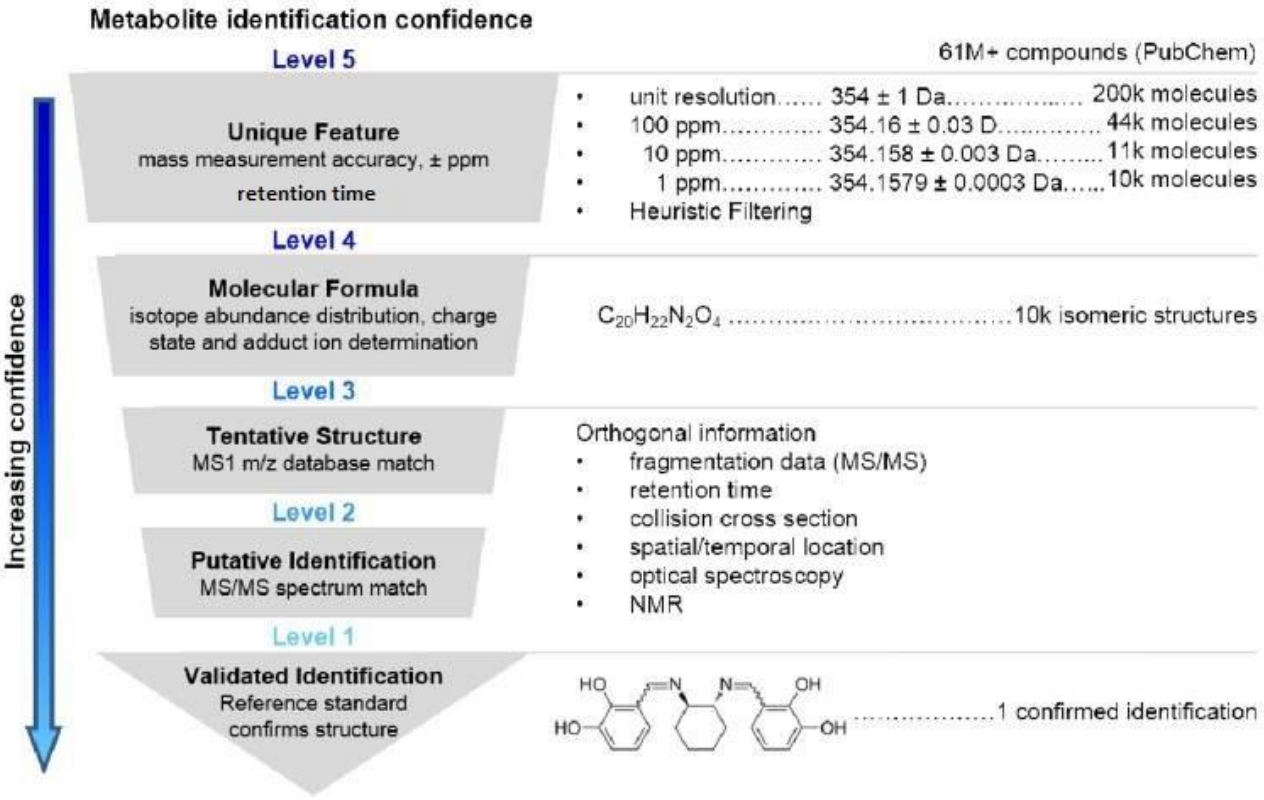

Figure 4. Confidence levels for metabolites identification (adapted from Schirimpe-Rutledge, 2016) [107]

1.2.5 Challenges in non-targeted metabolomics

In non-targeted metabolomics, the experiment design involving sample preparation steps, choice of separation method, data processing approach, and other analytical method parameters has great impact to the quality of biological conclusion that could be drawn from the study. Non-targeted metabolomics also faces other challenges that are well-described in contemporary studies. Due to the highly diversity of physicochemical properties, it is

impossible to assess every metabolite using a single sample preparation process. The large numbers of components derived from the growth medium/quenching buffer, and/or concentration variance between samples in different growth medium, in different harvesting stages make the challenges even greater. The fact is that none of current single analytical technique in NTA is suitable for the detection, identification, and quantification of all metabolites. Therefore, combinations of techniques, or advanced multidimensional (2D-LC, 2D-GC) approaches must be applied [18,98,108,131]. Another critical point in metabolomics is that all assessments will only be valid if the samples are studied under the same conditions with a validated method, and the change of recorded signal should be related to the metabolite concentration [108,132,133]. Nevertheless, for chromatographic separation, RT and *m/z* value variation of each metabolite feature are prone to happen due to random variation, signal intensity variation, accurate mass measurement drift, sample collection/preparation bias, sample instability, temperature fluctuation, column age, ion suppression, ionization efficiency, etc. These variations cannot be reliably predicted or easily modeled. Hence, uncertainty increases, and accumulates from the beginning to later stages of the analytical process [120,134]. Each step of assessment must be therefore well-understood, and designed carefully to avoid significantly affecting the data interpretation.

Other challenges for NTA are the high number of isomeric metabolites, lack of standard reference materials, difficulties in achieving highly confident metabolite identity confirmation, unpredictable fragmentation pattern for ESI, and finally the commonness of many metabolites across species create difficulties for discerning the biological source [107]. Moreover, poor data quality may fail to reveal biological differences between sample sets, or create false hypothesis. Thus, NTA needs to apply several analytical validation methods (or guidelines) to be considered as fit for purpose and avoid false conclusions due to inadequate data quality. To this end, using “biological QC” samples made from pooling aliquots of all samples, or QC sample consisting of the blank matrix spiked with known concentrations of various analytes, is an important strategy for ensuring data quality over long measurement sequences [135,136]. QC samples are analyzed during the measurement series in order to confirm repeatability of RT, peak area, and mass accuracy throughout the

measurement sequence. The FDA recommends for NTA by LC-MS values close to 30% RSD could be allowed [108]. While, Zelena *et al.* [63] noted an RSD of <20 % for the peak area of QC samples during analysis time was acceptable. Another challenge is software for interpretation. Myers *et al.* [119] provide a detailed comparison about chromatogram construction and peak detection between two open source software packages (XCMS and MZmine 2) on four sets of data. The comparison showed that two software packages employing the same continuous wavelet transform (CWT)-based centWave algorithm (a combination of density based feature detection approach and Wavelet based approach [137]), generated feature lists with significant differences including both false positive EIC peaks and failure to detect real EIC peaks. To the end, despite continuous effort across a wide range of metabolomics studies, a comprehensive, and detailed guideline for demonstrating a high analytical quality results for every metabolomics application is still lacking [138]. Since 2005, international metabolomics communities (i.e. Metabolomics Society, Metabolomics Standards Initiative - MSI) have gathered for the purpose of standardization of metabolomics studies. A lot of efforts have been made to establish standards in metabolomics including data evaluation, method validation and reporting, most importantly to support an extension of existing workflows or assure reproducibility of published work[139].

1.3 Selective enrichment in Metabolomics

1.3.1 Primary phosphorylated metabolites involved in central carbon metabolism

As discussed in the introduction, targeted metabolomics studies focused on subsets of metabolites are essential for understanding key pathways in cellular metabolism. One prominent example of this is the central carbon metabolism, which includes a complex series of enzymatic reactions to convert sugars into metabolic precursors for entire biomass generation of the cell [140]. The metabolism comprises several precursor metabolites that form the basis for biomass (e.g. D-glucose 6-phosphate (G6P) for glycogen, D-fructose 6-phosphate (F6P) for cell wall, D-glyceraldehyde 3-phosphate (GAP) for lipids, etc.), and one precursor (glycerate 1,3-bisphosphate) essential for the positive balance in glycolysis (Fig. 5). Glycolysis, pentose phosphate pathway, and tricarboxylic acid cycle (TCA) are the three pathways of the central carbon metabolism.

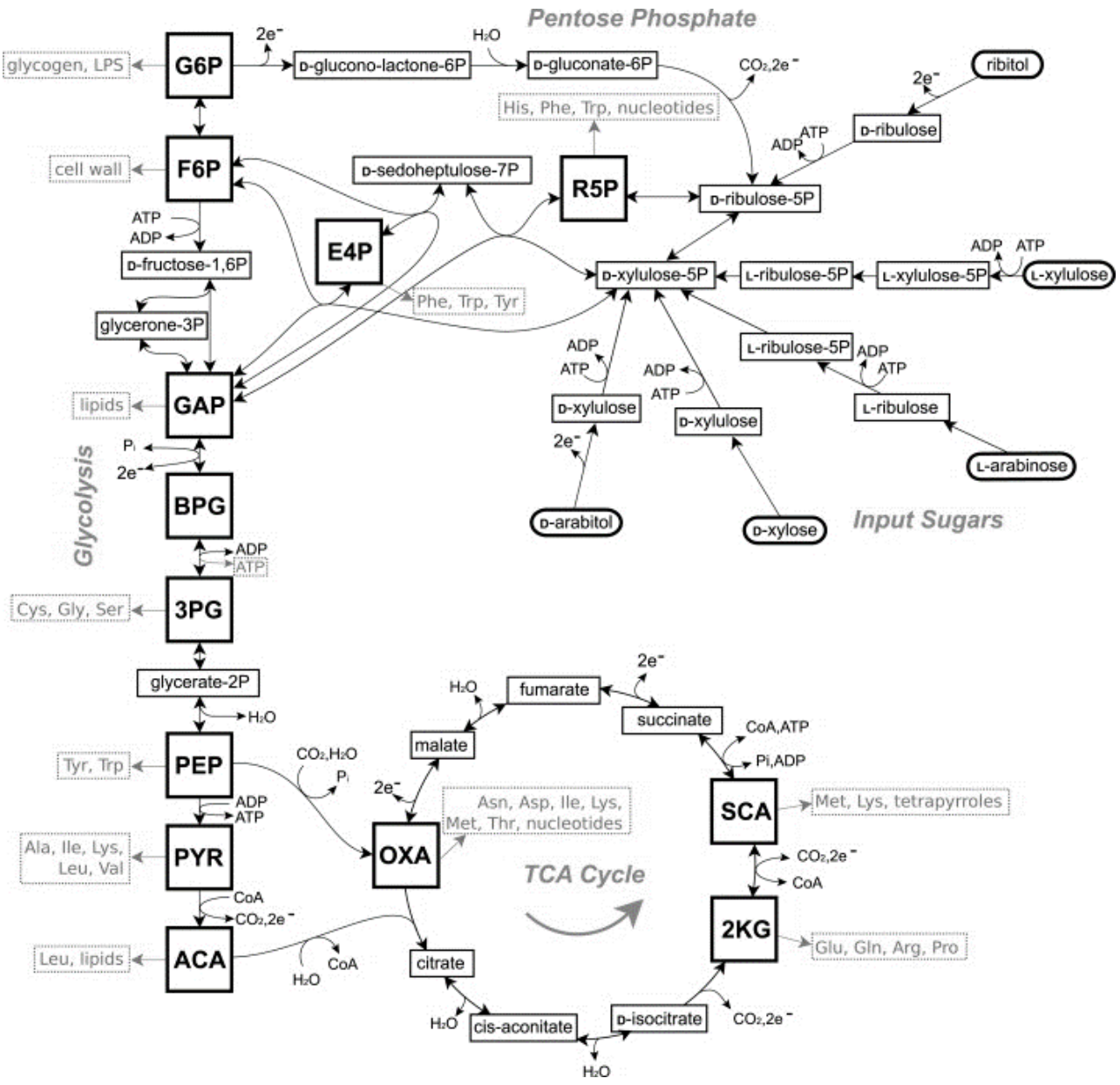

Figure 5. The central carbohydrate metabolic network (from Noor, 2010) [140]

Within the central carbon metabolism, phosphorylated metabolites play very important roles in substrate degradation, energy budget, cofactor regeneration, biosynthetic precursor supply, even regulation of tumor energetic metabolism, and are therefore essential in all life processes [94,141]. In glycolysis, glucose is converted into pyruvate, and generates energy in form of adenosine triphosphate (ATP) and reducing power in form of nicotinamide adenine dinucleotide (NADH) [142]. Its products includes many important precursor metabolites such as six-carbon compounds of G6P, F6P, and three-carbon compounds of

glycerol phosphates, glyceraldehyde 3-phosphate (GAP), glycerate 3-phosphate (Gro3P), phosphoenolpyruvate (PEP), and pyruvate. At the beginning, glucose is phosphorylated to G6P. G6P is polymerized and serves as a storage location for chemical energy in form of glycogen. If the body needs energy, it could be isomerized to form F6P and then phosphorylated to fructose 1,6-bisphosphate (FBP). FBP is quite common in cells as a key component in the glycolysis metabolic pathway. The enzyme aldolase splits FBP into dihydroxyacetone phosphate (DHAP) and glyceraldehyde phosphate. Then, FBP acts as an allosteric activator of pyruvate kinase enzyme, which later catalyzes the transfer of PEP (another important chemical compound involved in glycolysis and gluconeogenesis pathway) to adenosine diphosphate (ADP), producing pyruvate and ATP [143]. G6P also plays a role in the pentose phosphate pathway (PPP). PPP includes two different phases, irreversible oxidative phase and reversible non-oxidative phase. Oxidative phase is a main source of reducing power and metabolic intermediates for biosynthetic processes [144]. In this phase, G6P is converted into ribulose 5-phosphate (Rl5P), and producing NADPH (reducing equivalents). Later, Rl5P followed by reversible non-oxidative phase converted back to G6P using transketolase and transaldolase enzymes [145]. In this phase, pentose phosphates (i.e. ribose 5-phosphate (R5P), Rl5P, and xylulose-5-phosphate (Xl5P) are employed as building blocks of nucleotides and deoxynucleotides in deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) [142].

Clearly, phosphorylated compounds play an important role in biological systems, involving many important metabolic pathways. By understanding their variations in a biological system, many valuable biological conclusions can be made. However, challenges with the measurements of these metabolites are not only very low concentration level (trace, ultra-trace), but also many of them are isomeric compounds, highly polar, rapid in metabolic inter-conversion and are not easily chromatographically separated [146]. For example, R5P, Rl5P and Xl5P elute very close to each other due to their stereoisometric structures, while pentose phosphates and hexose phosphates can exist in both opened chain and closed ring forms. Both MS and NMR [147,148] methods can be used to analyze the phosphorylated compounds. As already mentioned, due to the very low sensitivity of ^{31}P NMR signal

(meaning that more sample and longer measurement time will be required), MS methods are currently essential for such metabolomics studies. Different separation techniques in combination with MS have been applied to overcome the challenges in analysis of phosphorylated metabolites within different type of biological samples, with LC-MS and GC-MS being the most common examples.

1.3.2 Quantitative analysis of primary phosphorylated metabolites of the central carbon metabolism

As stateed earlier, metabolomics aims at the (comprehensive) qualitative and quantitative analysis of the metabolome. Quantitative result from metabolomics measurement can be used as diagnostic- or bio- markers for an extensive range of biological studies. In the context of metabolomics, the term "metabolite profiles" was described first by Horning *et al.* [149] as multicomponent GC analyses that describe metabolic patterns for a certain group or class of metabolites. Since then, many methods involving both elemental and molecular mass spectrometry have been developed and validated for quantitative analysis of primary phosphorylated metabolites such as sugar phosphates from the central carbon metabolism.

Elemental mass spectrometry is an important analytical technique allowing accurate determination of isotope concentration [150]. Of most relevance to metabolomics, inductively coupled plasma mass spectrometry (ICP-MS) employs a high temperature (8000 K) ionization source, which is the plasma of ionized argon gas generated in an intense electromagnetic field. This ionization source allows efficient vaporization, drying, dissociation, atomization and ionization of all atoms, ions and molecules present in the sample into elemental constituents [151]. In principle, ICP-MS can detect most elements except the ones that have high ionization potentials or high background noise (e.g. H, O, N, F) or suffer from polyatomic interferences, which are usually formed in argon plasma (e.g. P, S, Se, Ca, Fe). These interferences can be overcome by use of high-resolution instruments

such as sector-field (SF) ICP-MS or collision/reaction cell technologies [152]. Thus, for phosphorus measurement, ICP-MS can be used in combination with chromatographic separation techniques for the analysis of phosphorylated compounds related to central carbon metabolism [153,154]. Chu *et al.* [146] applied anion exchange chromatography in combination with ICP-MS for separation and accurate quantification of several sugar phosphates in cell extract including G6P, M6P, F6P, R5P, 6PGA, 3PG, 2PG, and FBP. In this study, the sugar phosphates were separated on Dionex CarboPAC PA1 column, then on-line transferred to ICP-MS for detection of *m/z* value of 47 (PO^+). The reported method has a very good short-term and long-term repeatability with relative standard deviation (in term of peak area) below 3.0% and 10.0% respectively. The authors also optimized the method, resulting the absolute on-column LODs and LOQs of interested compounds in sub-pmol L^{-1} range. However, phosphorus measurement using ICP-MS has also has some limitations. Due to the hard ionization source, the advantage of ICP-MS regarding phosphorylated compounds analysis is a disadvantage for speciation as all structural information of compounds is lost. Furthermore, phosphorus measurement by ICP-MS suffers from polyatomic interference caused by atomic or molecular ions, which have the same mass-to-charge as phosphorus (e.g. $^{15}N^{16}O^+$, $^{14}N^{17}O^+$, $^{14}N^{16}O^1H^+$). Current ICP-MS instrumental software can correct for all known atomic isobaric interferences, however it does not correct for most polyatomic interferences [155]. Therefore, the reaction cell technique using oxygen gas, or collision cell quadrupole mass spectrometry are required for selective phosphorus measurement. Moreover, phosphorus only has only one stable isotope ^{31}P, which make absolute quantification via an isotope dilution strategy impossible.

In contrast to elemental mass spectrometry, molecular mass spectrometry employing soft ionization techniques e.g. ESI (for LC), CI (for GC), provides more information about compound structure. Molecular mass spectrometry in combination with chromatographic separation techniques (e.g. LC, and GC) has extensively expanded the capacity of chemical analysis for highly complex biological samples. Compared to GC-MS, LC-MS can be applied for a wider range of molecules without complex sample preparation [18]. For example, Vizán *et al.* [156] proposed a LC-MS/MS method for quantification of

phosphorylated carbohydrates in HT29 human colon adenocarcinoma cell line to understand the role of sugar phosphates in regulating tumor energetic metabolism. Huck *et al.* [145] profiled PPP intermediates in blood spots using liquid chromatography combined with MS and successfully established reference values for intermediates of pentose phosphate pathway in human blood spots. In 2007, Antonio *et al.* [157] developed an on-line LC-ESI-MS/MS method using a porous graphitic carbon (PGC) stationary phase for sensitive targeted analysis of the main glycolytic intermediates, sugars and sugar phosphates in plants. Luo *et al.* [158] reported a highly selective and sensitive LC-MS/MS method for identification and quantification of intracellular metabolites in the central carbon metabolism using the volatile ion pair modifier tributylammonium acetate (TBAA). However, the disadvantage for LC, under typical reversed-phase conditions is that many phosphorylated central carbon metabolites have poor separation and sensitivity due to their retention near to column void volume. Addition of ion-pairing agents i.e. tributylamine (TBA) to the mobile phase may help to extend coverage of these metabolites [53]. Nevertheless, the severe drawback of ion-pairing reagents is sustainable mass spectrometer contamination, which will limit use of the instrument for other methods [65,159]. Covalent coating of ion-pairing agents on a RP-stationary phase could be another choice [65], but has not yet proven rugged enough for commercialization. Another way to extend the coverage is using different stationary phases i.e. porous graphitic carbon (PGC), or HILIC [105,160]. PGC stationary phases offer strong retention of phosphorylated metabolites, but these columns are easily chemical modified over time, causing significantly reduction in performance and RT stability [65]. HILIC phases also offer good retention of relevant metabolites, but suffers from decreased chromatographic performance, lack of selectivity to discriminate key metabolite isomers, poor reproducibility, and longer equilibrium time compared to RP phases [70,159]. A final option for LC is to increase the hydrophobicity, which results improvement in retention and ESI ionization efficiency of targeted metabolites through derivatization [161]. However, derivatization steps increase the complexity of both analytical workflow, and sample matrix. Furthermore, no established protocols for selective derivatization of phosphate group-bearing metabolites is currently available.

Despite the advantages of LC-MS for broadest coverage of phosphorylated metabolites based on measurement of underivatized samples, GC-MS has proved to be the gold standard for separation and quantification of key phosphorylated metabolites. For example, Koek *et al.* [3] set up a GC-MS method consisting of an oximation and silylation derivatization reaction for the analysis of microbial metabolomes. The method was extensively validated using many other microorganisms, and showed the ability to analyze many metabolite classes including many sugars and sugar phosphates. Vielhauer *et al.* [162] established a simplified GC-isotope dilution MS method for absolute quantification of intra- and extra- celluar metabolites including not only the phosphorylated sugars but also other metabolite classes. Mairinger *et al.* [93] successfully reported the use of GC-QTOFMS with chemical ionization for ^{13}C-based flux for analysis of 42 free intracellular metabolites including sugar phosphates. Chu *et al.* [92] reported a method for isotopologue analysis of sugar phosphates in yeast cell extracts using GC-TOFMS, which employed automated ethoxymation (in which the keto groups (-C=O) will become –C=N-O-R, therefore keeps sugars stay in open chain structure yielding to less derivatives and hence less chromatographic peaks) and followed by trimethylsilylation (in which active hydrogen in hydroxyl- , carboxylic- , amine- and thiol groups are replaced by trimethylsilyl groups). The proposed method successfully separated many of sugar phosphates in cell extracts and allowed accurate determination of isotopologue with maximum bias, and average bias of 4.6% and 1.4% respectively.

Despite the already significant advances made, the analysis of intracellular central carbon metabolites by MS-based methods still faces several practical challenges such as: the low concentration of phosphorylated metabolites, complex matrix components, separation of highly polar sugar related metabolites and isomeric compound. These challenges again emphasize the important of sample preparation steps prior to the analysis in metabolomics.

1.3.3 Cleanup and selective enrichment methods in metabolomics

As mentioned earlier, challenges for primary phosphorylated metabolites involving central carbon metabolism, includes wide range of metabolite concentration (from sub µmol L^{-1} to mmol L^{-1}) and a highly complex sample matrix. Indeed, coupling MS with chromatographic separation techniques could help to partly remove sample matrix, but may not be enough in many cases. Therefore, prior to the separation, further sample cleanup, enrichment, and derivatization steps for specific classes of metabolites, especially those present in low concentrations or exhibiting poor stability, may be involved [13]. The most popular method to enrich metabolites in samples prior to MS-based measurements is removal of extraction solvent from the initial volume and subsequent re-constitution in a lower volume of (more) suitable solvent. However, metabolite losses may occur due to oxidation or decomposition during these steps, while both desired and interfering components are enriched [163,164]. For this reason, development of sample cleanup/enrichment procedures may be required for sensitive and selective MS-based analysis of targeted metabolite classes.

A review from Causon *et al.* in 2016 [13] about sample preparation strategies for MS-based metabolomics studies in industrial biotechnology, classified methods of enrichment in to 3 main categories including liquid-liquid extraction (LLE), solid-phase extraction (SPE), and dispersive SPE (dSPE). In LLE (or solvent extraction), compounds are separated based on their relative solubility in two liquids that are not dissolved in each other. The choice of extraction is critical to reduce unspecific extraction from sample matrix i.e. halogenated solvents are commonly combined with hydrophilic solvents for extraction of polar compounds, while non-ionized analytes are usually extracted by organic solvents than charged species [164]. In addition, pH should be controlled in the range of pK_a ±2 as a general rule. The extraction is primary limited to lipid analysis where polarity trends can be best exploited. SPE has become one of the most essential sample preparation techniques, widely used for metabolite extraction from biological fluids in metabolomics [165,166]. SPE and dSPE involve the affinity of solutes dissolved in a mobile phase to solid phase (stationary phase). The process may include several steps i.e. conditioning, loading, washing, and eluting. First, sample is loaded onto SPE material, then retained analytes are consequently

eluted using a solvent or solvent mixtures with sufficient elution strength. The interaction between sorbent and analyte could be explained by several mechanisms, i.e. Van der Waals interactions, dipole–dipole interactions, hydrogen bonding, electrostatic forces, or combination of several interactions. The difference between dSPE and SPE is that sorbent material is added directly to organic phase in dSPE, while SPE involves flowing of sample over material packed into a cartridge. Generally, dSPE aims to promote the separation of compounds of interest from matrix and uses a centrifugation step to remove the solid phase. In addition to offline SPE, online sample extraction using a pre-column coupled to LC-MS also has been successfully used for metabolomics investigations [95]. Both SPE and dSPE have some potential disadvantages such as the weakness for direct enrichment of limited retention of polar metabolites (sugars, sugar related compounds, etc.), non-specific interactions and instability under extreme pH of C18 modified silica phase.

Applying cleanup and selective enrichment strategy prior to chromatographic separation will make the overall workflow more complex, increases the chance to introduce error. However, using a suitable and effective cleanup and enrichment method indeed helps to remove high concentration matrix components and promotes the sensitivity and selectivity for low concentration phosphorylated metabolites of interest. Among strategies mentioned earlier, sample cleanup and enrichment for phosphorylated metabolites employing SPE is one of the most popular. Not only because of a wide range of SPE materials for cleanup and enrichment purpose has been developed and validated for a longtime, but also the methods shows the potential to overcome challenges in phosphorylated metabolites analysis. Especially, SPE based on metal oxide affinity chromatography was reported since many years ago as highly potential enrichment method for phosphor related compounds.

1.3.4 Metal oxide affinity chromatography for selective enrichment of primary phosphorylated metabolites

Of relevance to the enrichment of phosphorylated compounds, the review of Fíla *et al.* [167] summarizes several techniques applied in phosphoprotein and phosphopeptide enrichment, including immunoprecipitation (using phosphoamino acid-selective antibodies), immobilized metal affinity chromatography (IMAC, employing a matrix of resins with associated positive metal ions can catch negatively charged phosphate group), phos-tag (IMAC variation, in which the matrix composed of Phos-Tag, chemically named 1,3-bis [bis(pyridin-2-ylmethyl)amino]propan-2-olato dizinc(II) complex, that is anchored to agarose), metal oxide affinity chromatography (MOAC, using matrix composed of metal oxide or hydroxide), polymer-based metal ion affinity capture (PolyMAC, using a soluble dendrimer attaches two types of side group), phosphopeptide chemical modification, and phosphopeptide precipitation (using calcium phosphate precipitation). The review highlighted that the main problem for these enrichment strategies is the non-specific binding and the challenge to improve these protocols still remains. The authors also concluded that the most promising methods for phosphoproteomics studies are IMAC and MOAC, where TiO_2/ZrO_2-MOAC is considered more promising due to its ruggedness, selectivity, and sensitivity.

MOAC using materials such as titanium and zirconium oxides has been reported as a useful strategy for solid phase extraction (SPE)-based enrichment for phosphorylated peptide and phosphoprotein due to the particular affinity of metal oxides surfaces toward phosphate-group containing compounds [167–172]. The binding mode of the phosphate anion, mono-substituted phosphates and carboxylic acids to the TiO_2 surface was explained as bidentate nature, which bases on the donation of two electron pairs to a metal atom, or particular retention mechanism due to the fact that metal oxide could act as a Lewis acid [173,174]. In the past decade, TiO_2- and ZrO_2- based materials have also been recognized as an alternative stationary phase for LC due to its high adsorption capacity, selectivity and chemical stability when used under extreme pH and temperature conditions [175–177]. Due to nature of the phosphate-metal oxide interaction, the phosphorylated enrichment is more selective under

low pH conditions (pH 2-3) as phosphate group remains negatively charged, while the protonation of carboxylic acids should prevent their binding to the sorbent [167,178]. Ikeguchi *et al.* [179] reported online enrichment of 6 organic phosphate compounds that included guanosine 3',5'-cyclic monophosphate (cGMP); guanosine monophosphate (GMP), guanosine diphosphate (GDP), adenosine 3',5'-cyclic monophosphate (cAMP); adenosine 5'-monophosphate (AMP), O-phospho-DL-tyrosine (P-Tyr) using a titanium oxide pre-column, in which the organic phosphates were retained under acidic conditions and eluted under alkaline condition. Other groups also reported different approaches using metal oxide materials, allowing the separate analysis of phosphorylated and non-phosphorylated peptides from an enzymatic digest of proteins [168,180].

While there are already many works dedicated to phosphoproteomics, the enrichment of phosphorylated primary metabolites using MOAC has not received so much attention. In 2004, Sekiguchi's group demonstrated the use of TiO_2 columns to trap sugar phosphates (galactose 1-phosphate, glucose 1-phosphate, sucrose 6-phosphate, F6P, glucose 1,6-diphosphate, UDP-glucose) present in plant extracts for analysis using anion exchange chromatography with pulsed amperometric detector and MS/MS [95,181]. The results also highlighted some difficulties regarding the quantitative determination of some sugar phosphates (e.g. degradation of FBP during microwave-assisted extraction, moderate recovery of F6P, several peaks for erythrose 4-phosphate). Since then, new commercial MOAC (TiO_2, ZrO_2) products were available from many manufacturers, but their usage for metabolomics applications has not yet been reported.

1.4 Reference

[1] R.J. DeBerardinis, C.B. Thompson, Cellular metabolism and disease: what do metabolic outliers teach us?, Cell. 148 (2012) 1132–1144. doi:10.1016/j.cell.2012.02.032.Cellular.

[2] O. Fiehn, Metabolomics--the link between genotypes and phenotypes., Plant Mol. Biol. 48 (2002) 155–71. http://www.ncbi.nlm.nih.gov/pubmed/11860207 (accessed February 8, 2017).

[3] M. Koek, B. Muilwijk, M.J. Werf, T. Hankemeier, Microbial metabolomics with gas chromatography mass spectrometry, Anal Chem. 78 (2006) 1272–1281. doi:10.1021/ac051683+.

[4] A. Alonso, S. Marsal, A. Julià, Analytical methods in untargeted metabolomics: state of the art in 2015., Front. Bioeng. Biotechnol. 3 (2015) 23. doi:10.3389/fbioe.2015.00023.

[5] A. Zhang, H. Sun, H. Xu, S. Qiu, X. Wang, Cell Metabolomics, Omi. A J. Integr. Biol. 17 (2013) 495–501. doi:10.1089/omi.2012.0090.

[6] R.K. Naviaux, Metabolic features of the cell danger response, Mitochondrion. 16 (2014) 7–17. doi:10.1016/j.mito.2013.08.006.

[7] S. Padmanabhan, Handbook of Pharmacogenomics and Stratified Medicine, 2014. doi:10.1016/C2010-0-67325-1.

[8] R. Zenobi, Single-cell metabolomics: Analytical and biological perspectives, Science (80-.). 342 (2013) 0–11. doi:10.1126/science.1243259.

[9] D.A. Dickinson, H.J. Forman, Cellular glutathione and thiols metabolism, Biochem. Pharmacol. 64 (2002) 1019–1026. doi:10.1016/S0006-2952(02)01172-3.

[10] G. L. Cantoni, S-adenosylmethionine; a new intermediate formed enzymatically from

l-methionine and adenosinetriphosphate, J. Biol. 204 (1953) 403–416.

[11] J.K. Nicholson, J. Connelly, J.C. Lindon, E. Holmes, Metabonomics: a platform for studying drug toxicity and gene function, Nat. Rev. Drug Discov. 1 (2002) 153–161. doi:10.1038/nrd728.

[12] O. Schneewind, D.M. Missiakas, Protein secretion and surface display in Gram-positive bacteria., Philos. Trans. R. Soc. Lond. B. Biol. Sci. 367 (2012) 1123–39. doi:10.1098/rstb.2011.0210.

[13] T.J. Causon, S. Hann, Review of sample preparation strategies for MS-based metabolomic studies in industrial biotechnology, Anal. Chim. Acta. 938 (2016) 18–32. doi:10.1016/j.aca.2016.07.033.

[14] G.L.A.H. Jones, E. Sang, C. Goddard, R.J. Mortishire-Smith, B.C. Sweatman, J.N. Haselden, K. Davies, A.A. Grace, K. Clarke, J.L. Griffin, A Functional Analysis of Mouse Models of Cardiac Disease through Metabolic Profiling, J. Biol. Chem. 280 (2005) 7530–7539. doi:10.1074/jbc.M410200200.

[15] J. Allen, H.M. Davey, D. Broadhurst, J.K. Heald, J.J. Rowland, S.G. Oliver, D.B. Kell, High-throughput classification of yeast mutants for functional genomics using metabolic footprinting, Nat. Biotechnol. 21 (2003) 692–696. doi:10.1038/nbt823.

[16] L.M. Raamsdonk, B. Teusink, D. Broadhurst, N. Zhang, A. Hayes, M.C. Walsh, J.A. Berden, K.M. Brindle, D.B. Kell, J.J. Rowland, H. V. Westerhoff, K. van Dam, S.G. Oliver, A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations, Nat. Biotechnol. 19 (2001) 45–50. doi:10.1038/83496.

[17] O. Fiehn, J. Kopka, P. Dörmann, T. Altmann, R.N. Trethewey, L. Willmitzer, Metabolite profiling for plant functional genomics, Nat. Biotechnol. 18 (2000) 1157–1161. doi:10.1038/81137.

[18] A.R. Fernie, R.N. Trethewey, A.J. Krotzky, Metabolite profiling: from diagnostics to

systems biology, Nat. Publ. Gr. 5 (2004) 1–7. doi:10.1038/nrm1451.

[19] E.J. Want, A. Nordstrm, H. Morita, G. Siuzdak, A. Nordstro, From Exogenous to Endogenous : The Inevitable Imprint of Mass Spectrometry in Metabolomics, (2007) 459–468. doi:10.1021/pr060505.

[20] C.M. Metallo, M.G. Vander Heiden, Understanding metabolic regulation and its influence on cell physiology, Mol. Cell. 49 (2014) 388–398. doi:10.1016/j.molcel.2013.01.018.Understanding.

[21] L. Fuentes, S. Lebenkoff, K. White, C. Gerdts, K. Hopkins, J.E. Potter, D. Grossman, P.E. Project, R. Sciences, Metabolism and epigenetics: a link cancer cells exploit Alessandro, Curr Opin Biotechnol. 93 (2016) 292–297. doi:10.1016/j.contraception.2015.12.017.Women.

[22] C. Rombouts, M. De Spiegeleer, L. Van Meulebroek, W.H. De Vos, L. Vanhaecke, Validated comprehensive metabolomics and lipidomics analysis of colon tissue and cell lines, Anal. Chim. Acta. (2019). doi:10.1016/J.ACA.2019.03.020.

[23] D.G. Watson, A rough guide to metabolite identification using high resolution liquid chromatography mass spectrometry in metabolomic profiling in metazoans., Comput. Struct. Biotechnol. J. 4 (2013) e201301005. doi:10.5936/csbj.201301005.

[24] N. Araníbar, B.K. Singh, G.W. Stockton, K.-H. Ott, Automated Mode-of-Action Detection by Metabolic Profiling, Biochem. Biophys. Res. Commun. 286 (2001) 150–155. doi:10.1006/bbrc.2001.5350.

[25] N.W. Lutz, N. Yahi, J. Fantini, P.J. Cozzone, A new method for the determination of specific 13C enrichment in phosphorylated [1-13C]glucose metabolites. 13C-coupled, 1H-decoupled 31P-NMR spectroscopy of tissue perchloric acid extracts., Eur. J. Biochem. 238 (1996) 470–475.

[26] N.J. Andreas, M.J. Hyde, M. Gomez-Romero, M.A. Lopez-Gonzalvez, A.

Villaseñor, A. Wijeyesekera, C. Barbas, N. Modi, E. Holmes, I. Garcia-Perez, Multiplatform characterization of dynamic changes in breast milk during lactation, Electrophoresis. 36 (2015) 2269–2285. doi:10.1002/elps.201500011.

[27] A. Andersson, A. Lindgren, B. Hultberg, Effect of thiol oxidation and thiol export from erythrocytes on determination of redox status of homocysteine and other thiols in plasma from healthy subjects and patients with cerebral infarction, Clin. Chem. 41 (1995) 361–366.

[28] K. Yonezawa, S. Nishiumi, S. Nishiumii, J. Kitamoto-Matsuda, T. Fujita, K. Morimoto, D. Yamashita, M. Saito, N. Otsuki, Y. Irino, M. Shinohara, M. Yoshida, K. Nibu, Serum and tissue metabolomics of head and neck cancer., Cancer Genomics Proteomics. 10 (n.d.) 233–8. http://www.ncbi.nlm.nih.gov/pubmed/24136976 (accessed October 18, 2018).

[29] G. Ye, Y. Liu, P. Yin, Z. Zeng, Q. Huang, H. Kong, X. Lu, L. Zhong, Z. Zhang, G. Xu, Study of Induction Chemotherapy Efficacy in Oral Squamous Cell Carcinoma Using Pseudotargeted Metabolomics, J. Proteome Res. 13 (2014) 1994–2004. doi:10.1021/pr4011298.

[30] L. Dang, D.W. White, S. Gross, B.D. Bennett, M.A. Bittinger, E.M. Driggers, V.R. Fantin, H.G. Jang, S. Jin, M.C. Keenan, K.M. Marks, R.M. Prins, P.S. Ward, K.E. Yen, L.M. Liau, J.D. Rabinowitz, L.C. Cantley, C.B. Thompson, M.G. Vander Heiden, S.M. Su, Cancer-associated IDH1 mutations produce 2-hydroxyglutarate, Nature. 462 (2009) 739–744. doi:10.1038/nature08617.

[31] K. Kuroda, M. Ueda, Arming Technology in Yeast-Novel Strategy for Whole-cell Biocatalyst and Protein Engineering., Biomolecules. 3 (2013) 632–50. doi:10.3390/biom3030632.

[32] E. Kondo, P.J. Marriott, R.M. Parker, K.A. Kouremenos, P. Morrison, M. Adams, Metabolic profiling of yeast culture using gas chromatography coupled with

orthogonal acceleration accurate mass time-of-flight mass spectrometry: Application to biomarker discovery, Anal. Chim. Acta. 807 (2014) 135–142. doi:10.1016/J.ACA.2013.11.004.

[33] R.M. Seifar, A. ten Pierick, P.T.N. van Dam, Analysis of Metabolites from the Tricarboxylic Acid Cycle for Yeast and Bacteria Samples Using Gas Chromatography Mass Spectrometry, in: Methods Mol. Biol., 2018: pp. 277–282. doi:10.1007/978-1-4939-7592-1_20.

[34] G.A.N. Gowdaa, D. Rafterya, G.A.N. Gowda, D. Raftery, Can NMR solve some significant challenges in metabolomics?, J Magn Reson. 1 (2016) 144–160. doi:10.3109/02656736.2015.1007538.Histotripsy.

[35] H. Gu, Z. Pan, B. Xi, V. Asiago, B. Musselman, D. Raftery, Principal component directed partial least squares analysis for combining nuclear magnetic resonance and mass spectrometry data in metabolomics: Application to the detection of breast cancer, Anal. Chim. Acta. 686 (2011) 57–63. doi:10.1016/j.aca.2010.11.040.

[36] C.K. McClure, Structural Chemistry Using NMR Spectroscopy, Organic Molecules, 3rd ed., Elsevier Ltd., 2016. doi:10.1016/B978-0-12-803224-4.00294-6.

[37] T.W.-M. Fan, A.N. Lane, Applications of NMR spectroscopy to systems biochemistry., Prog. Nucl. Magn. Reson. Spectrosc. 92–93 (2016) 18–53. doi:10.1016/j.pnmrs.2016.01.005.

[38] H. Gowda, J. Ivanisevic, C.H. Johnson, M.E. Kurczy, H.P. Benton, D. Rinehart, T. Nguyen, J. Ray, J. Kuehl, B. Arevalo, P.D. Westenskow, J. Wang, A.P. Arkin, A.M. Deutschbauer, G.J. Patti, G. Siuzdak, Interactive XCMS online: Simplifying advanced metabolomic data processing and subsequent statistical analyses, Anal. Chem. 86 (2014) 6931–6939. doi:10.1021/ac500734c.

[39] T.J. Passe, H.C. Charles, P. Rajagopalan, K.R. Krishnan, P.R. THEODORE .I. PASSE, H. CECIL CHARLES, K. RANGA K, Nuclear magnetic resonance

spectroscopy: A review of neuropsychiatric applications, Prog. Neuro-Psychopharmacology Biol. Psychiatry. 19 (1995) 541–563. doi:10.1016/0278-5846(95)00101-Z.

[40] E.J. Want, G. O'Maille, C.A. Smith, T.R. Brandon, W. Uritboonthai, C. Qin, S.A. Trauger, G. Siuzdak, Solvent-dependent metabolite distribution, clustering, and protein extraction for serum profiling with mass spectrometry, Anal. Chem. 78 (2006) 743–752. doi:10.1021/ac051312t.

[41] J.C. Chatham, S.J. Blackband, Nuclear Magnetic Resonance Spectroscopy and Imaging in Animal Research, ILAR J. 42 (2001) 189–208. doi:10.1093/ilar.42.3.189.

[42] J.C. Lindon, J.K. Nicholson, I.D. Wilson, Direct coupling of chromatographic separations to NMR spectroscopy, Prog. Nucl. Magn. Reson. Spectrosc. 29 (1996) 1–49. doi:10.1016/0079-6565(95)01027-0.

[43] G.A. Nagana Gowda, Y.N. Gowda, D. Raftery, Expanding the Limits of Human Blood Metabolite Quantitation Using NMR Spectroscopy, Anal. Chem. 87 (2015) 706–715. doi:10.1021/ac503651e.

[44] E. Werner, J.F. Heilier, C. Ducruix, E. Ezan, C. Junot, J.C. Tabet, Mass spectrometry for the identification of the discriminating signals from metabolomics: Current status and future trends, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 871 (2008) 143–163. doi:10.1016/j.jchromb.2008.07.004.

[45] B. Zhou, J.F. Xiao, L. Tuli, H.W. Ressom, LC-MS-based metabolomics, Mol Biosyst. 8 (2013) 470–481. doi:10.1016/j.immuni.2010.12.017.Two-stage.

[46] M. Commisso, A. Anesi, S. Dal Santo, F. Guzzo, Performance comparison of electrospray ionization and atmospheric pressure chemical ionization in untargeted and targeted liquid chromatography/mass spectrometry based metabolomics analysis of grapeberry metabolites, Rapid Commun. Mass Spectrom. 31 (2017) 292–300. doi:10.1002/rcm.7789.

[47] K.K. Pasikanti, P.C. Ho, E.C.Y. Chan, Gas chromatography/mass spectrometry in metabolic profiling of biological fluids, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 871 (2008) 202–211. doi:10.1016/j.jchromb.2008.04.033.

[48] R.A. Zubarev, A. Makarov, Orbitrap Mass Spectrometry, Anal. Chem. 85 (2013) 5288–5296. doi:10.1021/ac4001223.

[49] C.F. Klitzke, Y.E. Corilo, K. Siek, J. Binkley, J. Patrick, M.N. Eberlin, Petroleomics by Ultrahigh-Resolution Time-of-Flight Mass Spectrometry, Energy & Fuels. 26 (2012) 5787–5794. doi:10.1021/ef300961c.

[50] M. Scigelova, M. Hornshaw, A. Giannakopulos, A. Makarov, Fourier transform mass spectrometry, Mol. Cell. Proteomics. (2011) 1–19. doi:10.1074/ mcp.M111.009431.

[51] A. Makarov, Eduard Denisov, Alexander Kholomeev, Wilko Balschun, Oliver Lange, A. Kerstin Strupat, S. Horning, Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer, (2006). doi:10.1021/AC0518811.

[52] K. Richardson, J. Hoyes, A novel multipass oa-TOF mass spectrometer, Int. J. Mass Spectrom. 377 (2015) 309–315. doi:10.1016/J.IJMS.2014.08.031.

[53] W. Lu, B.D. Bennett, J.D. Rabinowitz, Analytical strategies for LC-MS based targeted metabolomics, J.Chromatogr.B. . (2008) 236–242. doi:10.1016/j.jchromb.2008.04.031.Analytical.

[54] J.C. Lindon, J.K. Nicholson, E. Holmes, The Handbook of Metabonomics and Metabolomics, 2007. doi:10.1016/B978-0-444-52841-4.X5000-0.

[55] S. Böcker, M.C. Letzel, Z. Lipták, A. Pervukhin, SIRIUS: Decomposing isotope patterns for metabolite identification, Bioinformatics. 25 (2009) 218–224. doi:10.1093/bioinformatics/btn603.

[56] T. Kind, O. Fiehn, Metabolomic database annotations via query of elemental

compositions: mass accuracy is insufficient even at less than 1 ppm., BMC Bioinformatics. 7 (2006) 234. doi:10.1186/1471-2105-7-234.

[57] K.K. Murray, R.K. Boyd, M.N. Eberlin, G.J. Langley, L. Li, Y. Naito, Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013), Pure Appl. Chem. 85 (2013) 1515–1609. doi:10.1351/PAC-REC-06-04-06.

[58] J. Wang, T.T. Christison, K. Misuno, L. Lopez, A.F. Huhmer, Y. Huang, S. Hu, Metabolomic Pro fi ling of Anionic Metabolites in Head and Neck Cancer Cells by Capillary Ion Chromatography with Orbitrap Mass Spectrometry, Anal. Chem. 86 (2014) 5116–5124. doi:10.1021/ac500951v.

[59] X. Liu, Z. Ser, A.A. Cluntun, S.J. Mentch, J.W. Locasale, A strategy for sensitive, large scale quantitative metabolomics., J. Vis. Exp. (2014). doi:10.3791/51358.

[60] S. Weidt, J. Haggarty, R. Kean, C.I. Cojocariu, P.J. Silcock, R. Rajendran, G. Ramage, K.E.V. Burgess, A novel targeted/untargeted GC-Orbitrap metabolomics methodology applied to Candida albicans and Staphylococcus aureus biofilms, Metabolomics. 12 (2016) 1–10. doi:10.1007/s11306-016-1134-2.

[61] E. Grata, D. Guillarme, G. Glauser, J. Boccard, P. Carrupt, J.-L. Veuthey, S. Rudaz, J.-L. Wolfender, Metabolite profiling of plant extracts by ultra-high-pressure liquid chromatography at elevated temperature coupled to time-of-flight mass spectrometry, J. Chromatogr. A. 1216 (2009) 5660–5668. doi:10.1016/j.chroma.2009.05.069.

[62] G.X. Xie, Y. Ni, M.M. Su, Y.Y. Zhang, A.H. Zhao, X.F. Gao, Z. Liu, P.G. Xiao, W. Jia, Application of ultra-performance LC-TOF MS metabolite profiling techniques to the analysis of medicinal Panax herbs, Metabolomics. 4 (2008) 248–260. doi:10.1007/s11306-008-0115-5.

[63] E. Zelena, W.B. Dunn, D.I. Broadhurst, S. Francis-McIntyre, K.M. Carroll, P. Begley, S. O'Hagan, J.D. Knowles, A. Halsall, I.D. Wilson, D.B. Kell, Development of a Robust and Repeatable UPLC−MS Method for the Long-Term Metabolomic

Study of Human Serum, Anal. Chem. 81 (2009) 1357–1364. doi:10.1021/ac8019366.

[64] Y. Qi, Y. Song, H. Gu, G. Fan, Y. Chai, Global Metabolic Profi ling Using Ultra-Performance Liquid Chromatography/Quadrupole Time-of-Flight Mass Spectrometry Yunpeng, in: Methods Mol. Biol., 2012: pp. 271–298. doi:10.1002/9781119977438.ch12.

[65] D. Siegel, H. Permentier, D. Reijngoud, R. Bischoff, Chemical and technical challenges in the analysis of central carbon metabolites by liquid-chromatography mass spectrometry ☪, J. Chromatogr. B. 966 (2014) 21–33. doi:10.1016/j.jchromb.2013.11.022.

[66] M.M. Delgado-Povedano, M. Calderón-Santiago, F. Priego-Capote, M.D. Luque De Castro, Study of sample preparation for quantitative analysis of amino acids in human sweat by liquid chromatography-tandem mass spectrometry, Talanta. 146 (2016) 310–317. doi:10.1016/j.talanta.2015.07.066.

[67] B.B. Schneider, D.J. Douglas, D.D.Y. Chen, Ion fragmentation in an electrospray ionization mass spectrometer interface with different gases, Rapid Commun. Mass Spectrom. 15 (2001) 249–257. doi:10.1002/rcm.221.

[68] L.E. Sojo, G. Lum, P. Chee, Internal standard signal suppression by co-eluting analyte in isotope dilution LC-ESI-MS, Analyst. 128 (2003) 51–54. doi:10.1039/b209521c.

[69] M. Watanabe, M. Kusano, A. Oikawa, A. Fukushima, M. Noji, K. Saito, Physiological roles of the beta-substituted alanine synthase gene family in Arabidopsis., Plant Physiol. 146 (2008) 310–20. doi:10.1104/pp.107.106831.

[70] K. Contrepois, L. Jiang, M. Snyder, Optimized Analytical Procedures for the Untargeted Metabolomic Profiling of Human Urine and Plasma by Combining Hydrophilic Interaction (HILIC) and Reverse-Phase Liquid Chromatography

(RPLC)–Mass Spectrometry, Mol. Cell. Proteomics. 14 (2015) 1684–1695. doi:10.1074/mcp.M114.046508.

[71] J.T. Bjerrum, ed., Metabonomics Methods and Protocols, Humana Press, 2015. doi:10.1007/978-1-4939-2377-9.

[72] Soga T, Ueno Y, Naraoka H, Ohashi Y, Tomita M, Nishioka T, Simultaneous determination of anionic intermediates for Bacillus subtilis metabolic pathways by capillary electrophoresis electrospray ionization m... - PubMed - NCBI, (n.d.). https://www.ncbi.nlm.nih.gov/pubmed/12038746 (accessed November 22, 2018).

[73] N.L. Kuehnbaum, J.B. Gillen, A. Kormendi, K.P. Lam, A. DiBattista, M.J. Gibala, P. Britz-McKibbin, Multiplexed separations for biomarker discovery in metabolomics: Elucidating adaptive responses to exercise training, Electrophoresis. 36 (2015) 2226–2236. doi:10.1002/elps.201400604.

[74] L. Hajba, A. Guttman, Recent advances in column coatings for capillary electrophoresis of proteins, TrAC Trends Anal. Chem. 90 (2017) 38–44. doi:10.1016/J.TRAC.2017.02.013.

[75] R. Ramautar, G.W. Somsen, G.J. de Jong, CE–MS for metabolomics: Developments and applications in the period 2014–2016, Electrophoresis. 38 (2017) 190–202. doi:10.1002/elps.201600370.

[76] K. Harada, E. Fukusaki, Profiling of primary metabolite by means of capillary electrophoresis-mass spectrometry and its application for plant science, Plant Biotechnol. 26 (2009) 47–52. doi:10.5511/plantbiotechnology.26.47.

[77] M.R.N. Monton, T. Soga, Metabolome analysis by capillary electrophoresis-mass spectrometry, J. Chromatogr. A. 1168 (2007) 237–246. doi:10.1016/j.chroma.2007.02.065.

[78] E. Gorrochategui, J. Jaumot, S. Lacorte, R. Tauler, Data analysis strategies for

targeted and untargeted LC-MS metabolomic studies: overview and workflow, Trends Anal. Chem. (2016). doi:10.1016/j.trac.2016.07.004.

[79] E. Zarate, V. Boyle, U. Rupprecht, S. Green, S.G. Villas-Boas, P. Baker, F.R. Pinu, Fully automated trimethylsilyl (TMS) derivatisation protocol for metabolite profiling by GC-MS, Metabolites. 7 (2017) 1–13. doi:10.3390/metabo7010001.

[80] M.R. Nezami Ranjbar, Y. Luo, C. Di Poto, R.S. Varghese, A. Ferrarini, C. Zhang, N.I. Sarhan, H. Soliman, M.G. Tadesse, D.H. Ziada, R. Roy, H.W. Ressom, GC-MS Based Plasma Metabolomics for Identification of Candidate Biomarkers for Hepatocellular Carcinoma in Egyptian Cohort., PLoS One. 10 (2015) e0127299. doi:10.1371/journal.pone.0127299.

[81] M.M. Rahman, A.M. Abd El-Aty, J.-H. Choi, H.-C. Shin, S.C. Shin, J.-H. Shim, Basic Overview on Gas Chromatography Columns, in: Anal. Sep. Sci., Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2015: pp. 823–834. doi:10.1002/9783527678129.assep024.

[82] E. Stauffer, J.A. Dolan, R. Newman, Gas Chromatography and Gas Chromatography—Mass Spectrometry, Fire Debris Anal. (2008) 235–293. doi:10.1016/b978-012663971-1.50012-9.

[83] W. Pongsuwan, E. Fukusaki, T. Bamba, T. Yonetani, T. Yamahara, A. Kobayashi, Prediction of Japanese Green Tea Ranking by Gas Chromatography/Mass Spectrometry-Based Hydrophilic Metabolite Fingerprinting, J. Agric. Food Chem. 55 (2007) 231–236. doi:10.1021/JF062330U.

[84] S. Carneiro, R. Pereira, I. Rocha, Yeast Metabolomics: Sample Preparation for a GC/MS-Based Analysis, in: Methods Mol. Biol., 2014: pp. 197–207. doi:10.1007/978-1-4939-0563-8_12.

[85] T. Yoshida, Analytical Method for Pyrethroid Metabolites in Urine of the Non-Occupationally Exposed Population by Gas Chromatography/Mass Spectrometry, J.

Chromatogr. Sci. 55 (2017) 873–881. doi:10.1093/chromsci/bmx048.

[86] M. Kim, N.R. Song, J.-H. Choi, J. Lee, H. Pyo, Simultaneous analysis of urinary phthalate metabolites of residents in Korea using isotope dilution gas chromatography–mass spectrometry, Sci. Total Environ. 470–471 (2014) 1408–1413. doi:10.1016/j.scitotenv.2013.07.037.

[87] P. Favetta, C. Dufresne, M. Désage, O. Paissé, J.P. Perdrix, R. Boulieu, J. Guitton, Detection of new propofol metabolites in human urine using gas chromatography/mass spectrometry and liquid chromatography/mass spectrometry techniques, Rapid Commun. Mass Spectrom. 14 (2000) 1932–1936. doi:10.1002/1097-0231(20001030)14:20<1932::AID-RCM114>3.0.CO;2-P.

[88] S.G. Musharraf, S. Mazhar, A.J. Siddiqui, M.I. Choudhary, Atta-ur-Rahman, Metabolite profiling of human plasma by different extraction methods through gas chromatography–mass spectrometry—An objective comparison, Anal. Chim. Acta. 804 (2013) 180–189. doi:10.1016/j.aca.2013.10.025.

[89] J. Lisec, N. Schauer, J. Kopka, L. Willmitzer, A.R. Fernie, Gas chromatography mass spectrometry–based metabolite profiling in plants, Nat. Protoc. 1 (2006) 387–396. doi:10.1038/nprot.2006.59.

[90] U. Roessner, C. Wagner, J. Kopka, R.N. Trethewey, L. Willmitzer, Technical advance: simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry., Plant J. 23 (2000) 131–42. http://www.ncbi.nlm.nih.gov/pubmed/10929108 (accessed November 23, 2018).

[91] M.A. Farag, A.R. Khattab, A. Ehrlich, M. Kropf, A.G. Heiss, L.A. Wessjohann, Gas Chromatography/Mass Spectrometry-Based Metabolite Profiling of Nutrients and Antinutrients in Eight *Lens* and *Lupinus* Seeds (Fabaceae), J. Agric. Food Chem. 66 (2018) 4267–4280. doi:10.1021/acs.jafc.8b00369.

[92] D.B. Chu, C. Troyer, T. Mairinger, K. Ortmayr, S. Neubauer, G. Koellensperger, S.

Hann, Isotopologue analysis of sugar phosphates in yeast cell extracts by gas chromatography chemical ionization time-of-flight mass spectrometry, Anal. Bioanal. Chem. 407 (2015) 2865–2875. doi:10.1007/s00216-015-8521-9.

[93] T. Mairinger, M. Steiger, J. Nocon, D. Mattanovich, G. Koellensperger, S. Hann, Gas Chromatography-Quadrupole Time-of-Flight Mass Spectrometry-Based Determination of Isotopologue and Tandem Mass Isotopomer Fractions of Primary Metabolites for 13C-Metabolic Flux Analysis, Anal. Chem. 87 (2015) 11792–11802. doi:10.1021/acs.analchem.5b03173.

[94] H.F.N. Kvitvang, K.A. Kristiansen, P. Bruheim, Assessment of capillary anion exchange ion chromatography tandem mass spectrometry for the quantitative profiling of the phosphometabolome and organic acids in biological extracts, J. Chromatogr. A. 1370 (2014) 70–79. doi:10.1016/j.chroma.2014.10.029.

[95] Y. Sekiguchi, N. Mitsuhashi, Y. Inoue, H. Yagisawa, T. Mimura, Analysis of sugar phosphates in plants by ion chromatography on a titanium dioxide column with pulsed amperometric detection, J. Chromatogr. A. 1039 (2004) 71–76. doi:10.1016/j.chroma.2004.02.015.

[96] P. Aronov, Y. Kawahara, J. Cox, T. Christison, L. Lopez, Techincal note: Determination of NADH and NADPH Using Ion Chromatography and High Resolution Accurate Mass Spectrometry, 2014.

[97] T. Christison, J. Wang, Untargeted Metabolic Profiling of Oral Cancer Cells Using Capillary Ion Chromatography Coupled with an Orbitrap Mass Spectrometer, (n.d.).

[98] F.J. Naser, N.G. Mahieu, L. Wang, J.L. Spalding, S.L. Johnson, G.J. Patti, Two complementary reversed-phase separations for comprehensive coverage of the semipolar and nonpolar metabolome, Anal. Bioanal. Chem. (2017) 1–11. doi:10.1007/s00216-017-0768-x.

[99] J.M. Onorato, R. Langish, A. Bellamine, P. Shipkova, Applications of HILIC for

targeted and non-targeted LC/MS analyses in drug discovery, J. Sep. Sci. 33 (2010) 923–929. doi:10.1002/jssc.200900659.

[100] S. Tufi, M. Lamoree, J. de Boer, P. Leonards, Simultaneous analysis of multiple neurotransmitters by hydrophilic interaction liquid chromatography coupled to tandem mass spectrometry, J. Chromatogr. A. 1395 (2015) 79–87. doi:10.1016/J.CHROMA.2015.03.056.

[101] D.-Q. Tang, X.-X.Y. Ll Zou, Choon Nam Ong, HILIC-MS FOR METABOLOMICS: AN ATTRACTIVE AND COMPLEMENTARY APPROACH TO RPLC-MS, Mass Spectrom. Rev. (2016) 574–600. doi:DOI 10.1002/mas.21445.

[102] T. Sun, S.J. Wetzel, M.E. Johnson, B.A. Surlow, J. Patton-Vogt, Development and validation of a hydrophilic interaction liquid chromatography-tandem mass spectrometry method for the quantification of lipid-related extracellular metabolites in Saccharomyces cerevisiae, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 897 (2012) 1–9. doi:10.1016/j.jchromb.2012.03.034.

[103] A. Jungbauer, R. Hahn, Chapter 22 Ion-Exchange Chromatography, Methods Enzymol. 463 (2009) 349–371. doi:10.1016/S0076-6879(09)63022-6.

[104] P.M. Cummins, K.D. Rochfort, Brendan F. O'Connor, Ion-Exchange Chromatography: Basic Principles and Application, Protein Chromatogr. 1485 (2010) 209–223. doi:10.1002/9783527630158.

[105] B. Buszewski, S. Noga, Hydrophilic interaction liquid chromatography (HILIC)--a powerful separation technique., Anal. Bioanal. Chem. 402 (2012) 231–47. doi:10.1007/s00216-011-5308-5.

[106] A.J. Alpert, Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds., J. Chromatogr. 499 (1990) 177–196. http://www.ncbi.nlm.nih.gov/pubmed/2324207 (accessed February 26, 2019).

[107] A.C. Schrimpe-Rutledge, S.G. Codreanu, S.D. Sherrod, J.A. McLean, Untargeted Metabolomics Strategies???Challenges and Emerging Directions, J. Am. Soc. Mass Spectrom. 27 (2016) 1897–1905. doi:10.1007/s13361-016-1469-y.

[108] S. Naz, M. Vallejo, A. García, C. Barbas, Method validation strategies involved in non-targeted metabolomics, J. Chromatogr. A. 1353 (2014) 99–105. doi:10.1016/j.chroma.2014.04.071.

[109] L.D. Roberts, A.L. Souza, R.E. Gerszten, C.B. Clish, Targeted Metabolomics, 2012. doi:10.1002/0471142727.mb3002s98.Targeted.

[110] W. de Koning, K. van Dam, A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH., Anal. Biochem. 204 (1992) 118–23. http://www.ncbi.nlm.nih.gov/pubmed/1514678 (accessed November 23, 2018).

[111] A.B. Canelas, C. Ras, A. ten Pierick, J.C. van Dam, J.J. Heijnen, W.M. van Gulik, Leakage-free rapid quenching technique for yeast metabolomics, Metabolomics. 4 (2008) 226–239. doi:10.1007/s11306-008-0116-4.

[112] S.G. Villas-Boas, J. Højer-Pedersen, M. Akesson, J. Smedsgaard, J. Nielsen, Global metabolite analysis of yeast: Evaluation of sample preparation methods, Yeast. 22 (2005) 1155–1169. doi:10.1002/yea.1308.

[113] C. Wittmann, J.O. Krömer, P. Kiefer, T. Binz, E. Heinzle, Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria, Anal. Biochem. 327 (2004) 135–139. doi:10.1016/j.ab.2004.01.002.

[114] O. Yanes, R. Tautenhahn, G.J. Patti, G. Siuzdak, Expanding coverage of the metabolome for global metabolite profiling, Anal. Chem. 83 (2011) 2152–2161. doi:10.1021/ac102981k.

[115] A.K. Mathew, V.C. Padmanaban, Metabolomics: The apogee of the omics trilogy,

Int. J. Pharm. Pharm. Sci. 5 (2013) 45–48. doi:10.1038/nrm3314.

[116] W.M.A. Niessen, D. Falck, Introduction to Mass Spectrometry, a Tutorial, in: 2015: pp. 1–54. doi:doi:10.1002/9783527673391.ch1.

[117] M. Mann, S. Ong, M. Gr, H. Steen, O.N. Jensen, A. Pandey, Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome., Trends Biotechnol. 20 (2002) 261–8. doi:10.1016/S0167-7799(02)01944-3.

[118] H. Tsugawa, T. Cajka, T. Kind, Y. Ma, B. Higgins, K. Ikeda, M. Kanazawa, J. VanderGheynst, O. Fiehn, M. Arita, MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis, Nat. Methods. 12 (2015) 523–526. doi:10.1038/nmeth.3393.

[119] O.D. Myers, S.J. Sumner, S. Li, S. Barnes, X. Du, Detailed Investigation and Comparison of the XCMS and MZmine 2 Chromatogram Construction and Chromatographic Peak Detection Methods for Preprocessing Mass Spectrometry Metabolomics Data, Anal. Chem. 89 (2017) 8689–8695. doi:10.1021/acs.analchem.7b01069.

[120] R. Smith, D. Ventura, J.T. Prince, LC-MS alignment in theory and practice: A comprehensive algorithmic review, Brief. Bioinform. 16 (2013) 104–117. doi:10.1093/bib/bbt080.

[121] E.G. Armitage, J. Godzien, V. Alonso-Herranz, ??ngeles L??pez-Gonz??lvez, C. Barbas, Missing value imputation strategies for metabolomics data, Electrophoresis. 36 (2015) 3050–3060. doi:10.1002/elps.201500352.

[122] R. Goodacre, S. Vaidyanathan, W.B. Dunn, G.G. Harrigan, D.B. Kell, Metabolomics by numbers: Acquiring and understanding global metabolite data, Trends Biotechnol. 22 (2004) 245–252. doi:10.1016/j.tibtech.2004.03.007.

[123] M. Otto, Chemometrics: statistics and computer application in analytical chemistry,

3rd ed., 2016. doi:10.1007/s00216-017-0503-7.

[124] M.A. Samaraweera, L.M. Hall, D.W. Hill, D.F. Grant, Evaluation of an Artificial Neural Network Retention Index Model for Chemical Structure Identification in Non-Targeted Metabolomics, Anal. Chem. (2018) acs.analchem.8b03118. doi:10.1021/acs.analchem.8b03118.

[125] A. Sussulini, W.D. Jonathan, W.T. Mary, S.J. Lauren, Metabolomics: From Fundamentals to Clinical Applications, Mycotoxins Food Saf. 965 (2017) 223. doi:10.1007/978-3-319-47656-8.

[126] P. Arapitsas, A. Della Corte, H. Gika, L. Narduzzi, F. Mattivi, G. Theodoridis, Studying the effect of storage conditions on the metabolite content of red wine using HILIC LC-MS based metabolomics, Food Chem. 197 (2016) 1331–1340. doi:10.1016/j.foodchem.2015.09.084.

[127] R. Bade, L. Bijlsma, T.H. Miller, L.P. Barron, J.V. Sancho, F. Hernández, Suspect screening of large numbers of emerging contaminants in environmental waters using artificial neural networks for chromatographic retention time prediction and high resolution mass spectrometry data analysis, Sci. Total Environ. 538 (2015) 934–941. doi:10.1016/j.scitotenv.2015.08.078.

[128] E.L. Schymanski, J. Jeon, R. Gulde, K. Fenner, M. Ruff, H.P. Singer, J. Hollender, Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence, Environ. Sci. Technol. 48 (2014) 2097–2098. doi:10.1021/es5002105.

[129] T. Kind, O. Fiehn, Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry., BMC Bioinformatics. 8 (2007) 105. doi:10.1186/1471-2105-8-105.

[130] D.S. Wishart, Y.D. Feunang, A. Marcu, A.C. Guo, K. Liang, R. Vázquez-Fresno, T. Sajed, D. Johnson, C. Li, N. Karu, Z. Sayeeda, E. Lo, N. Assempour, M. Berjanskii,

S. Singhal, D. Arndt, Y. Liang, H. Badran, J. Grant, A. Serra-Cayuela, Y. Liu, R. Mandal, V. Neveu, A. Pon, C. Knox, M. Wilson, C. Manach, A. Scalbert, HMDB 4.0: the human metabolome database for 2018, Nucleic Acids Res. 46 (2018) D608–D617. doi:10.1093/nar/gkx1089.

[131] K. Ortmayr, G. Koellensperger, S. Hann, Complementing reversed-phase selectivity with porous graphitized carbon to increase the metabolome coverage in an on-line two- dimensional LC-MS setup for metabolomics, R. Soc. Chem. (2015). doi:10.1039/c5an00206k.

[132] L. Si-Hung, T. Causon, S. Hann, Comparison of Fully-Wettable RPLC Stationary Phases for LC-MS-based Cellular Metabolomics, Electrophoresis. 13 (2017). doi:10.1002/elan.

[133] H.G. Gika, G.A. Theodoridis, M. Earll, R.W. Snyder, S.J. Sumner, I.D. Wilson, Does the Mass Spectrometer Define the Marker? A Comparison of Global Metabolite Profiling Data Generated Simultaneously via UPLC-MS on Two Different Mass Spectrometers, Anal. Chem. 82 (2010) 8226–8234. doi:10.1021/ac1016612.

[134] J. Godzien, M. Ciborowski, S. Angulo, C. Barbas, From numbers to a biological sense: How the strategy chosen for metabolomics data treatment may affect final results. A practical example based on urine fingerprints obtained by LC-MS, Electrophoresis. 34 (2013) 2812–2826. doi:10.1002/elps.201300053.

[135] T. Sangster, H. Major, R. Plumb, A.J. Wilson, I.D. Wilson, A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis, Analyst. 131 (2006) 1075. doi:10.1039/b604498k.

[136] J. Godzien, V. Alonso-Herranz, C. Barbas, E.G. Armitage, J.G.V. Alonso-herranz, C. Barbas, E. Grace, Controlling the quality of metabolomics data: new strategies to get the best out of the QC sample, Metabolomics. 11 (2015) 518–528. doi:10.1007/s11306-014-0712-4.

[137] R. Tautenhahn, C. Böttcher, S. Neumann, Highly sensitive feature detection for high resolution LC/MS, BMC Bioinformatics. 9 (2008) 504. doi:10.1186/1471-2105-9-504.

[138] M.K.R. Engskog, J. Haglöf, T. Arvidsson, C. Pettersson, LC – MS based global metabolite profiling : the necessity of high data quality, Metabolomics. 12 (2016). doi:10.1007/s11306-016-1058-x.

[139] O. Fiehn, D. Robertson, J. Griffin, M. van der Werf, B. Nikolau, N. Morrison, L.W. Sumner, R. Goodacre, N.W. Hardy, C. Taylor, J. Fostel, B. Kristal, R. Kaddurah-Daouk, P. Mendes, B. van Ommen, J.C. Lindon, S.-A. Sansone, The metabolomics standards initiative (MSI), Metabolomics. 3 (2007) 175–178. doi:10.1007/s11306-007-0070-6.

[140] E. Noor, E. Eden, R. Milo, U. Alon, Central Carbon Metabolism as a Minimal Biochemical Walk between Precursors for Biomass and Energy, Mol. Cell. 39 (2010) 809–820. doi:10.1016/J.MOLCEL.2010.08.031.

[141] Q. Qin, B. Wang, M. Chang, Z. Zhou, X. Shi, G. Xu, Highly efficient solid-phase derivatization of sugar phosphates with titanium-immobilized hydrophilic polydopamine-coated silica, J. Chromatogr. A. 1457 (2016) 125–133. doi:10.1016/j.chroma.2016.06.046.

[142] D. (Dietmar) Schomburg, G. Michal, Biochemical pathways : an atlas of biochemistry and molecular biology, John Wiley & Sons, 2012. https://www.wiley.com/en-us/Biochemical+Pathways%3A+An+Atlas+of+Biochemistry+and+Molecular+Biology%2C+2nd+Edition-p-9780470146842 (accessed October 17, 2018).

[143] G. Valentini, L. Chiarelli, R. Fortin, M.L. Speranza, A. Galizzi, A. Mattevi, The allosteric regulation of pyruvate kinase., J. Biol. Chem. 275 (2000) 18145–52. doi:10.1074/jbc.M001870200.

[144] N.J. Kruger, A. Von Schaewen, The oxidative pentose phosphate pathway: Structure and organisation, Curr. Opin. Plant Biol. 6 (2003) 236–246. doi:10.1016/S1369-5266(03)00039-6.

[145] J.H.J. Huck, E.A. Struys, N.M. Verhoeven, C. Jakobs, M.S. Van Der Knaap, Profiling of pentose phosphate pathway intermediates in blood spots by tandem mass spectrometry: Application to transaldolase deficiency, Clin. Chem. 49 (2003) 1375–1380. doi:10.1373/49.8.1375.

[146] D.B. Chu, K. Klavins, G. Koellensperger, S. Hann, Speciation analysis of sugar phosphates via anion exchange chromatography combined with inductively coupled plasma dynamic reaction cell mass spectrometry – optimization for the analysis of yeast cell extracts, J. Anal. At. Spectrom. 29 (2014) 915. doi:10.1039/c4ja00043a.

[147] F. Desmoulin, J.-P. Galons, P. Canioni, 31P Nuclear Magnetic Resonance Study of a Human Colon Adenocarcinoma Cultured Cell Line, (n.d.).

[148] T.R. BROWN, R.A. GRAHAM, B.S. SZWERGOLD, W.J. THOMA, R.A. MEYER, Phosphorylated Metabolites in Tumors, Tissues, and Cell Lines, Ann. N. Y. Acad. Sci. 508 (1987) 229–240. doi:10.1111/j.1749-6632.1987.tb32907.x.

[149] E.C. Horning, M.G. Horning, Metabolic profiles: gas-phase methods for analysis of metabolites., Clin. Chem. 17 (1971) 802–809.

[150] A. Sanz-Medel, M. Montes-Bayón, M. del Rosario Fernández de la Campa, J.R. Encinar, J. Bettmer, Elemental mass spectrometry for quantitative proteomics, Anal. Bioanal. Chem. 390 (2008) 3–16. doi:10.1007/s00216-007-1615-2.

[151] E. Maes, K. Tirez, G. Baggerman, D. Valkenborg, L. Schoofs, J.R. Encinar, I. Mertens, The use of elemental mass spectrometry in phosphoproteomic applications, Mass Spectrom. Rev. 35 (2016) 350–360. doi:10.1002/mas.21440.

[152] J.S.B. and, H.-J. Dietze, J.A.M. and, A. Montaser*, Ultratrace and Isotope Analysis

of Long-Lived Radionuclides by Inductively Coupled Plasma Quadrupole Mass Spectrometry Using a Direct Injection High Efficiency Nebulizer, (1999). doi:10.1021/AC9900883.

[153] N. Arashida, R. Nishimoto, M. Harada, K. Shimbo, N. Yamada, Highly sensitive quantification for human plasma-targeted metabolomics using an amine derivatization reagent, Anal. Chim. Acta. 954 (2017) 77–87. doi:10.1016/J.ACA.2016.11.068.

[154] B. Klencsár, S. Li, L. Balcaen, F. Vanhaecke, High-performance liquid chromatography coupled to inductively coupled plasma – Mass spectrometry (HPLC-ICP-MS) for quantitative metabolite profiling of non-metal drugs, TrAC Trends Anal. Chem. 104 (2018) 118–134. doi:10.1016/J.TRAC.2017.09.020.

[155] T.W. May, R.H. Wiedmeyer, M. Chaudhary-webb, D.C. Paschal, W.C. Elliott, H.P. Hopkins, a M. Ghazi, B.C. Ting, I. Romieu, O. Vicente, E. Pelfort, L. Martinez, R. Olsina, E. Marchevsky, H.P. Chen, D.T. Miller, J.C. Morrow, A table of polyatimic interferences in ICP-MS, At. Spectrosc. 19 (1998) 150–155.

[156] P. Vizán, G. Alcarraz-Vizán, S. Díaz-Moralli, J.C. Rodríguez-Prados, M. Zanuy, J.J. Centelles, O. Jáuregui, M. Cascante, Quantification of intracellular phosphorylated carbohydrates in HT29 human colon adenocarcinoma cell line using liquid chromatography-electrospray ionization tandem mass spectrometry, Anal. Chem. 79 (2007) 5000–5005. doi:10.1021/ac070170v.

[157] C. Antonio, T. Larson, A. Gilday, I. Graham, E. Bergström, J. Thomas-Oates, Quantification of sugars and sugar phosphates in Arabidopsis thaliana tissues using porous graphitic carbon liquid chromatography-electrospray ionization mass spectrometry, J. Chromatogr. A. 1172 (2007) 170–178. doi:10.1016/j.chroma.2007.10.011.

[158] B. Luo, K. Groenke, R. Takors, C. Wandrey, M. Oldiges, Simultaneous

determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography-mass spectrometry, J. Chromatogr. A. 1147 (2007) 153–164. doi:10.1016/j.chroma.2007.02.034.

[159] F. Michopoulos, N. Whalley, G. Theodoridis, I.D. Wilson, T.P.J. Dunkley, S.E. Critchlow, Targeted profiling of polar intracellular metabolites using ion-pair-high performance liquid chromatography and -ultra high performance liquid chromatography coupled to tandem mass spectrometry: Applications to serum, urine and tissue extracts, J. Chromatogr. A. 1349 (2014) 60–68. doi:10.1016/j.chroma.2014.05.019.

[160] K. Ortmayr, T.J. Causon, S. Hann, G. Koellensperger, Increasing selectivity and coverage in LC-MS based metabolome analysis, TrAC Trends Anal. Chem. 82 (2016) 358–366. doi:10.1016/j.trac.2016.06.011.

[161] P. Deng, Y. Zhan, X. Chen, D. Zhong, Derivatization methods for quantitative bioanalysis by LC–MS/MS, Bioanalysis. 4 (2012) 49–69. doi:10.4155/bio.11.298.

[162] O. Vielhauer, M. Zakhartsev, T. Horn, R. Takors, M. Reuss, Simplified absolute metabolite quantification by gas chromatography–isotope dilution mass spectrometry on the basis of commercially available source material, J. Chromatogr. B. 879 (2011) 3859–3870. doi:10.1016/J.JCHROMB.2011.10.036.

[163] J. Schaub, C. Schiesling, M. Reuss, M. Dauner, Integrated Sampling Procedure for Metabolome Analysis, Biotechnol. Prog. 22 (2008) 1434–1442. doi:10.1021/bp050381q.

[164] C. Bylda, R. Thiele, U. Kobold, D.A. Volmer, Recent advances in sample preparation techniques to overcome difficulties encountered during quantitative analysis of small molecules from biofluids using LC-MS/MS, Analyst. 139 (2014) 2265. doi:10.1039/c4an00094c.

[165] K. Dettmer, P.A. Aronov, B.D. Hammock, Mass spectrometry-based metabolomics.,

Mass Spectrom. Rev. 26 (2007) 51–78. doi:10.1002/mas.20108.

[166] X. Zhang, M. Romm, X. Zheng, E.M. Zink, Y.-M. Kim, K.E. Burnum-Johnson, D.J. Orton, A. Apffel, Y.M. Ibrahim, M.E. Monroe, R.J. Moore, J.N. Smith, J. Ma, R.S. Renslow, D.G. Thomas, A.E. Blackwell, G. Swinford, J. Sausen, R.T. Kurulugama, N. Eno, E. Darland, G. Stafford, J. Fjeldsted, T.O. Metz, J.G. Teeguarden, R.D. Smith, E.S. Baker, SPE-IMS-MS: An automated platform for sub-sixty second surveillance of endogenous metabolites and xenobiotics in biofluids, Clin. Mass Spectrom. 2 (2016) 1–10. doi:10.1016/J.CLINMS.2016.11.002.

[167] J. Fíla, D. Honys, J. Fílla, D. Honys, Enrichment techniques employed in phosphoproteomics., Amino Acids. 43 (2012) 1025–47. doi:10.1007/s00726-011-1111-z.

[168] M.W.H. Pinkse, P.M. Uitto, M.J. Hilhorst, B. Ooms, A.J.R. Heck, Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns, Anal. Chem. 76 (2004) 3935–3943. doi:10.1021/ac0498617.

[169] U.K. Aryal, A.R.S. Rossz, Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry, Rapid Commun. Mass Spectrom. 24 (2009) 3567–3577. doi:10.1002/rcm.

[170] H.K. Kweon, K. Håkansson, Selective zirconium dioxide-based enrichment of phosphorylated peptides for mass spectrometric analysis, Anal. Chem. 78 (2006) 1743–1749. doi:10.1021/ac0522355.

[171] S.R. Schmidt, F. Schweikart, M.E. Andersson, Current methods for phosphoprotein isolation and enrichment, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 849 (2007) 154–162. doi:10.1016/j.jchromb.2006.09.016.

[172] N. Sugiyama, T. Masuda, K. Shinoda, A. Nakamura, M. Tomita, Y. Ishihama,

Phosphopeptide Enrichment by Aliphatic Hydroxy Acid-modified Metal Oxide Chromatography for Nano-LC-MS/MS in Proteomics Applications, Mol. Cell. Proteomics. 6 (2007) 1103–1109. doi:10.1074/mcp.T600060-MCP200.

[173] P.A. Connor, A.J. McQuillan, Phosphate adsorption onto TiO2 from aqueous solutions: an in situ internal reflection infrared spectroscopic study, Langmuir. 15 (1999) 2916–2921. doi:10.1021/la980894p.

[174] N.C. Jeong, J.S. Lee, E.L. Tae, Y.J. Lee, K.B. Yoon, Acidity scale for metal oxides and Sanderson's electronegativities of lanthanide elements, Angew. Chemie - Int. Ed. 47 (2008) 10128–10132. doi:10.1002/anie.200803837.

[175] S.S. Jensen, M.R. Larsen, Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques, Rapid Commun. Mass Spectrom. 21 (2007) 3635–3645. doi:10.1002/rcm.3254.

[176] A. Leitner, M. Sturm, W. Lindner, Tools for analyzing the phosphoproteome and other phosphorylated biomolecules: A review, Anal. Chim. Acta. 703 (2011) 19–30. doi:10.1016/j.aca.2011.07.012.

[177] T.S. Kephart, P.K. Dasgupta, Hot eluent capillary liquid chromatography using zirconia and titania based stationary phases, Anal. Chim. Acta. 414 (2000) 71–78. doi:10.1016/S0003-2670(00)00821-7.

[178] A. Leitner, Phosphopeptide enrichment using metal oxide affinity chromatography, TrAC - Trends Anal. Chem. 29 (2010) 177–185. doi:10.1016/j.trac.2009.08.007.

[179] Y. Ikeguchi, H. Nakamura, Determination of organic phosphates by column-switching high performance anion-exchange chromatography using on-line preconcentration on titania, Anal. Sci. 13 (1997) 479–483. doi:10.2116/analsci.13.479.

[180] I. Kuroda, Y. Shintani, M. Motokawa, S. Abe, M. Furuno, Phosphopeptide-selective

column-switching RP-HPLC with a titania precolumn., Anal. Sci. 20 (2004) 1313–1319. doi:10.2116/analsci.20.1313.

[181] Y. Sekiguchi, N. Mitsuhashi, T. Kokaji, H. Miyakoda, T. Mimura, Development of a comprehensive analytical method for phosphate metabolites in plants by ion chromatography coupled with tandem mass spectrometry, J. Chromatogr. A. 1085 (2005) 131–136. doi:10.1016/j.chroma.2005.01.098.

2. Publications

2.1 Chromatographic selectivity for broad non-targeted metabolome coverage

In this study, a range of commercially-available RPLC columns suitable for LC-MS based metabolome assessment were taken with an aim of carrying out a practical assessment of chromatographic performance and also the range and repeatability of metabolome coverage possible using a uniform LC-TOFMS platform. Seven different analytical columns packed with 3 µm and sub-2 µm fully-wettable RPLC particles were investigated to determine quality of chromatographic separation, inter-sample molecular feature alignment, metabolome coverage, and overall suitability for various metabolomics workflows. For each column, a wide range of relevant central carbon metabolism standards as well as yeast samples were used to ascertain baseline chromatographic performance and the suitability of columns for realistic metabolomics workflows. A special focus in this study concerned column performance achieved for polar metabolites. As part of ongoing efforts to move towards validation in non-targeted metabolomics approaches, this work aims to provide an informed overview on the role of RPLC column selection upon the quality of LC-based metabolomics workflows using cell extracts relevant to fermentation biotechnology as a suitable example.

Le Si-Hung
Tim J. Causon
Stephan Hann

Division of Analytical Chemistry, Department of Chemistry, University of Natural Resources and Life Sciences (BOKU Vienna), Vienna, Austria

Received April 9, 2017
Revised June 20, 2017
Accepted June 24, 2017

Research Article

Comparison of fully wettable RPLC stationary phases for LC-MS-based cellular metabolomics

Reversed-phase LC combined with high-resolution mass spectrometry (HRMS) is one of the most popular methods for cellular metabolomics studies. Due to the difficulties in analyzing a wide range of polarities encountered in the metabolome, 100%-wettable reversed-phase materials are frequently used to maximize metabolome coverage within a single analysis. Packed with silica-based sub-3 μm diameter particles, these columns allow high separation efficiency and offer a reasonable compromise for metabolome coverage within a single analysis. While direct performance comparison can be made using classical chromatographic characterization approaches, a comprehensive assessment of the column's performance for cellular metabolomics requires use of a full LC-HRMS workflow in order to reflect realistic study conditions used for cellular metabolomics. In this study, a comparison of several reversed-phase LC columns for metabolome analysis using such a dedicated workflow is presented. All columns were tested under the same analytical conditions on an LC-TOF-MS platform using a variety of authentic metabolite standards and biotechnologically relevant yeast cell extracts. Data on total workflow performance including retention behavior, peak capacity, coverage, and molecular feature extraction repeatability from these columns are presented with consideration for both non-targeted screening and differential metabolomics workflows using authentic standards and *Pichia pastoris* cell extract samples.

Keywords:
Biotechnology / Column comparison / Fully wettable column / HPLC / Metabolite / Pichia pastoris / Reversed-phase / TOFMS / Yeast

DOI 10.1002/elps.201700157

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Metabolomics studies are focused on the analysis of low molecular weight intermediates and products of biological processes in a wide range of samples including tissue, biofluids, or cell extracts [1–3]. RPLC remains the most popular mode of HPLC for practical broad scale metabolome coverage as the retention characteristics of these materials are well-suited to a range of moderately polar and nonpolar metabolites. Moreover, advances in column chemistry leading to the development of fully wettable phases that can be routinely used with 100% aqueous mobile-phase conditions without risk of phase collapse [4], have further extended the effective polarity range available in a single analysis [5]. There are several ways to make silica-based columns fully wettable such as having more available space between the alkyl chains that allows the chains to bend without forming a matted surface [6], or using polar-embedded groups [7]. However, due to the difference of packing material, packing quality, and the particle size, each column has its own characteristics and differences in chromatographic separation performance [8]. Furthermore, the introduction of sub-2 μm particles and systems able to work at very high pressures (UHPLC) has yielded significant improvements in both the efficiency of high resolution separations and decrease of analysis time for the analysis of complex biological matrices [9]. Therefore, choosing the right columns can be seen as a first analytical consideration in method development for LC-based metabolomics.

Correspondence: Dr. Stephan Hann, Division of Analytical Chemistry, Department of Chemistry, University of Natural Resources and Life Sciences (BOKU Vienna), Muthgasse 18, Vienna 1190, Austria
E-mail: Stephan.Hann@boku.ac.at

Abbreviations: BRE, batch recursive feature extraction; **BTF**, batch-targeted feature extraction; **HRMS**, high-resolution mass spectrometry; **HILIC**, hydrophilic interaction liquid chromatography; **MFE**, batch molecular feature extraction

Several previous studies have focused on metabolome coverage, peak capacity, and overall performance of different columns suitable for LC-based metabolomics applications [10]. For example, in the study of Wernisch et al., the comparison focused on investigation of metabolome coverage, retention time profile, peak shape, and selectivity from different hydrophilic interaction liquid chromatography (HILIC) columns [11]. Walles et al. also compared separation efficiency and precision of two metabolites: norverapamil and o-demethylverapamil in human-liver microsomal incubates analysis on two sub-2 μm particle columns, with results indicating that there was no substantial improvement when changing from 1.8 μm to 1.7 μm particle size [12]. Analysis results of reduced glutathione, glutathione disulfide, and ophthalmic acid analysis by Lee-Sun et al. indicated the performance of Acquity HSS T3 1.8 μm was better than Acquity UPLC BEH C18 1.7 μm columns [13]. Bajad et al. calculated a score for different columns based on sensitivity, peak sharpness, peak symmetry, and retention time [14]. In a similar study, Contrepois et al. applied a score system to classify five different LC columns into three groups: good, acceptable, and unacceptable in which a good score was characterized by a good retention time to avoid ion suppression in the void volume zone, a narrow elution profile to provide optimal sensitivity and simplify accurate peak integration, and an intense MS signal to be accurately extracted, aligned, quantified, and identified [15]. A critical point of undertaking such comparisons for metabolomics is that assessment using such workflows will only be valid as long as all the samples are studied under the same conditions using a well-controlled method and the change in the signal should be related to the concentration of the components [16, 17]. Furthermore, for method quality control purposes, the concept of using a "biological QC" sample made from pooling aliquots of all samples [18,19] and a QC sample consisting of the blank matrix spiked with known concentrations of various analytes are critical for ensuring data quality over long measurement sequences. QC samples are analyzed during the run in order to confirm repeatability of retention time, peak area, and mass accuracy in samples throughout the measurement sequence. The FDA recommends for single analyte tests that tolerance limits are set such that the measured response detected in two-thirds of QC samples is within 15% of the QC mean, except for compounds with concentrations at or near the LOQ [20]. However, Zelena et al. noted an RSD of <20% for the peak area of QC samples during analysis time was acceptable [21], while Shama Naz et al. accepted an RSD of <30% in nontargeted metabolomics analysis [16].

In addition to the role of RPLC column quality, the ability to provide suitable metabolome assessment depends on the quality and repeatability of high-resolution mass spectra for molecular features that can be obtained. For example, Godzien et al. reported a substantial reduction from approximately 30 000 found chromatographic features in a first step extraction down to 20 confirmed compounds following exhaustive filtering and metabolite annotation. Therefore, each step of assessment must be well-understood and data processed carefully to avoid significantly affecting the data interpretation due to the enhancement and accumulation of error from the beginning to later stages [22]. To this end, comparison of column performance for LC-based metabolomics necessitates a holistic study encompassing the influence of both analytical aspects, with a focus on chromatographic performance and analytical method parameters (e.g. precision under repeatability conditions) with relation back to the original objective of the analytical task (i.e., maximization of metabolome coverage).

Thus in the present study, a range of commercially available RPLC columns suitable for LC-MS based metabolome assessment were taken with an aim of carrying out a practical assessment of chromatographic performance and also the range and repeatability of metabolome coverage possible using a uniform LC-TOFMS platform. Seven different analytical columns packed with 3 μm and sub-2 μm fully wettable RPLC particles were investigated to determine quality of chromatographic separation, intersample molecular feature alignment, metabolome coverage, and overall suitability for various metabolomics workflows. For each column, a wide range of relevant central carbon metabolism standards as well as yeast samples were used to ascertain baseline chromatographic performance and the suitability of columns for realistic metabolomics workflows. A special focus in this study concerned column performance achieved for polar metabolites. As part of ongoing efforts to move toward validation in nontargeted metabolomics approaches [23], this work aims to provide an informed overview on the role of RPLC column selection upon the quality of LC-based metabolomics workflows using cell extracts relevant to fermentation biotechnology as a suitable example.

2 Experimental

2.1 Chemicals

All solvents used for mobile phase preparation such as H_2O, methanol, formic acid were LC-MS grade and purchased from Sigma Aldrich (Vienna, Austria). All standard substances were purchased from Sigma Aldrich and Merck. First, single standard solutions were prepared in LC-MS grade water containing suitable additives (0.1 M HCl, 0.1 M NaOH) when necessary. The full list of metabolites used in this study is provided in the Supporting Information Table 1.

2.2 Sample preparation

Three types of test sample were prepared for LC-TOFMS experiments, i.e. a mixture of 90 analytical grade authentic metabolite standards, *Pichia pastoris* cell extract samples, and *P. pastoris* cell extract samples spiked with the metabolite mixture.

Equimolar mixtures of 90 metabolites were prepared from single standard solutions with a final concentration

of 50 μM in 0.1% v/v formic acid. This global QC mixture is used in our laboratory to support validation of targeted metabolomics workflows using HILIC, reversed-phase, and GC-based protocols. All equimolar metabolite standard mixtures were immediately centrifuged, evaporated under reduced pressure using a GeneVac EZ-2 solvent evaporation system and stored at −20°C. For each column assessment, the 50 μM metabolite standards mixtures was diluted down to 1 μM using 0.1% v/v formic acid and measured immediately.

Ethanolic cell extracts from *P. pastoris* were obtained from ISOtopic Solutions (Vienna, Austria). The calculated amount of *P. pastoris* cell dry weight per injection was 37 μg. Two additional dilutions of the *P. pastoris* samples were also prepared in 0.1% v/v formic acid corresponding to 7.4 μg and 1.85 μg of biomass per injection. Spiked cell extract samples were prepared by adding an aliquot from the 90 metabolite standard mixture to yeast samples. The final concentration of spiked metabolites was 1 μM with 37 μg of yeast biomass per injection.

2.3 Instrumentation

An Agilent model 1290 Infinity II HPLC system and analytical columns from several manufacturers (Agilent, Fortis Technologies, Macherey-Nagel, Phenomenex, Sigma-Aldrich, and Waters) were utilized in this study. All columns were 150 mm in length with an internal diameter of 2.1 mm. Three columns containing 3 μm particles: Atlantis T3 C18 (**ATL**) (Waters, Milford, USA), Discovery HSF5 (**DIS**) (Sigma-Aldrich), and Fortis H_2O (**FOR**) (Fortis Technologies). The particle size for Acquity HSS T3 (**ACQ**) (Waters), Zorbax SB-AQ (**ZOR**) (Agilent), and Nucleodur C18 Gravity-SB (**NUC**) (Macherey-Nagel) is 1.8 μm, while the Luna Omega (**LUN**) (Phenomenex) contains 1.6 μm particles.

For MS detection an Agilent model 6230 TOFMS from Agilent Technologies equipped with a Dual Jetstream ESI interface was used.

2.4 LC-TOFMS method

In order to allow a fair comparison of all columns over different measurement days and sequences, samples were always prepared immediately prior to measurement and stored at 4°C during the entire measurement sequence. An inline 0.2 μm filter from IDEX Health & Science (USA) was used to protect the columns.

The HPLC method used in this study was adapted from previous work [24]. Briefly, mobile phase A consisted of water with 0.1% v/v formic acid, mobile phase B contained 100% methanol. Initial gradients conditions (100% A) were kept for 2 min, then increased from 0% to 40% B in the next 8 min, and was held for 2 min before starting the cleaning step. Following a cleaning step and suitable reequilibration time (5.9 min), the total analysis time was 20 min per injection. A constant flow rate of 250 μL/min was used, an injection volume of 5 μL, and the column temperature was 45°C.

For TOFMS detection, mass spectra between 50 and 1700 m/z were recorded in positive polarity mode with following settings: drying gas temperature 120°C, drying gas flow 10 L/min, nebulizer pressure 40 psig, sheath gas temperature 350°C, sheath gas flow 12 L/min, capillary voltage 3500 V, and fragmentor voltage 120 V. The 2 GHz extended dynamic range detection mode with an acquisition rate of three TOF spectra/s was used for all measurements.

2.5 Quality control

In order to ensure that batch-to-batch and day-to-day variation of TOFMS performance were not biasing comparisons, the metabolite mixture was measured using the ATL column at the beginning and the end of each standards-mixture batch, and spiked yeast samples were measured during each batch of *P. pastoris* measurements. Measurement of QC samples was needed to ascertain that the repeatability of retention time, peak area, and extracted accurate mass of some stable metabolite compounds confirmed the compliance of results obtained on that measurement day [19]. The QC data were evaluated with a maximum allowance of ±3% for retention time, ±2 ppm for extracted accurate mass, and ±15% of QC mean for peak area to reflect current FDA guidelines [20]. Supporting Information Fig. 1 demonstrates the peak area precision obtained over a period of 7 days.

2.6 Chromatographic performance

Chromatographic peak capacity was considered for both the analysis of authentic standards and yeast extracts using peak width and retention time data of compounds that could be reliably determined with each column used in the study. Peak capacity (n_c) was calculated according to Snyder [25]:

$$n_c = 1 + \frac{t_g}{FWHM} \quad (1)$$

Where t_g is the gradient time and FWHM is the average full peak width at half maximum.

2.7 Data processing and evaluation

In this study, two main metabolomics datasets were collected for each column. The comparisons made from these datasets are further divided considering requirements for different analytical strategies employed for metabolome assessment. All data evaluation was performed using Agilent PCDL Manager B.07.00, Agilent MassHunter Profinder B.08.00.

Firstly, a dataset based solely on measurements of authentic metabolite standards in a simple matrix (0.1% v/v formic acid) was treated using a batch-targeted feature extraction (BTF) workflow. The BTF workflow includes only one stage of feature extraction and is based on primary (monoisotopic) ions, adducts, and isotopologue patterns derived from the input sum formula. This set of data was used to reveal

basic information and baseline performance of each column for a wide range of metabolites relevant to the application at hand. For development of targeted and nontarget screening metabolomics workflows, PCDL Manager was used to create a column-specific in-house compound database based on the input chemical formulae and experimentally measured retention times of the authentic metabolite standards for each column in the positive mode. These databases could then be applied using a BTF in Profinder for extracting each compound's ion chromatogram (i.e. a sum of extracted signals representing isotopologues, dimers, and adducts of the target compound). The extracted compound chromatograms obtained for the targeted metabolites were considered for determination of the precision under repeatability conditions ($n = 6$ replicate injections per column) for retention time, peak area, mass accuracy, and compound target score. The extraction width was ±20 ppm, while matching tolerance for mass accuracy (bias) and retention time were ±10 ppm and ±0.5 min, respectively. Finally, yeast samples spiked with authentic metabolites were also analyzed and evaluated under the same conditions to assess the performance of columns with a real matrix in the context of nontarget screening using the developed column-specific in-house compound databases. Molecular features were further filtered (target score ≥75, mass precision ±2 ppm) to ensure that spiked compounds were correctly annotated. The quality of batch molecular feature extraction (MFE) is represented by the so-called MFE and target scores, which are a weighted averages reflecting how well the software-extracted molecular feature matches typical (or target) monoisotopic mass, isotopologue pattern, retention time, and matching of a target chemical formula to MS data when available.

The second dataset collected was based on the analysis of ethanolic yeast extracts prepared in different dilutions in order to estimate coverage, and also to gauge the effect of a complex sample matrix upon performance of the column and potential use in a nontargeted differential analysis workflow. Results for this set were extracted using a batch recursive feature extraction (BRE) workflow in MassHunter Profinder. The BRE includes two stages: The MFE and Batch Find by Ion feature extraction (FbI). MFE removes unwanted information such as background noise, and summarizes all extracted features in each sample considering accuracy of mass measurements, the grouping of ions related to the charge-state envelope, isotopologue pattern, and the presence of adducts or dimers [19]. Then, these molecular features are aligned across all of the selected samples using mass and retention time characteristics. The total compound abundance is calculated as a sum of areas stemming from all relevant single ion features. Next, FbI uses the median values derived from the MFE process to perform a targeted extraction to improve the reliability of finding and reporting features from complex datasets used for differential analysis. In the BRE, samples were grouped, aligned, and checked for extraction quality based on narrow retention time and mass-to-charge windows (±0.15 min and 20 ppm ±2.0 mDa, respectively). Extracted ion chromatograms with symmetric extraction windows (m/z ± 10 ppm and expected retention time ±0.5 min) were then extracted for all compounds. In order to be considered as a compound of interest for further evaluation in a differential analysis workflow, appearance in all of the yeast samples with an MFE score and target score greater than 75 were required.

Finally, separate comparisons using solvent blanks and three different biomass concentrations were used to make a general evaluation of these aspects for performance and coverage for each column.

3 Results

3.1 Initial column screening and study design

Isocratic experiments with UV/VIS detection were conducted to assess the primary suitability of columns considered of interest to LC-MS based metabolome assessment (i.e. retention capacity and number of theoretical plates). The mobile phase flow-rate chosen for the remainder of study was found to represent a good compromise for throughput and sensitivity of ESI-MS detection for 2.1 mm id columns (250 µL/min). While UHPLC conditions for columns packed with sub-2 µm particles could be employed to yield faster analysis with the same quality of separation (i.e. peak capacity, resolution), this parameter was kept constant to allow fair comparison for practical metabolomics workflow using the same LC-TOFMS platform including the electrospray conditions. Based on the performance (i.e. retention capacity, peak shape, efficiency) during initial screening (data not shown), seven columns packed with fully wettable silica-based particles were selected for further testing using the LC-TOFMS platform.

Using a standardized LC-TOFMS method, columns were tested with a mixture containing 90 relevant intracellular metabolites representing the central carbon, energy, and redox metabolism. As already mentioned in the experimental section, this mixture represents an enhanced QC standard that is used to check all LC-MS and GC-MS based separations. Evidently, several compounds in this mixture were not amenable to reversed-phase liquid chromatography, and did not show significant retention on these columns. Measurements of this mixture (six technical replicates per column) were used to assess the separation quality, repeatability, and estimate the metabolome coverage obtainable using each column. All separation conditions including column temperature, flow rate, and gradient profile were kept constant as all columns were characterized by the same dimensions, while void volumes were estimated to be 1.2 min ± 15% for all columns (see Supporting Information Table 1). To further assess performance and repeatability for the analysis of biotechnologically relevant cellular extracts, six technical replicates for the analysis of yeast samples with 37 µg biomass per injection, three each for 7.4 and 1.85 µg biomass, and three for spiked yeast samples were collected. In order to allow a fair comparison of all columns using chromatographic and MS-based parameters, 12 representative metabolites, all of which

showed an acceptable chromatographic retention time and a high score (MFE score ≥75) regarding molecular feature extraction, were selected and identified in both the metabolite standard mixture and spiked yeast samples for all columns assessed in this study.

3.2 Chromatographic performance

3.2.1 Retention time distribution

This evaluation considered the number of compounds from a set of authentic metabolites that could be reliably retained, extracted, and identified using the columns assessed in this study. In case of the metabolite standards mixture, of the 90 metabolites contained in this mixture, approximately 40% of metabolites eluted within less than 2.5 min for all columns assessed that is close to the void volume of approximately 1.2 min for all investigated columns. Within this retention range are several polar metabolite classes including some amino acids, sugar phosphates, and other sugar-related compounds. The fraction of sample eluting within or close to the column void volume is prone to suffer from ion suppression effects, or, in some cases, ion enhancement in ESI-MS due to coelution of matrix components [26]. In this study, many of the early-eluting compounds were characterized by broader peak widths and lower software-based compound extraction reliability due to retention time misalignment.

For the majority of columns in this study, more than 50% of the total number of software-extracted compounds was retained for less than 3 min except on the ATL and ACQ columns. The ATL was characterized by the lowest fraction of metabolites eluting in this region (<40%). Overall, the retention behavior of the columns studied was found to be very similar with almost 80% of total compounds extracted eluting within 6 min in all cases. All retention data for software-extracted compounds are provided in the Supporting Information Table 1. Furthermore, based on the number of compounds eluted in each retention time window, a compound elution distribution profile for each column could be created (Fig. 1). Ideally, assuming that each 1 min window should contain 8.33% of the total number of compounds, an elution distribution factor was calculated based on the summed square of residuals between each retention distribution profile and the ideal profile. Thus, the smaller the factor is, the better the performance. Using this assessment, the ATL (75.6) and DIS (76.2) columns were found to be the closet to the ideal elution profile compared to other columns, e.g. NUC (80.3), ACQ (81.7), FOR (95.2), LUN (92.1), and ZOR (108.2).

3.2.2 Peak capacity

Among the 12 compounds reliably found on each column, only compounds with a retention time greater than 3.5 min were chosen to calculate peak capacity (thymine,

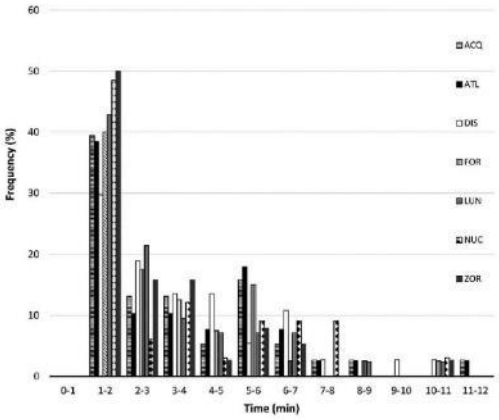

Figure 1. Frequency distribution reflecting the retention profile of metabolites reliably software-extracted from 90 metabolite mixture for the seven columns assessed in this study.

phenylalanine, uridine, inosine, guanosine, and S-adenosylhomocysteine)cc. In most cases, column peak capacity showed no significant difference between standard-mixture and spiked samples except for LUN (Table 1). ACQ showed the highest peak capacity in both cases (166 and 173), while ZOR and DIS had the lowest peak capacity. It was interesting to note that there was also no apparent strong relationship between particle size and the calculated peak capacity. Finally, based on these results, we suggest that comparing peak capacity values as a standalone metric using such a wide range of metabolite classes should be interpreted carefully due to peak shape variations observed when using a complex mixture of standards, in comparison to a real sample matrix.

3.2.3 Retention time repeatability

Considering the metabolite standard mixture, the average retention time RSD was 0.34% across all columns and compounds, and the maximum retention time RSD was 1.23% (for L-cystathionine in case of NUC). NUC and ZOR were characterized by a slightly higher average retention time

Table 1. Chromatographic performance parameters assessed for columns used in this study

	ACQ	ATL	DIS	FOR	NUC	LUN	ZOR
n_c (standards-mixture)	166	147	110	142	150	132	114
n_c (spiked sample)	173	156	107	156	156	163	130
Positive mode coverage	36	37	35	38	30	40	37
Recovered in spiked samples	34	34	30	36	27	37	32

Metabolome coverage was calculated by considering the number of reliably extracted compounds from the 90 metabolite mixture. The fraction of these metabolites correctly annotated in the spiked samples is also given.

RSD than other columns (0.85 and 0.81%, respectively). With the analytical platform analyzing only metabolite standards undisturbed over a single batch (several hours of continuous measurement), the retention time stability was found to be very good for all columns used in this study. It is noteworthy that column equilibration with the ethanolic extracts was not critical concerning retention time repeatability and that the injection of three samples for system equilibration prior to each sequence was sufficient. This is due to the fact that the matrix content of ethanolic cell extracts is generally much lower as the matrix content of plasma or other biofluids.

3.3 Assessment of columns with full data processing workflows

3.3.1 Positive mode coverage of metabolite standards mixture

The targeted metabolomics workflow revealed the number of metabolites in the standard mixture that could be reliably extracted for each column in the positive ionization mode. The range of compound coverage obtained was 30–40 compounds from the 90 metabolite standards injected. However, of the 90 metabolites contained in this sample, assessment of chemical structures and confirmation with negative mode ionization measurements indicated that approximately 50% of these metabolites cannot reliably form a protonated molecular ion under these measurement conditions employed. Moreover, some metabolites (e.g. sugar phosphates) were observed to suffer from insufficient retention and/or poor resolution on all columns (see Supporting Information). Among the sub-2 μm particle columns, LUN had the highest number of the metabolites that could be reliably extracted by software (40), while the lowest number was software-extracted for the NUC (30). Among the 3 μm particle columns, FOR showed the highest coverage of the test compounds (38). Interestingly, the ATL (37) exhibited almost the same degree of compound coverage as ACQ (36). In general, it can be said that little difference in the coverage using the same HPLC gradient was observed for the different columns and also with respect to the particle size. In terms of critical pair (isobaric) separations, isomeric sugar phosphates could not be reliably separated on any of the columns considered, while only ZOR could not separate leucine and isoleucine under the gradient conditions employed in this study. This might be in part due to the lower retention capacity of this column at the moderately high column temperature employed in this study (45°C). Conversely, the pentafluorophenyl phase (DIS) exhibits extremely strong retention of these compounds with a peak-to-peak resolution of approximately 3.

3.3.2 Mass precision

In term of MS quality parameters, all columns tested exhibited good repeatability with mass precision of below 2.0 ppm, which was calculated by standard deviation divided by the average monoisotopic mass of n replicates for the 12 metabolites that could be software-extracted reliably on all columns. Average mass precision for all columns was 0.27 ppm, and the worst precision was 0.88 ppm for S-adenosyl-homocysteine in the case of NUC. Among the seven columns, NUC showed a slightly worse than average mass precision (0.52 ppm), LUN had the best average mass precision (0.21 ppm), but it can be concluded that the choice of column did not apparently influence the accuracy of TOFMS detection.

3.3.3 Peak area repeatability

Considering the set of 12 representative metabolites from the standard mixture that could be reliable analyzed using all seven columns, the average peak area RSD for all columns was less than 5.0% except for NUC (6.2%). Some particular compounds stood out as outliers in this dataset. For example, poor repeatability was observed for thymine using NUC (12.7%), L-cystathionine and cystine for FOR (8.3%), S-adenosyl-homocysteine using ZOR (8.2%), respectively. It is worth noting here that many sulfur-containing metabolites are known to be prone to degradation and are usually assessed with deriviatization-based workflows [27].

3.3.4 Quality control

For quality control of the standards mixture measurements, data of several compounds measured with the ATL column alongside each batch were assessed. All tested metabolites were characterized by good batch-to-batch repeatability with maximum RT RSD of 2.2%, a maximum extracted accurate mass RSD of 1.6 ppm in the case of 5'UMP, and a maximum RT bias of 1.2% (leucine). In terms of peak area, there was significant observable changes over the measurement days, but this variation was nonlinear (i.e. not all the metabolites in the sample followed the same trend). Examining the detected peak areas of several stable metabolites revealed values within 20% RSD (see Supporting Information Fig. 1) and therefore the observed variation was considered to be acceptable for comparing columns on the same platform.

3.4 Performance test for a nontarget screening approach

The same six targeted compounds that had retention time greater than 3.5 min were used for peak capacity determination in spiked samples. Peak capacity values were observed to be marginally improved over the standards-mixture in a number of the studied columns (Table 1). This may be in part caused by the higher concentration of the metabolites, mild changes in chromatographic processes due to the sample matrix, and/or ionization enhancement effects arising from the sample matrix.

However, lower repeatability of software-based compound extraction was observed for the spiked samples compared to the metabolite standards alone. The maximum retention time RSD was 1.5% for 3′AMP using ACQ. In terms of extracted mass, the worst extracted mass precision was 0.96 ppm in case of S-adenosyl-homocysteine for NUC. In term of peak area, ACQ exhibited the lowest average peak area RSD of 2.1% calculated for the 12 selected compounds. For the other columns, the average peak area RSD ranged from 4.6% to 10.9%. Poor repeatability was observed for some particular compounds such as inosine and guanosine for ATL (8.6%, 9.3%), S-adenosyl-homocysteine for DIS, LUN, NUC (13.5, 25.5, and 20.1%), guanosine for NUC and ZOR (37.5%, 52.8%), and thymine for ZOR (12.8%). All of these cases were manually checked with software to ensure that no integration errors had occurred. In some cases, the width of the mass extraction window was found to be a major contribution to this variation (i.e. employing a wider mass extraction window reduced the RSD substantially). The increased complexity of the sample matrix can lead to suppression of reference ion signal and/or interferences with the target compound ions.

Furthermore, the results from analysis of spiked samples revealed a shifting of retention times away from those observed in the standard mixture recorded in separate batches (Fig. 2). The minimum and maximum average RT biases were 0.3% (ACQ) and 2.9% (DIS), respectively. A maximum RT bias of 6.98% for adenine using DIS was observed, which, although very large, is not unexpected for a weakly retained compound. Retention time matching (bias against column-specific in-house compound databases) was variable for the different columns with values reaching >5% in some cases mostly within the elution time window of 4–6 min. This may indicate that, in some cases, the columns should be given more time to reequilibrate or this could be due to the retention of methanol under initial 100% aqueous conditions disturbing compounds eluting near the solvent front and thus increasing the RT bias in the middle of the gradient. This result again highlights that the selectivity of the high-resolution mass spectrum and retention time is often not sufficient for metabolite identity confirmation using databases based on authentic standards in some cases, and MS/MS or a method with complementary selectivity may be required [28]. Moreover, the building of column-specific in-house compound databases is better carried out by using measurements of spiked samples, rather than authentic standards alone. Nevertheless, based on the number of standard compounds that could be found using the column-specific in-house databases following spiking, the software-based compound recovery in a nontarget screening workflow could be calculated, with results showing that approximately 90% of metabolites could be reliably identified in spiked yeast extracts for all tested columns (Table 1).

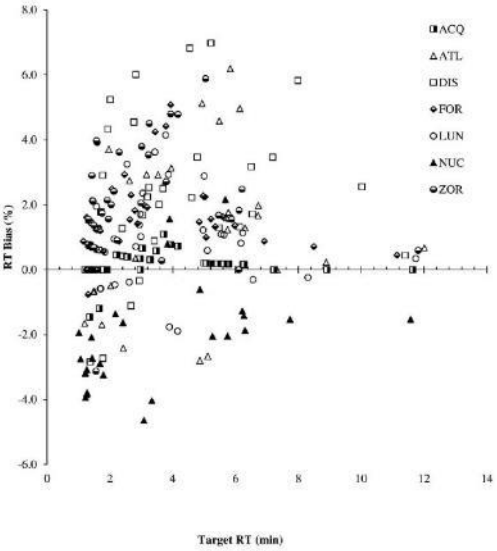

Figure 2. RT bias versus retention time of seven columns based on database value against the measured value from spiked yeast samples.

3.5 Practical implications of column choice for nontargeted metabolome assessment

Nontargeted differential analysis relies on a relative quantification of software-extracted, aligned, and averaged molecular features found across multiple samples. For such assessments, data need to be carefully checked manually, and statistically processed to avoid model overfitting and/or result distortion due to the risks associated with small numbers of observations with a much larger number of variables [22].

An initial assessment with moderate filtering was applied to generate a baseline dataset for completely nontargeted molecular feature extraction for metabolome assessment. This assessment considered only the six technical replicates of yeast sample (37 μg of extracted biomass per injection) without retention time restrictions. Confirmed molecular features had to be found in five out of six replicates, while MFE and target scores had to both be greater than 75. Using this workflow, the number of molecular features that could be found in each column varied from 720 to 966 features. LUN showed the highest number of molecular features (966), while ZOR had the lowest number of features (720). The number of molecular features that could be extracted using the other columns were 826 for ACQ, 840 for ATL, 878 for DIS, 812 for FOR, and 820 for NUC, indicating that no clear correlation between particle size and number of molecular features extracted was discernible. While the number of molecular features determined using each column appears promising for nontargeted LC-MS metabolomics workflows, a large percentage of the molecular features eluted between

1.2 and 1.5 min whereby the elution of matrix components can severely impede relative quantification.

To investigate additional factors including variation in injected biomass with consideration of ion suppression effects for early-eluting compounds and potential background compound arising from HPLC-grade solvents, a second assessment using three different biomass concentrations and corresponding blanks was carried out. Decreasing the amount of injected (extracted) biomass reduces the overall matrix load on the HPLC column and also the possibility of medium to long-term source contamination in the MS. However, these improvements come at the clear cost of losing metabolite signals as ion counting statistics and absolute sensitivity limitations of the analytical platform are reached. Furthermore, principally superfluous/unwanted molecular features found to be present in blank samples can also be expected to be found in the yeast extract samples regardless of biomass concentration. However, ion suppression arising from increasing the injected biomass (viz. sample matrix) was observed to reduce the number of background compounds reliably found for all columns studied. With consideration of a reliable retention time window for relative quantification tasks (i.e. ≥3.5 min), the number of molecular features that could be found in the highest biomass concentration was considered as the reference number of molecular features in the *Pichia* sample. DIS and LUN had the highest number of total features with 624 and 576, while the number of molecular features of other columns were 455 for ACQ, 400 for NUC, 384 for ZOR, 295 for FOR and 343 for ATL, respectively. Reducing the biomass sample caused an expected decrease in number of molecular features that could be reliably analyzed. When one fifth of the biomass was injected, approximately 50% of total molecular features were lost, and with one-twentieth injected, approximately 20% of molecular features were remaining for all columns (Fig. 3).

Blank samples also were treated with BRE separately using the same software settings and filtering. DIS accounted for highest number of molecular features from the blank samples (81), the runner up was ACQ (35), while less than 20 molecular features were determined for the other columns. The results seem to indicate that the PFP group (DIS) may lead to increased retention of background compounds, thereby increasing the total number of molecular features that could be consistently found.

In summary, the chosen biomass concentration employed for a particular LC-MS metabolomics method depends finally on a combination of the influence of sample matrix on chromatographic and MS aspects of the total measurement. Moreover, the total number of samples to be analyzed and the ruggedness of the method are additional concerns for which due consideration must be given during method development. It can be concluded that the analysis of solvent or procedural blanks was found to be critical for all columns used in this study, even though the total number of molecular features arising from the solvent was very low in many cases.

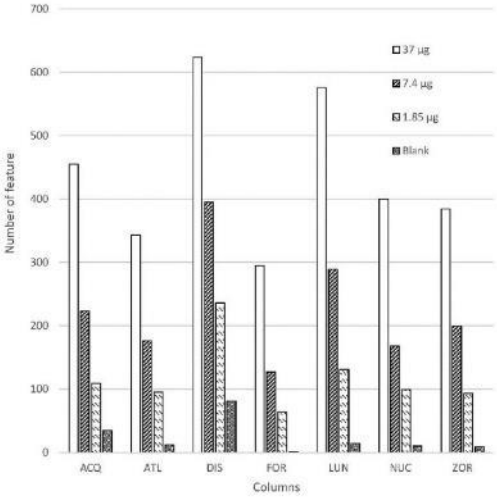

Figure 3. Number of molecular features determined for different total amounts of extracted yeast biomass (three technical replicates) and comparison to solvent blanks (three technical replicates).

4 Conclusions

For the LC-TOFMS platform employed in this study, results indicated that the 3 µm columns had comparable performance to the sub-2 µm columns when considering the entire analytical platform. Interestingly, particle size and chromatographic peak capacity were not found to exhibit a strong correlation with estimates of metabolome coverage using this common analytical platform and method. However, it should also be noted that the performance of all individual columns examined in this study could be further optimized by adjustment of gradient, flow-rate, temperature, and reequilibration time. Furthermore, the long-term stability of columns for real samples was not assessed in this work, which is an additional consideration for future studies. Overall, DIS and LUN were the two columns that showed the best performance, while ACQ and ATL also provided very good characteristics. DIS performed very well in this study primarily due to its different retention characteristics, which could be very beneficial for the analysis of polar metabolites.

This work also highlights that building of compound libraries for metabolite identification based on measurement of standards alone is prone to errors with retention time matching, and we suggest that such libraries should be built on spiked samples wherever possible. Furthermore, some additional challenges still remain for LC-HRMS in cellular metabolomics such as the difficulty of retention time matching when beginning measurements with 100% aqueous conditions, lack of retention of important classes of metabolites, and challenges for validating differential analysis workflows. Based on the results of this study and the general

shortcomings encountered with all columns, we suggest that improving retention capacity for polar metabolite classes and separation selectivity are more critical for improving coverage, repeatability, and method ruggedness rather than selection of the optimum HPLC column. Finally, it can be concluded that all columns tested in this study were found to be suitable for LC-MS metabolomics workflows as described in this study. However, based on separation considerations, we suggest that PFP phases present an interesting choice for further studies in this area, particularly as part of multidimensional methods using chromatography or ion mobility as a second separation step in order to improve overall metabolome coverage.

The authors gratefully acknowledge Hedda Drexler (Division of Analytical Chemistry, BOKU Vienna) for support with LC-TOFMS measurements. This work was funded by the ASEA-UNINET on behalf of the Austrian Federal Ministry of Science. Vienna Business Agency and EQ BOKU VIBT GmbH are acknowledged for providing mass spectrometry instrumentation. The authors thank Dr. Gerrit Hermann (ISOtopic Solutions) for providing the yeast samples.

The authors have declared no conflict of interest.

5 References

[1] Nicholson, J. K., Connelly, J., Lindon, J. C., Holmes, E., Nat. Rev. Drug Discov. 2002, 1, 153–161.

[2] Fiehn, O., Plant Mol. Biol. 2002, 48, 155–171.

[3] Schneewind, O., Missiakas, D. M., Philos. Trans. R Soc. Lond. B. Biol. Sci. 2012, 367, 1123–1139.

[4] Pettersson, S. W., Persson, B. S., Nyström, M., J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 2004, 803, 159–165.

[5] Ortmayr, K., Causon, T. J., Hann, S., Koellensperger, G., TrAC Trends Anal. Chem. 2016, 82, 358–366.

[6] Herbert, N. R., Ebdon, D., The HPLC analysis of polar analytes with aqueous mobile phases. The 25th International Symposium on High Performance Liquid Chromatography and Related Techniques (HPLC), June 17-22, 2011, Maastricht, The Netherlands. 2001.

[7] Majors, R. E., Przybyciel, M., LCGC Eur. 2002, 15, 652–657.

[8] Plenis, A., Oledzka, I., Baczek, T., J. Pharm. Biomed. Anal. 2013, 78–79, 161–169.

[9] Grata, E., Guillarme, D., Glauser, G., Boccard, J., Carrupt, P.-A., Veuthey, J.-L., Rudaz, S., Wolfender, J.-L., J. Chromatogr. A 2009, 1216, 5660–5668.

[10] Cajka, T., Fiehn, O., Anal. Chem. 2016, 88, 524–545.

[11] Wernisch, S., Pennathur, S., Anal. Bioanal. Chem. 2016, 408, 6079–6091.

[12] Walles, M., Gauvin, C., Morin, P.-E., Panetta, R., Ducharme, J., J. Sep. Sci. 2007, 30, 1191–1199.

[13] New, L.-S., Chan, E. C. Y., J. Chromatogr. Sci. 2008, 46, 209–214.

[14] Bajad, S. U., Lu, W., Kimball, E. H., Yuan, J., Peterson, C., Rabinowitz, J. D., J. Chromatogr. A 2006, 1125, 76–88.

[15] Contrepois, K., Jiang, L., Snyder, M., Mol. Cell. Proteomics 2015, 14, 1684–1695.

[16] Naz, S., Vallejo, M., García, A., Barbas, C., J. Chromatogr. A 2014, 1353, 99–105.

[17] Gika, H. G., Theodoridis, G. A., Earll, M., Snyder, R. W., Sumner, S. J., Wilson, I. D., Anal. Chem. 2010, 82, 8226–8234.

[18] Sangster, T., Major, H., Plumb, R., Wilson, A. J., Wilson, I. D., Analyst 2006, 131, 1075.

[19] Alonso-Herranz, J. G. V., Barbas, C., Grace, E., Metabolomics 2015, 11, 518–528.

[20] Food and Drug Administration, Guidance for Industry: Bioanalytical method validation., 2013.

[21] Zelena, E., Dunn, W. B., Broadhurst, D. I., Francis-McIntyre, S., Carroll, K. M., Begley, P., O'Hagan, S., Knowles, J. D., Halsall, A., Wilson, I. D., Kell, D. B., Anal. Chem. 2009, 81, 1357–1364.

[22] Godzien, J., Ciborowski, M., Angulo, S., Barbas, C., Electrophoresis 2013, 34, 2812–2826.

[23] Gika, H. G., Wilson, I. D., Theodoridis, G. A., J. Chromatogr. B 2014, 966, 1–6.

[24] Neubauer, S., Haberhauer-Troyer, C., Klavins, K., Russmayer, H., Steiger, M. G., Gasser, B., Sauer, M., Mattanovich, D., Hann, S., Koellensperger, G., J. Sep. Sci. 2012, 35, 3091–3105.

[25] Wang, X., Stoll, D. R., Schellinger, A. P., Carr, P. W., Anal. Chem. 2006, 78, 3406–34016.

[26] Chambers, E., Wagrowski-Diehl, D. M., Lu, Z., Mazzeo, J. R., J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 2007, 852, 22–34.

[27] Hansen, R. E., Winther, J. R., Anal. Biochem. 2009, 394, 147–158.

[28] Kind, T., Fiehn, O., BMC Bioinformatics 2007, 8, 105.

2.2 Selectivity cleanup/enrichment of phosphorylate primary metabolites

In this study, several commercially-available MOAC materials were tested for selective enrichment of intracellular phosphorylated metabolites in yeast cell extract samples. The best material was chosen through a preliminary screening based on phosphorylated metabolites trapping ability, and the removal of unwanted compounds in real samples. The sample cleanup and phosphorylated metabolites enrichment workflow using the best SPE material was then optimized and validated using GC-MS/MS. As part of efforts to overcome challenges in sensitive and quantitative analysis of central carbon metabolites, this work proposes a validated efficiency sample cleanup and selective phosphorylated metabolite enrichment method using ZrO_2-based SPE material.

Contents lists available at ScienceDirect

Talanta

journal homepage: http://ees.elsevier.com

Sensitive quantitative analysis of phosphorylated primary metabolites using selective metal oxide enrichment and GC- and IC- MS/MS

Le Si-Hung, Christina Troyer, Tim Causon, Stephan Hann*

Institute of Analytical Chemistry, Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria

ARTICLE INFO

Keywords
Mass spectrometry
Metabolomics
Quantification
Sample preparation
Solid phase extraction
Sugar phosphate

ABSTRACT

In this study, we present a novel selective cleanup/enrichment method based on metal oxide solid phase extraction combined with quantitative gas chromatography-tandem mass spectrometry and ion exchange chromatography-tandem mass spectrometry for the analysis of phosphorylated metabolites in yeast cell extracts relevant to biotechnological processes. Following screening of several commercially available metal oxide-based enrichment materials, all steps of the enrichment process (loading, washing and elution) were optimized for both the selective enrichment of 12 phosphorylated compounds from the glycolysis and pentose phosphate pathways, and the simultaneous removal of highly abundant matrix components such as organic acids and sugars. The full analytical workflow was then validated to meet the demands of accurate quantification of phosphorylated metabolites in yeast (*Pichia pastoris*) cell extracts using the best performing material and cleanup/enrichment method combined with quantification strategies based on internal standardization with isotopically labeled internal standards and external calibration. A good recovery (>70%) for 5 of the 12 targeted phosphorylated compounds with RSDs of less than 6.0% was obtained while many sugars, organic acids and amino acids were removed (>99% of glucose, and >95% of aspartate, succinate, glutamate, alanine, glycine, serine, threonine, proline, and valine). The use of isotopically labeled internal standards added to the samples prior to SPE, enables accurate quantification of the metabolites as it compensates for errors introduced during sample pretreatment *and* GC-MS or LC-MS analysis. To the best of our knowledge, this is the first time an effective and selective metal oxide-based affinity chromatography cleanup/enrichment method was designed and applied successfully for intracellular phosphorylated metabolites.

1. Introduction

Metabolomics studies focus on the analysis of a wide range of low molecular weight intermediates and products of biological processes in samples including tissues, biofluids or cell extracts [1–3]. Within this diverse range, phosphorylated metabolites such as intermediates of glycolysis and the pentose-phosphate pathway play very important roles in substrate degradation, energy budget, cofactor regeneration, biosynthetic precursor supply as well as regulation of tumor energetic metabolism, and are therefore essential in all life processes. Regarding the challenging quantitative analysis of phosphorylated compounds, several techniques such as capillary electrophoresis (CE)-mass spectrometry (MS), liquid chromatography (LC)-MS and gas chromatography (GC)-MS have been applied in the past [4–6]. However, for separation and quantification of the most important sugar phosphates, GC-MS remains the gold standard for metabolomics [7,8]. GC-MS combines high separation efficiency and sensitivity as well as selective detection and hence covers - applying efficient derivatization strategies - a broad range of polar metabolites such as organic acids, amino acids, sugars, sugar phosphates and phosphorylated compounds within a single chromatographic run. However, due to the low concentrations of phosphorylated metabolites in cellular samples and potential matrix interferences for GC derivatization [9], the quantitative determination of phosphorylated compounds in biological samples remains analytically challenging. Consequently, reduction of sample complexity using selective enrichment methods is a step forward for their quantitative analysis, since not only derivatization in GC-MS is matrix-dependent, but also the electrospray ionization (ESI) process in LC-MS suffers from matrix-related suppression.

Metal oxide-based affinity chromatography (MOAC) using materials such as titanium and zirconium oxides has been reported as a useful strategy for solid phase extraction (SPE)-based enrichment for phosphorylated peptides [10–12] and phosphoproteins [13–15] due to the par-

* Corresponding author. Division of Analytical Chemistry, Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Muthgasse 18, A-1190, Vienna, Austria.
Email address: Stephan.Hann@boku.ac.at (S. Hann)

https://doi.org/10.1016/j.talanta.2019.120147
Received 23 May 2019; Received in revised form 9 July 2019; Accepted 10 July 2019
Available online xxx
0039-9140/

ticular affinity of metal oxide surfaces toward phosphate-group containing compounds. The binding mode of the phosphate anion, mono-substituted phosphates and carboxylic acids to the TiO_2 surface was explained as being of bidentate nature based on the donation of two electron pairs to a metal atom, and the fact that metal oxide can act as Lewis acids [16–19]. In the past decade, materials based on TiO_2 and ZrO_2 have also been used as alternative stationary phases for liquid chromatography due to their adsorption capacity, selectivity and chemical stability when used under extreme pH and temperature conditions [20–22].

Due to the nature of the phosphate-metal oxide interaction, the enrichment of metabolites containing phosphate functional groups is more selective under low pH conditions (pH 2-3) [23], as non-specific interactions will be minimized. Theoretically, the phosphate group remains negatively charged under low pH conditions, while the protonation of carboxylic acids should prevent binding to the sorbent [15]. However, controlling pH alone is not sufficient to steer selectivity in matrix-loaded biological samples and optimization of all SPE steps is thus required. Hence, mixtures of trifluoroacetic acid (TFA) or formic acid (FA) combined with acetonitrile (ACN) or methanol (MeOH) at levels of 50–80% (v/v) were initially applied. Later, the use of 2,5-dihydroxybenzoic acid, lactic acid, β-hydroxypropanoic acid and glycolic acid were also reported to control pH and to prevent non-specific hydrophobic interactions with the sorbent [11,20,24–28]. However, the results regarding efficiency of preventing non-specific binding using these mixtures were quite contradictory and hence highlight some of the difficulties in selecting the optimal SPE conditions, particularly with regard to the role of the sample matrix.

One crucial step during SPE is the washing of the metal oxide after the binding of the phosphorylated analytes prior to their elution. Inappropriate washing conditions can lead to losses of analytes, incomplete removal of residual chemicals or non-specifically bound matrix compounds [26]. As washing solution, pure deionized water, combinations of several organic acids (e.g. TFA, acetic acid, lactic acid), combinations of ACN/MeOH or ammonium bicarbonate in ACN were used [29–32]. After washing, the bound compounds were eluted following the frequently used approach employing high pH (9-11) inorganic solvents such as NH_4OH [24,29,33,34], ammonium phosphate [32] or ammonium bicarbonate [10]. Kyono et al. reported that the elution efficiency was independent of pH, and improved for the hydroxyl acid-modified metal oxides when secondary monoamines such as piperidine or pyrrolidine were added to the eluents [31].

While much work has been dedicated to phosphoproteomics, the enrichment of phosphorylated primary metabolites using MOAC has not received much attention. In 2004, Sekiguchi's group demonstrated the use of TiO_2 columns to trap sugar phosphates (galactose 1-phosphate, glucose 1-phosphate, sucrose 6-phosphate, glucose 6-phosphate, fructose 6-phosphate, fructose 1,6-diphosphate, UDP-glucose) present in plant extracts for subsequent analysis using anion exchange chromatography with pulsed amperometric detection [35] and MS/MS [36]. These studies also highlighted some difficulties regarding the quantitative determination of some sugar phosphates (e.g. degradation of fructose 1,6-disphosphate during microwave-assisted extraction, moderate recovery of fructose 6-phosphate, several peaks for erythrose 4-phosphate). Since then, new commercial MOAC (TiO_2, ZrO_2) products have been made available from many manufacturers, but their usage for metabolomics applications has not yet been investigated.

In this work, we address the challenging quantitative determination of phosphorylated metabolites at low concentrations in complex biological matrices. We present a workflow combining GC- and IC-MS/MS analysis with a novel selective MOAC solid phase extraction method for cleanup/enrichment and quantification of 12 phosphorylated primary metabolites from glycolysis and pentose phosphate pathways. Several SPE materials designed for either application in phosphopeptide enrichment or phospholipid depletion workflows were tested for their suitability. Due to the fact that diverse errors can easily be introduced throughout the whole analytical workflow including solid phase extraction, derivatization and GC-MS/MS as well as IC-MS/MS analysis [37,38], an uniformly $U^{13}C$-labeled biological cell extract was employed as internal standard ($U^{13}C$ ISTD) for all analyses during method development and evaluation.

2. Material and methods

2.1. Chemicals

2-Phosphoglyceric acid disodium salt hydrate (2PG), 3-phosphoglyceric acid disodium salt (3PG), dihydroxyacetone phosphate lithium salt (DHAP), erythrose 4-phosphate sodium salt (E4P), fructose 6-phosphate disodium salt hydrate (F6P), fructose1,6-bisphosphate trisodium salt (FBP), glucose 6-phosphate dilithium salt (G6P), glycerol 3-phosphate bis(cyclohexylammonium) salt (Gro3P), mannose 6-phosphate disodium salt hydrate (M6P), phosphoenolpyruvic acid (PEP), ribose 5-phosphate disodium salt hydrate (R5P), D-sedoheptulose 7-phosphate lithium salt (S7P), alpha-aminoadipic acid (AAA), alpha-ketoglutaric acid (AKG), citric acid monohydrate (Cit), DL-isocitric acid trisodium salt (i-Cit), fructose (Fru), glucose (Glc), L-lysine (Lys), L-proline (Pro), fumaric acid (Fum), DL-serine (Ser), L-cystathionine (LCT), mannitol (Man-OL), mannose (Man), galactose (Gal), succinic acid (Suc), and L-tryptophan (Trp) were purchased from Sigma Aldrich (Vienna, Austria). L-aspartic acid (Asp), glycine (Gly), DL-malic acid (Mali), L-methionine (Met), and L-valine (Val) were purchased from Merk (Darmstadt, Germany). Single stock solutions were prepared in amber glass vials in LC-MS grade water (CHROMASOLV™ LC-MS, Honeywell) containing suitable additives (0.1 mol L^{-1} HCl, 0.1 mol L^{-1} NaOH) when necessary. All single stocks were stored at −80 °C. The working standards were prepared daily in LC-MS grade water. Formic acid (FA), and trifluoroacetic acid (TFA) were purchased from Sigma Aldrich. Ethoxyamine hydrochloride (EtOx) was purchased from Sigma Aldrich, N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS) from Thermo Scientific (Waltham, MA, USA). The EtOx solution was prepared daily by dissolving an appropriate amount of solid substance in water-free pyridine (Sigma Aldrich).

2.2. Internal standards

$U^{13}C$ fructose 1,6-bisphosphate ($U^{13}C$ FBP) was purchased from Omicron Biochemicals (South Bend, USA), $U^{13}C$-labeled metabolites internal standard derived from fully ^{13}C-labeled *P. pastoris* extraction ($U^{13}C$ ISTD) was obtained from ISOtopic Solutions (Vienna, Austria).

2.3. Reconstitution of P. Pastoris samples and $U^{13}C$ ISTD

Ethanolic cell extracts of *P. pastoris* were obtained from ISOtopic Solutions (Vienna, Austria) and were stored at −80 °C. These *P. pastoris* samples contain a dried extract of more than 2×10^9 cells (corresponding to approximately 15 mg cell dry weight (CDW)). Before use, the natural yeast extract and the $U^{13}C$ labeled yeast extract (ISTD) were reconstituted to 2.0 mL with LC-MS grade water. The $U^{13}C$ ISTD was further diluted by a factor of 10 with LC-MS grade water before further use in some experiments.

2.4. Quantitative determination of primary metabolites in P. pastoris

2.4.1. Quantitative GC-MS/MS analysis

Sample aliquots of 60 µL were pipetted into 200 µL flat bottom glass inserts in 1.5 mL GC-vials. After addition of 60 µL of EtOx in pyridine (c_{EtOx} = 18.7 mg mL^{-1}) for protection of carbonyl groups during evapo-

ration step [39], all samples were dried using a vacuum centrifuge and then transferred to a Gerstel MPS2 robot (Gerstel, Germany) for automated just-in-time online two-step derivatization before measurement with GC-MS/MS.

The two-step derivatization procedure was modified based on the derivatization parameters published by Koek et al. [9]. First, the dried sample was reconstituted in 18 µL EtOx in water-free pyridine (c_{EtOx} = 18.7 mg mL^{-1}) and incubated at 40 °C for 90 min for ethoximation, followed by silylation with 42 µL MSTFA with 1% TMCS at 40 °C for 50 min. The derivatized samples were kept at 4 °C for 5 min for sample cool down and were then injected into the GC-MS/MS system.

For the analysis by GC-MS/MS, an Agilent Technologies 7010B GC-MS/MS Triple Quadrupole system (Agilent, Waldbronn, Germany) with electron ionization (EI) source was used. A multipurpose sampler (MPS2, Gerstel, Germany) was used for automated two-step derivatization of the analytes and injection into a programmed temperature vaporizer (KAS6, Gerstel). The sensitivity and mass accuracy of the mass spectrometer was checked before starting analysis with an internal calibrant solution containing perfluorotributylamine (PFTBA) and perfluoro-5,8-dimethyl-3,6,9-trioxydodecane (PFDTD) applying the mass calibration protocol of the system. Multi Baffled PTV Liners with Siltek coating (Gerstel) were changed after approximately one hundred injections. The derivatives were separated on a nonpolar Optima 1 MS Accent analytical column (Macherey - Nagel, Germany, 60 m × 0.25 mm i.d., 0.25 µm film thickness, 100% dimethylpolysiloxane stationary phase). For protection of the analytical column, a deactivated nonpolar guard column (3 m × 0.25 mm i.d., Phenomenex) was used. For the analytical column, helium was used as carrier gas at a constant flow rate of 1.3 mL min^{-1}. Injection of 1.0 µL aliquots of sample solution was performed applying programmable temperature vaporization (PTV) (70 °C for 0.1 min, 12 °C min^{-1} to 260 °C (hold 1 min), 12 °C min^{-1} to 300 °C, 5 min hold). A splitless inlet mode was used with purge flow to split vent is 50 mL min^{-1} at 2 min. The GC temperature gradient was started at 70 °C (hold 1 min), then ramped at 15 °C min^{-1} to 190 °C, at 5 °C min^{-1} to 225 °C, at 3 °C min^{-1} to 255 °C and finally at 25 °C min^{-1} to 310 °C (hold 5 min). Total cycle time was 33.2 min. Information of all precursor and product ions as well as collision energy used can be found in supplementary material (Table S1).

2.4.2. Quantitative IC-MS/MS analysis of FBP and PEP

FBP was reported not to be stable during trimethylsilylation derivatization and hence suffers from poor derivatization efficiency (< 10%) [39]. Furthermore, PEP shows very low sensitivity when analyzed with GC-MS/MS. For the measurement of FBP and PEP an IC-MS/MS method was thus employed.

Before IC-MS/MS analysis, 375 µL aliquots of the effluents containing 30 µL of U^{13}C FBP (with a concentration of 25 µmol L^{-1}) were pipetted into 1.5 mL vials. An HPIC system (Thermo Fisher Scientific IC5000+) with suppressed conductivity and mass spectrometric detection was used. The instrument was operated in external water mode (in which a second pump system was used to supply the electrolytic membrane suppressor with regenerant water) and was coupled with a Thermo TSQ Vantage Triple Quadrupole Mass Spectrometer (Thermo Fisher Scientific). The samples were injected via an AS-AP autosampler with push-full injection mode using a 30 µL fixed loop (overfill factor 2). A Dionex IonPac AS11-HC column (4 µm, 2 × 250 mm) preceded by a Dionex IonPac AG11-HC guard column (4 µm, 2 × 50 mm) was used for anion exchange chromatography with an AERS 500e 2 mm electrolytic membrane suppressor (Thermo Fisher Scientific) placed between the column and the conductivity detector. A suppressor current of 30 mA was applied for the entire analytical run. A multistep electrolytically generated gradient program was used with an initial concentration of 10 mmol L^{-1} KOH ramped up to 60 mmol L^{-1} from 0 to 0.5 min, then held isocratic until 5.5 min, returned to 10 mmol L^{-1} KOH at 6.1 min and held until 13 min. The analytical and suppressor regenerant flow rate were both set to 0.2 mL min^{-1}. Temperature of the column compartment and cell heater were 30 °C and 35 °C, respectively. The suppressed effluent exiting the conductivity cell was directly connected to the heated electrospray ionization (HESI) source of a Thermo TSQ Vantage. Another pumping system was used to supply the ACN and mixed with effluent before entering HESI for enhancing ionization (flow rate of 0.1 mL min^{-1}). The vaporizer and capillary temperature were set to 350 °C, sprayer voltage to 3000 V, aux gas pressure to 15 arbitrary units, sheath gas pressure to 40 arbitrary units and collision gas pressure for multiple reaction monitoring to 1.5 mTorr. Using negative ionization mode, transitions for FBP were 339 → 241 (15 V CE) and 339 → 97 (20 V CE), transitions for PEP were 167 → 79 (37 V CE) and 167 → 63 (66 V CE). A MRM chromatogram of the PEP and FBP separation is shown in Supplementary material (Fig. S1).

2.5. SPE materials

The tested MOAC materials were HybridSPE®-Phospholipid (ZrO_2) cartridges (SUP) purchased from Sigma-Aldrich (Vienna, Austria), Phosphopeptide Purification ZrO_2 iTip (PET) purchased from GlySci (Columbia, USA), MagReSyn® TiO_2 (MTO) and MagReSyn® ZrO_2 (MZO) materials obtained from ReSyn Biosciences (Gauteng, South Africa). We additionally tested immobilized metal affinity chromatography (IMAC) materials, which are resins with associated metal ions [15]: Phosphopeptide Enrichment Spin Columns (PES) purchased from Takara (CA, USA), MagReSyn® Ti-IMAC (MTI) and MagReSyn® Zr-IMAC (MZI) obtained from ReSyn Biosciences (Gauteng, South Africa). Further information for all materials can be found in the Supplementary material (Table S2).

2.6. Preliminary comparison of MOAC materials with regard to breakthrough during loading

For the comparison of different cleanup/enrichment materials, three conditioning/loading procedures (referred to as A, B, C)were selected based on recommend procedures from respective manufacturers, then modified to fit for the purpose. These procedures were tested with *P. pastoris* cell extracts to assess their performance for selective cleanup/enrichment of phosphorylated metabolites while maintaining simplicity, speed, low cost and compatibility with the measurement methods employed. A summary of the three procedures is provided in Supplementary material (Table S3). In procedure A, the samples were acidified with concentrated FA to a final FA concentration of 2% and then diluted by a factor of 2 with 1% FA in MeOH (loading solution) prior to loading. The materials were conditioned with 780 µL of loading solution. Then, 780 µL of the diluted sample (amount of cell extract corresponding to 2.8 mg CDW) was loaded to each material. Procedure B was similar to procedure A with the exception that the loading solution was 1% (v/v) FA in 45% (v/v) MeOH, 45% (v/v) ACN, 9% (v/v) water. In procedure C, the samples were acidified with 10% TFA to a final TFA concentration of 0.4%, then diluted by a factor of 2 with a loading solution comprising 0.1% (v/v) TFA in 90% (v/v) ACN. A volume of 780 µL of loading solution was used for conditioning. For these initial comparisons, only GC-MS/MS was used for measurements. 60 µL aliquots of the eluents and 60 µL of the 1:10 diluted ISTD were pipetted into 200 µL flat bottom glass inserts in 1.5 mL GC-vial. After addition of 60 µL of EtOx in pyridine (c_{EtOx} = 18.7 mg mL^{-1}), all samples were dried using a vacuum centrifuge and then transferred to a Gerstel MPS2 robot (Gerstel, Germany) for automated just-in-time online two-step derivatization before measurement with GC-MS/MS.

2.7. Assessment of breakthrough, washing and elution recoveries

For MOAC material comparison and later method development, the SPE was performed with the pure samples diluted with loading solution only. Diluted $U^{13}C$ ISTD was spiked into the effluents of the loading, washing, and elution steps *after* SPE and prior to analysis. For the calculation of recovery, reference samples were spiked with the $U^{13}C$ ISTD and analyzed directly omitting the SPE steps. All samples used for recovery determination are referred to as R-samples.

For these experiments, the cartridges were first conditioned with conditioning solution. Yeast cell extracts were acidified according to the loading procedures, diluted by a factor of two with the loading solution and subsequently loaded onto the cartridges. 100 µL aliquots of the eluents and 100 µL of the 1:10 diluted ISTD were pipetted into 200 µL flat bottom glass inserts in 1.5 mL GC-vials and were then analyzed by GC-MS/MS. For IC-MS/MS analysis see section 2.4.2. The compound fractions in the breakthrough, washing and eluent solutions were then calculated using equations (1)–(3) (in which the compound fraction quantified in the eluent represent the overall compound recovery of the method for the respective analyses):

$$x_{b,\ \%} = \frac{c_b}{c_s} \times 100 \quad (1)$$

$$x_{w,\ \%} = \frac{c_w}{c_s \times \frac{V_s}{V_w}} \times 100 \quad (2)$$

$$x_{e,\ \%} = \frac{c_e}{c_s \times \frac{V_s}{V_e}} \times 100 \quad (3)$$

where x_b, x_w, x_e are the recoveries in effluents of loading, washing and eluting steps expressed as molar fractions (%), c_b, c_w, c_e are compound concentrations (µmol L^{-1}) in the effluents; V_s, V_w, V_e are the volumes of the sample, washing and elution fractions.

2.8. $U^{13}C$ ISTD addition for the correction of errors caused by the analytical workflow

For the assessment of precision under repeatability conditions of measurement, the yeast cell extracts were subjected to the optimized SPE procedure.

The samples were either spiked with the undiluted $U^{13}C$ ISTD *prior to* the sample loading. These samples are referred to as P-samples. 100 µL aliquots of the effluents were dried for GC-MS/MS analysis ($n = 5$).

In a second, alternate approach, effluents of unspiked samples were spiked with the diluted ISTD *after* SPE (R-samples). 100 µL aliquots of these samples were spiked with 100 µL diluted ISTD before GC-MS/MS analysis for determination of recovery and assessment of precision under repeatability conditions ($n = 3$).

2.9. Data processing and evaluation

Data acquisition and evaluation for GC-MS/MS were carried out with MassHunter Acquisition B.07.05.2479, MassHunter Quantitative QQQ B.09.00 (Agilent Technologies, CA, USA). IC-MS/MS data acquisition and evaluation were performed using Xcalibur 2.3 and Tracefinder 3.1 (Thermo Fisher Scientific).

3. Results and discussion

3.1. Selection of MOAC materials for cleanup/enrichment

The *P. pastoris* cell extract used for this set of experiments is known to have a complex matrix that includes highly concentrated non-phosphorylated compounds such as carboxylic acids, amino acids and sugars [40–42]. On the other hand, intracellular concentrations of phosphorylated metabolites are comparatively low from sub-µmol L^{-1} to few µmol L^{-1} [7,43]. The initial goal of SPE cleanup/enrichment was to achieve efficient enrichment in the loading step, and at the same time, high breakthrough for non-phosphorylated matrix compounds (i.e. amino acids, carboxylic acids and sugars). Although MTI, MZI, and PES are IMAC materials, they were nevertheless tested in our initial experiments. Based on the adsorption efficiency of the phosphorylated compounds of interest in the loading step and the degree of removal of non-phosphorylated compounds, the most suitable material and procedure were selected and further optimized. Parameters like loading amount, elution volume and elution order were tested with regard to their impact on analyte recovery in the eluent. The fractions of the phosphorylated compounds in the breakthrough solutions were calculated according to equation (1) for all materials and procedures. Results revealed more efficient adsorption of phosphorylated compounds on SUP, which is a ZrO_2-based cartridge containing a high amount of solid phase (see Supplementary material, Table S2) in comparison to the other materials (see Supplementary material, Figs. S2a and b). Furthermore, increasing ACN concentration in the loading solution (procedure B) resulted in slightly lower concentrations of the phosphorylated metabolites in the breakthrough solution of SUP. For the non-phosphorylated matrix compounds, all materials exhibited high breakthrough with more than 70% for the studied compounds using procedure A and B, except for SUP (see Table S4). Regardless of the procedure employed, SUP tended to efficiently bind Cit, Mali, and Lys. Adding ACN to the loading solution (procedure B) increased the breakthrough of Cit and Mali from 4% to 40% and from 17% to 44%, respectively. Breakthrough fractions of other non-phosphorylated compounds are provided in the Supplementary material (Table S4). These results are in accordance with the finding that increasing ACN concentration in the loading solution helps to prevent unwanted binding of some compounds for SUP [11,32,44]. For other materials, no effect was apparent. Possible explanations for the more efficient binding of Cit, Mali and Lys using SUP could be either the higher binding capacity due to the higher amount of SPE material (30 mg) in the cartridge compared to others materials, or the relatively strong Lewis basicity of Cit and Mali, which can influence retention on zirconia materials [44]. From the point of view of ion exchange-type interactions, the pK_a values of these acids may also influence retention processes [23], which would be reflected by the fact that Cit (pK_a of 3.1 and 4.7) and Mali (pK_a of 3.5 and 5.03) are more efficiently bound than Suc (pK_a of 4.2 and 5.6). Like SUP, the investigated PET tips contain a ZrO_2 phase, but a very low mass of material (75 µg), which could be the main reason for the high breakthrough of both phosphorylated and non-phosphorylated compounds on this material.

Using 0.1% TFA in 90% ACN as loading/conditioning solution (procedure C), all materials yielded good results concerning the removal of interfering matrix compounds with approximately 100% breakthrough for the studied non-phosphorylated metabolites, except for some organic acids on SUP (Table S4). For SUP cartridges, low breakthrough was observed for Cit (3%), Mali (26%), AKG (63%) and Lys (70%). Again, an attempt to explain the results for binding of interfering compounds based on pH only is unconvincing due to the fact that the pH of the sample was already lower than the pK_a values of the carboxylic acids (although adjusted to be greater than the pK_a value of the phosphorylated metabolites) in order to prevent unwanted binding. This

may indicate the important role of the analyte functional groups, since compounds that contain multiple carboxylic acid or amine groups may exhibit better retention on ZrO_2 materials [44–46]. However, for important phosphorylated metabolites, SUP exhibited best results with no detectable breakthrough for 3PG and G6P, 5% for 2PG, 7% for DHAP and 10% for M6P (Fig. 1).

Hence, compared to other tested materials, the use of SUP with procedure C showed the lowest breakthrough for the primary phosphorylated compounds, while at the same time provided partial removal of sugars, organic acids and amino acids. Therefore, SUP was chosen for further procedure optimization and validation.

3.2. Optimization of SPE cleanup/enrichment for phosphorylated metabolites in P. pastoris cell extracts using SUP

3.2.1. Optimization of washing solution composition

Various washing solutions were investigated for SUP including 0.1% TFA in 90% ACN, pure MeOH and pure H_2O. The compound losses during the washing step were calculated using equation (2). Washing cartridges with 0.1% TFA in 90% ACN, which is also the composition of the loading solution of procedure C, yielded the lowest losses (less than 5%) for all investigated phosphorylated metabolites (except for Gro3P), while washing with MeOH caused higher losses such as 12% of 3PG, 14% of G6P and 7% of M6P. Washing with H_2O caused losses of up to 38% for 3PG 35% for G6P, and 16% for M6P. Therefore, 0.1% TFA in 90% ACN remained the best choice for the washing step of the SPE cleanup/enrichment procedure.

3.2.2. Optimization of SPE elution

It was reported that the recovery of phosphopeptides can benefit from the addition of secondary monoamines such as piperidine and pyrrolidine to the eluent for the hydroxyl acid-modified metal oxides [31]. Since pyridine was used in our GC derivation procedure, pyridine was tested as elution solvent for the first time. Compound recovery in the elution step was calculated using equation (3). Preliminary results with 100 µL of 10% pyridine in MeOH showed low recovery of the analytes, in more detail 23% for 2PG, 14% for DHAP, 17% for F6P, 29% for Gro3P, 21% for G6P and 22% for M6P. Preliminary results showed that washing with H_2O and MeOH also eluted the phosphorylated metabolites from the cartridge, hence both of these solvents were included in the optimization of a sequential elution procedure.

One important factor for sequential elution is the order of elution due to the influence of the previous eluent upon subsequent elution steps. Therefore, two elution protocols with different orders were tested. Procedure E1 used the order: (1) H_2O, (2) 20% NH_4OH, (3) MeOH and (4) 10% pyridine in MeOH, while the order of the E2 procedure was: (1) 10% pyridine in MeOH, (2) MeOH, (3) H_2O and (4) 20% NH_4OH. Each elution step was carried out using a volume of 100 µL, resulting in a total elution volume of 400 µL, which was collected in a single vial. The results showed that procedure E1 led to significantly higher recoveries for all phosphorylated compounds with recoveries ranging from 46% (S7P) to 92% (Gro3P), while the maximum recovery of procedure E2 procedure was only 58% (Gro3P). Hence, elution protocol E1 was chosen for all further work.

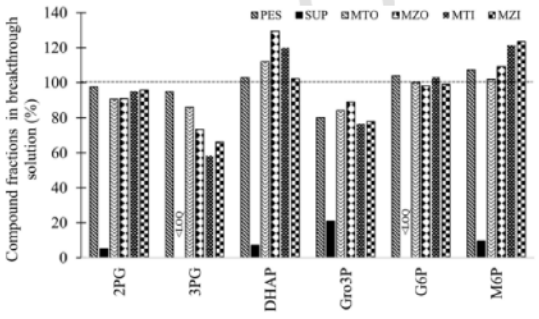

Fig. 1. Compound fractions in the breakthrough solution (%) using 0.1% TFA in 90% ACN as conditioning and loading solution (procedure C), $n = 3$ for SUP, $n = 1$ for all other materials, loaded amount of cell extract corresponding to 2.8 mg CDW of *P. pastoris*.

Next, the volume of the elution fractions needed to be adjusted to provide a compromise regarding repeatability, enrichment factor and elution efficiency. For this purpose, six cartridges were loaded with a CDW of 2.8 mg of *P. pastoris* using the loading/conditioning/washing conditions of procedure C. Then three cartridges were eluted employing elution procedure E1 with a total elution volume of 200 µL (50 µL per elution step), while the other three cartridges were eluted with a total volume of 400 µL (100 µL per elution step). The results showed higher amount (nmol) for most of the studied phosphorylated compounds in eluents when using 100 µL for each step (Fig. 2). Thus, a total elution volume of 400 µL was employed (4 × 100 µL) for all further experiments, resulting in an enrichment factor of approximately 2. This two-fold reduction of the enrichment factor in comparison to a total volume of 200 µL can easily be compensated by doubling the volume of the aliquots, which are dried for GC-MS/MS analysis.

3.2.3. Effect of the loaded sample amount

Six SUP cartridges were loaded with two different sample amounts (corresponding to CDW of 2.8 mg and 6.0 mg of *P. pastoris*, respectively) using procedure C, washed with same conditioning solution and eluted with elution procedure E1. As more sample loaded onto the cartridge implies a higher amount of loaded matrix, a cartridge will eventually reach its binding capacity and breakthrough will occur. Fig. 3 shows the concentration of phosphorylated metabolites (µmol L^{-1} CDW^{-1}) recovered from eluents using SUP cartridges with two loaded sample amounts. Results ($n = 3$) showed that the concentration of all compounds were higher using a sample amount of 2.8 mg CDW, and significantly lower using a sample amount of 6.0 mg CDW. For this reason, a CDW of 2.8 mg of yeast cell extract was used for all further experiments.

In summary, the optimized SPE conditions are as follows (Fig. 4): conditioning of the cartridge with 780 µL of 0.1% TFA in 90% ACN, acidification of the sample (sample amount = 2.8 mg CDW) with 10% TFA to a final TFA concentration of 0.4%, dilution of the sample by factor of two with the conditioning solution, loading of the sample onto the cartridge, washing of the cartridges with 780 µL of conditioning so-

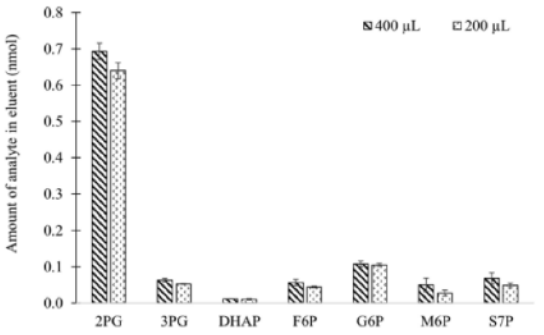

Fig. 2. Effect of total elution volume using four-step sequential elution on recovery of phosphorylated metabolites in the eluent from SUP cartridges ($n = 3$, loaded sample amount corresponding to 2.8 mg CDW).

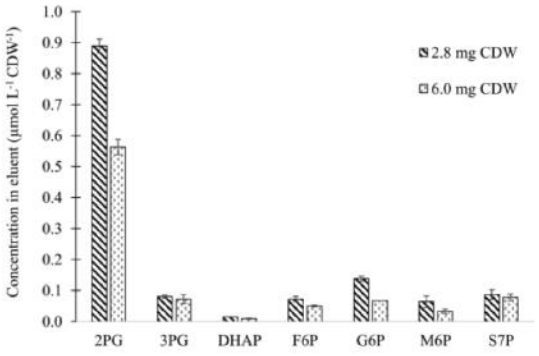

Fig. 3. Concentration of phosphorylated metabolites recovered in the eluent using four-step sequential elution in dependence of sample amount loaded onto SUP cartridges ($n = 3$).

lution, then sequential eluting using a four-steps elution with H_2O, 20% NH_4OH, MeOH, and 10% pyridine in MeOH (100 µL each).

3.2.4. Influence of matrix on GC-MS/MS

As the SPE cleanup procedure is expected to have a strong effect on the composition of the sample matrices, this should in turn positively affect the later GC derivatization. By calculating the peak area ratio (R_{13C}) of phosphorylated $U^{13}C$ ISTD compounds spiked in samples which have undergone the SPE procedure and also in samples which were not subjected to cleanup/enrichment ($n = 3$), the effect could be revealed in terms of derivatization efficiency (in which R_{13C} is equal to 1 means no effect, greater than 1 means enhancement, and smaller than 1 means reduction). The results showed a significant improvement in dervatization efficiency (according to R_{13C}) for some phosphorylated compounds such as 2PG ($R_{13C} = 2.0 \pm 0.6$) and 3PG (1.4 ± 0.3). However, the peak area ratio for other phosphorylated compounds ranged from 0.31 ± 0.04 (S7P) to 0.92 ± 0.06 (Gro3P), thus indicating a decrease in derivatization efficiency (see Fig. S3). Moreover, in comparison to the calibration standards, the GC retention times in samples, which were not subjected to SPE shift significantly (approximately 0.2 min) for some early eluting phosphorylated metabolites (Fig. 5)

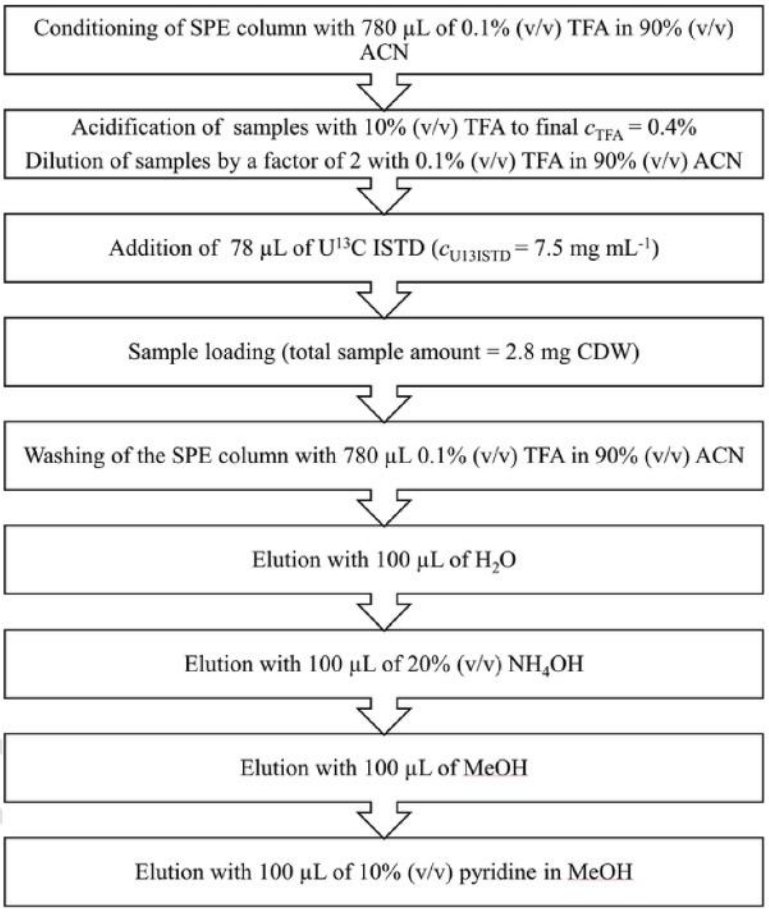

Fig. 4. Final SPE cleanup/enrichment procedure with $U^{13}C$ ISTD correction for the analysis of *P. pastoris* cell extract.

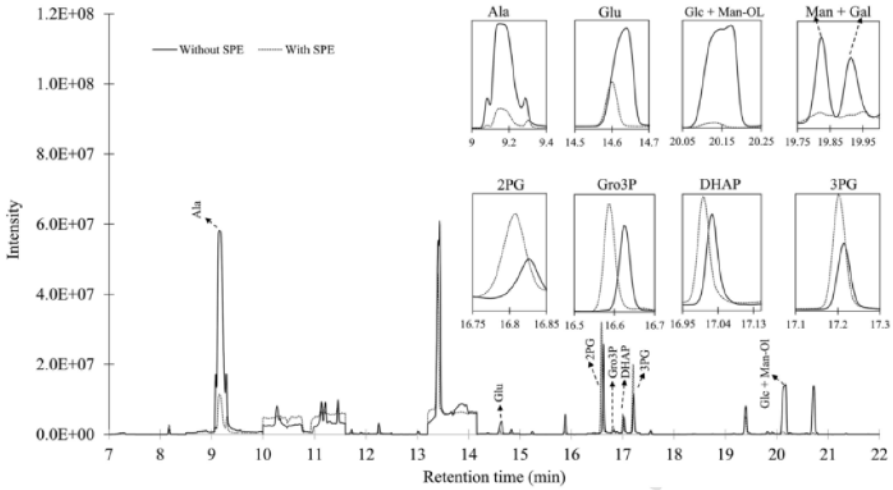

Fig. 5. GC-MS/MS total ion chromatogram (TIC of MRM transitions) obtained of *P. pastoris* cell extracts without SPE cleanup (solid line) and with SPE cleanup (dashed line) using SUP (loaded sample amount, cell extract corresponding to 2.8 mg of *P pastoris* CDW).

due to the matrix influence. In contrast to this, the eluent of the SPE cleanup procedure yielded the same retention times as measured in the standards, thereby assisting data evaluation.

In terms of removal of unwanted compounds, the results ($n = 3$) showed efficient removal of Glc as well as most of the measured amino acids. Organic acids like Glu, Asp, Fum and Suc were also substantially removed (see Fig. 5 and Table S5 in the Supplementary material).

Fig. 5 shows GC-MS/MS total ion chromatograms (TICs, calculated from the MRM transitions measured) for cell extracts corresponding to a CDW of 2.8 mg of *P. pastoris*. Both a cell extract which was measured directly as well as a cell extract, which was measured after SUP cleanup/enrichment are shown. A significant reduction in intensity of highly abundant matrix peaks such as organic acids (e.g. Suc, Fum, Glu) and sugars was observed, while the phosphorylated metabolite peaks (e.g. Gro3P, 2PG and 3PG) were enriched and the intensity is increased in comparison to the non-treated sample.

3.3. Assessment of recovery, precision and correction of errors using U^{13}C labeled isotopologues for internal standardization

For the compensation of errors contributing to the overall trueness and precision of the entire analytical workflow (from SPE to GC-MS/MS or IC-MS/MS analysis), it is necessary to add U^{13}C isotopologues as ISTD at the beginning of the sample preparation process *prior to* SPE.

The biologically generated U^{13}C *P. pastoris* cell extract offers the advantage that it contains all yeast metabolites contained in and extracted from the cells under production conditions. As *P. pastoris* can be grown on ^{13}C glucose as sole carbons source, the cell extract contains a multitude of metabolites from diverse metabolic pathways, thereby containing also the ^{13}C labeled isotopologues of all analytes of interest this study focuses on. When using internal standardization with ^{13}C isotopologues, one must be aware of two factors strongly affecting the accuracy of the correction and hence the entire workflow. First, only internal standardization with an isotopologue, which is chemically identical to the analyte, can compensate for compound-specific effects. Second, accurate internal standardization is based on the addition of the *same* internal standard to all samples and standards, hence all standards and samples must be spiked at the same time point in order to avoid alteration of unstable compounds in the sample or standard.

For the final assessment of the optimized procedure, two sets of samples (R and P), applying the same optimized SPE cleanup/enrichment conditions, were collected simultaneously. R-samples were used for the assessment of the recovery and repeatability of the SPE steps adding the internal standard *after* the SPE procedure, while P-samples were used to investigate the improvement of repeatability by addition of ISTD *prior to* sample loading on the SUP cartridge, thereby also compensating for SPE errors.

To assess precision under repeatability conditions of measurement compound fractions in the different SPE steps, all effluents from the SPE steps (R-samples) were collected separately and spiked as described in section 2.8 before measurement by GC-MS/MS or IC-MS/MS ($n = 3$).

The quantitative results for the SPE fractions showed that there was no measurable breakthrough for 3PG, G6P, PEP and FBP and less than 5% breakthrough for 2PG and R5P. The highest breakthrough among the remaining phosphorylated metabolites was 21% for Gro3P. There were no measurable washing losses for 2PG, 3PG, G6P, M6P, PEP, and FBP, and a maximum loss during washing of 4% for F6P (Fig. 6). Concerning recovery in the eluent, the highest recovery ($n = 3$) was 92% for Gro3P. Four phosphorylated compounds exhibited good recovery (>70%) including 2PG, 3PG, DHAP, and G6P, while moderate recoveries were obtained for F6P, FBP, M6P, PEP, R5P, and S7P (Table 1). Peak distortion and reduction in intensity was observed in both breakthrough and eluting samples for F6P, which may indicate that compound degradation occurred during the cleanup/enrichment. The sum of compound fractions in breakthrough, washing and eluting solutions being significantly less than 100% as observed for FBP, M6P, PEP, R5P, and S7P on SUP indicates chemical alteration (e.g. decomposition) and/or permanent binding of these compounds onto the ZrO_2 surface. GAP showed a recovery of more than 100% indicating the decomposition of a matrix constituent during SPE finally resulting in formation of GAP.

P-samples were used to check whether SPE introduces errors, which cannot be compensated by the ISTD. For this purpose, P-samples were spiked with U^{13}C ISTD and measured directly ($n = 3$) as well as their eluent ($n = 3$). Theoretically, the measured concentrations of the analytes should be similar for these two types of samples (independent of the fact that one set of samples has undergone SPE) since these concen-

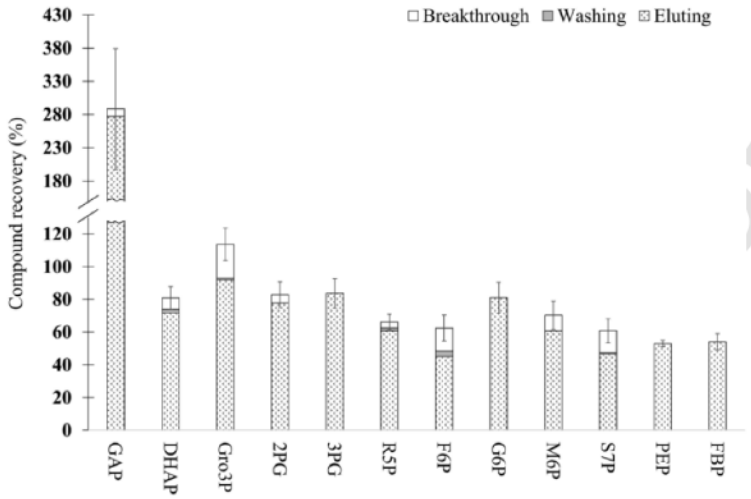

Fig. 6. Sum of fractions (%) of phosphorylated metabolites in the SPE solutions using SUP for cleanup/enrichment (measurements carried out by GC-MS/MS except for PEP and FBP).

Table 1
Recovery of phosphorylated metabolites in the SPE eluent using SUP and the optimized cleanup/enrichment procedure ($n = 3$), loaded amount of cell extract corresponding to 2.8 mg CDW of *P. pastoris*. Standard deviation was calculated using the law of error propagation.

	Compound recovery in eluent (%)	Method
GAP	277 ± 91	GC-MS/MS
DHAP	72 ± 7	GC-MS/MS
Gro3P	92 ± 8	GC-MS/MS
2PG	78 ± 7	GC-MS/MS
3PG	84 ± 9	GC-MS/MS
R5P	61 ± 5	GC-MS/MS
F6P	45 ± 7	GC-MS/MS
G6P	81 ± 9	GC-MS/MS
M6P	61 ± 8	GC-MS/MS
S7P	46 ± 5	GC-MS/MS
PEP	53 ± 2	IC-MS/MS
FBP	54 ± 5	IC-MS/MS

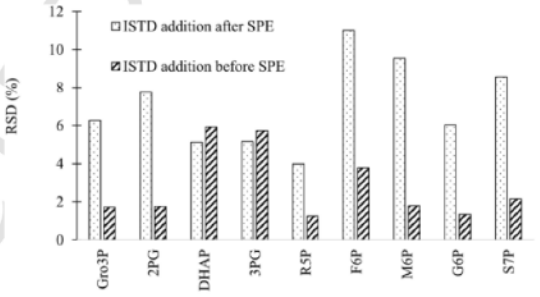

Fig. 7. Repeatability of GC-MS/MS analysis of investigated phosphorylated compounds after SPE ($n = 3$, loaded sample amount = 2.8 mg CDW). $U^{13}C$ ISTD addition before SPE (▨) compensates for errors of SPE as well as GC-MS/MS analysis, while addition of $U^{13}C$ ISTD after SPE (□) compensates only for errors of GC-MS/MS analysis. All compounds were corrected using the $U^{13}C$ ISTD.

trations were determined by external calibration and internal standardization before SPE. All errors introduced by SPE should be compensated by the use of the $U^{13}C$ ISTD. Hence, the ratio of the concentrations of the two types of samples should be 1, expressing the fact that the measured concentrations of the phosphorylated metabolites are the same for samples with and without SPE treatment, which is a precondition for reliable quantification. Among all studied phosphorylated compounds, only GAP was found to show significant deviations from 1 (see Supplementary material, Table S6). The GAP amount in the eluent of the SPE was increased by a factor of 4 in comparison to the original sample. The high recovery value for GAP in the eluent (Fig. 6) is most likely explained by the degradation of matrix constituents forming GAP, which cannot be compensated by the use of internal standards.

By adding the internal standard early in the sample preparation workflow, degradation of analytes, recovery of SPE, matrix-dependent derivatization efficiency, losses during injection (GC) and ion suppression (IC-ESI-MS) can be compensated. A comparison of RSDs obtained spiking the samples before and after the SPE procedure is shown in Fig. 7. The Maximum RSD value obtained in R-samples was 11% (F6P), while highest RSD value obtained in P-samples is only 6.0% (DHAP). The results showed that although there was no significant improvement in precision for DHAP and 3PG when the $U^{13}C$ ISTD was added prior to SPE, a significant improvement was observed for all other phosphorylated metabolite under investigation. Clearly, by spiking the $U^{13}C$ ISTD at the beginning of the workflow, a general improvement in precision under repeatability conditions of measurement could be achieved resulting in RSD values of less than 6.0% for all investigated phosphorylated compounds.

Adding the $U^{13}C$ ISTD to the yeast cell extract *before* SPE indeedenables accurate quantitative analysis of the phosphorylated primary metabolites (with the exception of GAP) and improved repeatability precision for GC-MS/MS and IC-MS/MS measurements by compensating for errors and short-term fluctuation during sample preparation and measurement.

4. Conclusions

The present work demonstrates the benefits of the combination of a phosphate selective cleanup/enrichment for quantitative assessment of phosphorylated metabolites in cell samples.

Although the enrichment factor of the proposed procedure is low (approximately a factor of 2) the removal of the majority of the matrix compounds improves the method significantly with regard to ruggedness, e.g. GC retention time shifts between sample and standards were fully eliminated. The observed impact of the matrix on the derivatization efficiency of the phosphorylated compounds, showing both shifts towards higher and towards lower efficiencies depending on the compound, clearly underlines the need of internal standardization with labeled isotopologues.

Evidently, the method could be scaled up to higher enrichment factors and hence sensitivity using higher sample volumes and SUP bed weights. As a drawback, the analysis cost per sample would increase significantly due to the higher consumption of $U^{13}C$ ISTD. Nevertheless, the efficient removal of sample matrix enables a significant enhancement of the amount of sample subjected to GC-MS/MS analysis employing the proposed procedure without further scale-up.

Finally, it could be shown that the compound-specific internal standardization using stable isotope labeled isotopologues enables quantification based on a ratio method as defined by the Consultative Committee for Amount of Substance (CCQM) in 1998, thereby improving precision under repeatability conditions. Hence it can be concluded that the precision under intermediate and reproducibility conditions of measurement will also be significantly improved by the proposed standardization based on $U^{13}C$ isotopologues.

Conflicts of interest

The authors declare no competing financial interest.

Acknowledgments

The authors gratefully acknowledge Philipp Tondl and Simone Panholzer for enabling the use of GC-MS/MS instrumentation, and Halimat Ahmatowa for enabling the use of IC-MS/MS instrumentation (Department of Analytical Chemistry, BOKU Vienna). This work was funded by the ASEA-UNINET on behalf of the Austrian Federal Ministry of Science, Vienna Business Agency and EQ BOKU VIBT GmbH are acknowledged for providing mass spectrometry instrumentation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.talanta.2019.120147.

References

[1] J.K. Nicholson, J. Connelly, J.C. Lindon, E. Holmes, Metabonomics: a platform for studying drug toxicity and gene function, Nat. Rev. Drug Discov. 1 (2002) 153–161, doi:10.1038/nrd728.
[2] O. Fiehn, Metabolomics-the link between genotypes and phenotypes, Plant Mol. Biol. 48 (2002) 155–171. http://www.ncbi.nlm.nih.gov/pubmed/11860207. [accessed 8 February 2017].
[3] O. Schneewind, D.M. Missiakas, Protein secretion and surface display in Gram-positive bacteria, Philos. Trans. R. Soc. Lond. B Biol. Sci. 367 (2012) 1123–1139, doi:10.1098/rstb.2011.0210.
[4] B. Luo, K. Groenke, R. Takors, C. Wandrey, M. Oldiges, Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography-mass spectrometry, J. Chromatogr. A 1147 (2007) 153–164, doi:10.1016/j.chroma.2007.02.034.
[5] R. Ramautar, G.W. Somsen, G.J. de Jong, CE–MS for metabolomics: developments and applications in the period 2014–2016, Electrophoresis 38 (2017) 190–202, doi:10.1002/elps.201600370.
[6] J.H.J. Huck, E.A. Struys, N.M. Verhoeven, C. Jakobs, M.S. Van Der Knaap, Profiling of pentose phosphate pathway intermediates in blood spots by tandem mass spectrometry: application to transaldolase deficiency, Clin. Chem. 49 (2003) 1375–1380, doi:10.1373/49.8.1375.
[7] D.B. Chu, C. Troyer, T. Mairinger, K. Ortmayr, S. Neubauer, G. Koellensperger, S. Hann, Isotopologue analysis of sugar phosphates in yeast cell extracts by gas chromatography chemical ionization time-of-flight mass spectrometry, Anal. Bioanal. Chem. 407 (2015) 2865–2875, doi:10.1007/s00216-015-8521-9.
[8] T. Mairinger, M. Steiger, J. Nocon, D. Mattanovich, G. Koellensperger, S. Hann, Gas chromatography-quadrupole time-of-flight mass spectrometry-based determination of isotopologue and tandem mass isotopomer fractions of primary metabolites for 13C-metabolic flux analysis, Anal. Chem. 87 (2015) 11792–11802, doi:10.1021/acs.analchem.5b03173.
[9] M. Koek, B. Muilwijk, M.J. Werf, T. Hankemeier, Microbial metabolomics with gas chromatography mass spectrometry, Anal. Chem. 78 (2006) 1272–1281, doi:10.1021/ac051683 + .
[10] M.W.H. Pinkse, P.M. Uitto, M.J. Hilhorst, B. Ooms, A.J.R. Heck, Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns, Anal. Chem. 76 (2004) 3935–3943, doi:10.1021/ac0498617.
[11] U.K. Aryal, A.R.S. Rossz, Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry, Rapid Commun. Mass Spectrom. 24 (2009) 3567–3577, doi:10.1002/rcm.
[12] H.K. Kweon, K. Håkansson, Selective zirconium dioxide-based enrichment of phosphorylated peptides for mass spectrometric analysis, Anal. Chem. 78 (2006) 1743–1749, doi:10.1021/ac0522355.
[13] S.R. Schmidt, F. Schweikart, M.E. Andersson, Current methods for phosphoprotein isolation and enrichment, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 849 (2007) 154–162, doi:10.1016/j.jchromb.2006.09.016.
[14] N. Sugiyama, T. Masuda, K. Shinoda, A. Nakamura, M. Tomita, Y. Ishihama, Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications, Mol. Cell. Proteom. 6 (2007) 1103–1109, doi:10.1074/mcp.T600060-MCP200.
[15] J. Fíla, D. Honys, J. Fílla, D. Honys, Enrichment techniques employed in phosphoproteomics, Amino Acids 43 (2012) 1025–1047, doi:10.1007/s00726-011-1111-z.
[16] P.A. Connor, A.J. McQuillan, Phosphate adsorption onto TiO2 from aqueous solutions: an in situ internal reflection infrared spectroscopic study, Langmuir 15 (1999) 2916–2921, doi:10.1021/la980894p.
[17] N.C. Jeong, J.S. Lee, E.L. Tae, Y.J. Lee, K.B. Yoon, Acidity scale for metal oxides and Sanderson's electronegativities of lanthanide elements, Angew. Chem. Int. Ed. 47 (2008) 10128–10132, doi:10.1002/anie.200803837.
[18] K.D. Dobson, A.J. McQuillan, In situ infrared spectroscopic analysis of the adsorption of aromatic carboxylic acids to TiO2, ZrO2, Al2O3, and Ta2O5 from aqueous solutions, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 56 (2000) 557–565, doi:10.1016/S1386-1425(99)00154-7.
[19] K.D. Dobson, A.J. McQuillan, In situ infrared spectroscopic analysis of the adsorption of aliphatic carboxylic acids to TiO2, ZrO2,Al2 O3, and Ta2 O5 from aqueous solutions, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 55 (1999) 1395–1405, doi:10.1016/S1386-1425(99)00154-7.
[20] S.S. Jensen, M.R. Larsen, Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques, Rapid Commun. Mass Spectrom. 21 (2007) 3635–3645, doi:10.1002/rcm.3254.
[21] A. Leitner, M. Sturm, W. Lindner, Tools for analyzing the phosphoproteome and other phosphorylated biomolecules: a review, Anal. Chim. Acta 703 (2011) 19–30, doi:10.1016/j.aca.2011.07.012.
[22] T.S. Kephart, P.K. Dasgupta, Hot eluent capillary liquid chromatography using zirconia and titania based stationary phases, Anal. Chim. Acta 414 (2000) 71–78, doi:10.1016/S0003-2670(00)00821-7.
[23] A. Leitner, Phosphopeptide enrichment using metal oxide affinity chromatography, TrAC Trends Anal. Chem. (Reference Ed.) 29 (2010) 177–185, doi:10.1016/j.trac.2009.08.007.
[24] M.R. Larsen, T.E. Thingholm, O.N. Jensen, P. Roepstorff, T.J.D. Jørgensen, Highly selective enrichment of phosphorylated peptides from peptide mixtures

using titanium dioxide microcolumns, Mol. Cell. Proteom. 4 (2005) 873–886, doi:10.1074/mcp.T500007-MCP200.
[25] T.E. Thingholm, T.J.D. Jørgensen, O.N. Jensen, M.R. Larsen, Highly selective enrichment of phosphorylated peptides using titanium dioxide, Nat. Protoc. 1 (2006) 1929–1935, doi:10.1038/nprot.2006.185.
[26] M. Mazanek, E. Roitinger, O. Hudecz, J.R.A. Hutchins, B. Hegemann, G. Mitulović, T. Taus, C. Stingl, J.M. Peters, K. Mechtler, A new acid mix enhances phosphopeptide enrichment on titanium- and zirconium dioxide for mapping of phosphorylation sites on protein complexes, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 878 (2010) 515–524, doi:10.1016/j.jchromb.2009.12.017.
[27] N. Sugiyama, T. Masuda, K. Shinoda, A. Nakamura, M. Tomita, Y. Ishihama, Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications, Mol. Cell. Proteom. 6 (2007) 1103–1109, doi:10.1074/mcp.T600060-MCP200.
[28] A. Gonzálvez, B. Preinerstorfer, W. Lindner, Selective enrichment of phosphatidylcholines from food and biological matrices using metal oxides as solid-phase extraction materials prior to analysis by HPLC-ESI-MS/MS, Anal. Bioanal. Chem. 396 (2010) 2965–2975, doi:10.1007/s00216-010-3527-9.
[29] P.K. Kanaujia, D. Pardasani, V. Tak, A.K. Purohit, D.K. Dubey, Selective enrichment of the degradation products of organophosphorus nerve agents by zirconia based solid-phase extraction, J. Chromatogr. A 1218 (2011) 6612–6620, doi:10.1016/j.chroma.2011.07.091.
[30] L. Qiao, C. Roussel, J. Wan, P. Yang, H.H. Girault, B. Liu, Specific on-plate enrichment of phosphorylated peptides for direct MALDI-TOF MS analysis, J. Proteome Res. 6 (2007) 4763–4769, doi:10.1021/pr0705284.
[31] Y. Kyono, N. Sugiyama, K. Imami, M. Tomita, Y. Ishihama, Successive and selective release of phosphorylated peptides captured by hydroxy acid-modified metal oxide chromatography, J. Proteome Res. 7 (2008) 4585–4593, doi:10.1021/pr800305y.
[32] H.C. Hsieh, C. Sheu, F.K. Shi, D.T. Li, Development of a titanium dioxide nanoparticle pipette-tip for the selective enrichment of phosphorylated peptides, J. Chromatogr. A 1165 (2007) 128–135, doi:10.1016/j.chroma.2007.08.012.
[33] M. Wakabayashi, Y. Kyono, N. Sugiyama, Y. Ishihama, Extended coverage of singly and multiply phosphorylated peptides from a single titanium dioxide microcolumn, Anal. Chem. 87 (2015) 10213–10221, doi:10.1021/acs.analchem.5b01216.
[34] E.S. Simon, M. Young, A. Chan, Z.Q. Bao, P.C. Andrews, Improved enrichment strategies for phosphorylated peptides on titanium dioxide using methyl esterification and pH gradient elution, Anal. Biochem. 377 (2008) 234–242, doi:10.1016/j.ab.2008.03.024.
[35] Y. Sekiguchi, N. Mitsuhashi, Y. Inoue, H. Yagisawa, T. Mimura, Analysis of sugar phosphates in plants by ion chromatography on a titanium dioxide column with pulsed amperometric detection, J. Chromatogr. A 1039 (2004) 71–76, doi:10.1016/j.chroma.2004.02.015.
[36] Y. Sekiguchi, N. Mitsuhashi, T. Kokaji, H. Miyakoda, T. Mimura, Development of a comprehensive analytical method for phosphate metabolites in plants by ion chromatography coupled with tandem mass spectrometry, J. Chromatogr. A 1085 (2005) 131–136, doi:10.1016/j.chroma.2005.01.098.
[37] H.F.N. Kvitvang, P. Bruheim, Fast filtration sampling protocol for mammalian suspension cells tailored for phosphometabolome profiling by capillary ion chromatography - tandem mass spectrometry, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 998–999 (2015) 45–49, doi:10.1016/j.jchromb.2015.06.018.
[38] T.J. Causon, S. Hann, Review of sample preparation strategies for MS-based metabolomic studies in industrial biotechnology, Anal. Chim. Acta 938 (2016) 18–32, doi:10.1016/j.aca.2016.07.033.
[39] O. Vielhauer, M. Zakhartsev, T. Horn, R. Takors, M. Reuss, Simplified absolute metabolite quantification by gas chromatography–isotope dilution mass spectrometry on the basis of commercially available source material, J. Chromatogr. B 879 (2011) 3859–3870, doi:10.1016/J.JCHROMB.2011.10.036.
[40] H.F.N. Kvitvang, K.A. Kristiansen, P. Bruheim, Assessment of capillary anion exchange ion chromatography tandem mass spectrometry for the quantitative profiling of the phosphometabolome and organic acids in biological extracts, J. Chromatogr. A 1370 (2014) 70–79, doi:10.1016/j.chroma.2014.10.029.
[41] A.B. Canelas, A. Ten Pierick, C. Ras, R.M. Seifar, J.C. Van Dam, W.M. Van Gulik, J.J. Heijnen, Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics, Anal. Chem. 81 (2009) 7379–7389, doi:10.1021/ac900999t.
[42] K. Ortmayr, G. Koellensperger, S. Hann, Complementing reversed-phase selectivity with porous graphitized carbon to increase the metabolome coverage in an on-line two- dimensional LC-MS setup for metabolomics, R. Soc. Chem. (2015), doi:10.1039/c5an00206k.
[43] D.B. Chu, K. Klavins, G. Koellensperger, S. Hann, Speciation analysis of sugar phosphates via anion exchange chromatography combined with inductively coupled plasma dynamic reaction cell mass spectrometry – optimization for the analysis of yeast cell extracts, J. Anal. At. Spectrom. 29 (2014) 915, doi:10.1039/c4ja00043a.
[44] J.A. Blackwell, P.W. Carr, Development of an Eluotropic Series for the Chromatography of Lewis Bases on Zirconium Oxide, American Chemical Society, 1992, doi:10.1021/ac00032a008.
[45] J. Yu, Z. El Rassi, Chromatographic properties of zirconia-based stationary phases for ion exchange chromatography having surface bound cationic functions, J. High Resolut. Chromatogr. 17 (1994) 705–712. https://onlinelibrary.wiley.com/doi/pdf/10.1002/jhrc.1240171005. [accessed 27 November 2018].
[46] T. Micksch, N. Liebelt, D. Scharnweber, B. Schwenzer, Investigation of the peptide adsorption on ZrO_2, TiZr, and TiO_2 surfaces as a method for surface modification, ACS Appl. Mater. Interfaces 6 (2014) 7408–7416, doi:10.1021/am500823m.

3. Summary and Conclusion

Analytical metabolomics can help to reveal a lot of useful information about metabolism in biological systems using both targeted and non-targeted approaches. Together with proteomics, transcriptomics and genomics, mass spectrometry-based metabolomics studies help to provide new information relevant to biological questions in many organisms. Several analytical challenges still exist for metabolomics studies due to the high diversity of the metabolome, the wide dynamic range of the metabolites´ concentration, drawbacks of each analytical platform, and the lack of harmonized validation guidelines and reports. While there have been substantial advances already made in the field, the work in this thesis highlights important analytical challenges and suggests methods to meet the requirements of current and future metabolomics studies.

In the case of non-targeted metabolomics, the availability of high-resolution mass spectrometry instrumentation used in combination with the broad coverage of reversed-phase liquid chromatography separation is identified to remain as a key pillar of broad metabolomics studies. The type of chromatographic stationary phases for extending metabolome coverage in context of RPLC combined with high resolution mass spectrometry was shown to yield only minor differences in overall method performance for non-targeted metabolomics. This is certainly, in part, due to the overall improvements in column technology and fundamental limitation for coverage of polar metabolites shared by columns with similar chemical properties. Improvements in retention capacity and selectivity for key isomers rather than chromatographic efficiency are identified as key areas of development for improving coverage with RPLC-HRMS.

Validated sample preparation workflows addressing key metabolite classes were also demonstrated to be essential in cases where detailed, quantitative data are required. Twelve phosphorylated metabolites present in low abundance in samples were successfully enriched in a validated workflow based on metal oxide affinity solid phase extraction. This sensitive and selective cleanup/enrichment method developed in this work dealt not only with glycolysis and pentose phosphate pathways, but also provides a basis for method development requiring selective enrichment of other phosphorylated metabolites.

4. Curriculum Vitae

europass Curriculum Vitae Si Hung LE

PERSONAL INFORMATION LE, Si Hung

Home address: 271A/14B Thinh Quang, Dong Da, Hanoi, Vietnam
Current address: Ferrogasse 68-70, 1180, Vienna, Austria
0436608364468
le.si-hung@boku.ac.at

Sex Male | Date of birth 14/11/1988 | Nationality Vietnamese

WORK EXPERIENCE

March 2016 – Now	**PhD in Analytical chemistry** Vienna, Austria ▪ Research on non-targeted metabolomics with accurate mass spectrometry in food- and biotechnology **Business or Sector:** University of Natural Resources and Life Sciences, Vienna
October 2013 – July 2015	**Researcher** Hanoi, Vietnam ▪ Research on manufacturing portable optical measurement device and test kit for rapid analysis of trace amounts of ammonium, nitrite and nitrate in the field **Business or Sector:** VNU University of Science
September 2013 – May 2014	**Researcher** Hanoi, Vietnam ▪ Analyses and assessment sources, exchange and transport of heavy metals and nutrients in water and sediment environments in downstream of Cau River basin: a case study in Hai Duong province **Business or Sector:** VNU University of Science
August 2013 – September 2013	**Internship** Kofu, Japan ▪ Determination source of nutrients (N and P) pollution in downstream of Cau River basin: a case study in Hai Duong province using dual-isotope analysis of Nitrate and water. ▪ Collecting sample, determining source of nutrients (N and P) pollution **Business or Sector:** ICRE, University of Yamanashi, Kofu, Japan
August 2012 – September 2012	**Researcher** Hanoi, Vietnam ▪ Analyses and assessment heavy metals pollution in Mekong river: A case study in Vientiane, Laos ▪ Collecting, measuring samples **Business or Sector:** VNU University of Science
March 2011 – February 2016	**Researcher** Hanoi, Vietnam ▪ Research on determination trace of heavy metals using ICP-MS and AAS ▪ Lecture on Analytical chemistry using advanced analytical techniques such as ICP-MS, F-AAS, etc **Business or Sector:** VNU University of Science

EDUCATION AND TRAINING

August 2018	Poster presentation at Austrian Proteomics and Metabolomics Research Symposium (APMRS 2018), Vienna, August 29 - 31, 2018 Title: *"Metal oxide-based enrichment for addressing sensitivity and selectivity issues in the quantitative analysis of phosphorylated primary metabolites"*
June 2018	Poster presentation at 66th Conference on Mass Spectrometry and Allied Topics (ASMS) 2018, San Diego, CA, USA, June 3-7, 2018 Title: *"Sensitive and quantitative analysis of phosphorylated primary metabolites using selective metal oxide enrichment and GC-MS/MS"*
September 2017	Oral presentation at 23rd International Symposium on Separation Sciences (ISSS 2017), Vienna, AUSTRIA, Sept 19 - 22, 2017 Title: *"Comparison of fully-wettable RPLC stationary phases for LC-MS based cellular metabolomics"*
May 2017	Oral presentation at 13th ASAC JunganalytikerInnenforum 2017, Vienna, May 12-13, 2017 Title: *"Assessment of Fully-Wettable RPLC columns for LC-MS-based Cellular Metabolomics"*
March 2016 - Now	PhD – Thesis title: *"Non-targeted metabolomics with accurate mass spectrometry in food- and biotechnology"* University of Natural Resources and Life Sciences (BOKU), Vienna Supervisor: Assoc. Prof. Dr. Stephan Hann
April 2015	MSc – Thesis title: *"Assessment of sources, exchange and transport of nutrients in water environments in downstream of Cau River basin"* VNU University of Science, Hanoi, Vietnam (chemvnu.edu.vn)
February 2014	Oral presentation at *"International Young Researchers Workshop on River Basin Environment and Management"* Title: *" Analyses and assessment sources of heavy metals and nutrients in water and sediment environments in downstream of Cau River basin: a case study in Hai Duong province"* AIT University, Bangkok, Thailand
August 2013 – September 2013	Completed the ICRE Internship Topic: *"Light stable isotope analysis for addressing contamination sources"* University of Yamanashi, Kofu, Japan (http://www2.yamanashi.ac.jp)
June 2011	Bachelor of Science in Chemistry VNU University of Science, Hanoi, Vietnam (chemvnu.edu.vn) - Trace metal analysis - Environmental chemistry - Material chemistry

PERSONAL SKILLS

Mother tongue(s): Vietnamese

Other language(s)

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken interaction	Spoken production	Writing
English	C2	C1	C2	C2	B2

TOEFL iBT / Test date: July 24st 2010 / Registration No: 09903097
Listening: 22, Reading: 25, Speaking: 15, Writing: 21, Total: 83

Levels: A1/2: Basic user - B1/2: Independent user - C1/2 Proficient user
Common European Framework of Reference for Languages

Communication skills

- Good communication skills gained through my teaching and travelling
- Good teamwork skill gained through my research work

Job-related skills

- Analytical chemistry:
  - *Non-targeted metabolomics with accurate and tandem mass spectrometry (LC-QTOF, GC-Q-TOF, LC-QQQ, GC-QQQ) in food- and biotechnology analysis (e.g. analysis of microbial cell factory extracts). Additional experience with development of new ion mobility-mass spectrometry methods for metabolomics.*
  - *Analysis of light isotopes.*
  - *Analysis and evaluation of toxic compounds in different kinds of water and sediment samples using F-AAS, ICP-MS.*
- Environmental chemistry: synthesis of various materials for waste water and ground water treatments.

Computer skills

- Good skills in using Microsoft Office™ tools, Minitab and some dedicated software for Analytical Chemistry and related topics (e.g. Agilent MassHunter suite: Profinder, Mass Profiler Professional, Qualitative Analysis, Quantitative Analysis, IM Browser), Skyline.

Other skills

- Photography
- Travelling
- Martial arts

ADDITIONAL INFORMATION

Publications

1. Si-Hung, L; Troyer, C; Causon, TJ; Hann, S, (2019), "Sensitive and quantitative analysis of phosphorylated primary metabolites using selective metal oxide enrichment and GC-MS/MS", doi:10.1016/j.talanta.2019.120147, Talanta.
2. Causon, TJ; Si-Hung, L; Newton, K; Kurulugama, R; Fjeldsted, J; Hann, S, (2019) "Fundamental study of ion trapping and multiplexing using drift tube-ion mobility time-of-flight mass spectrometry for non-targeted metabolomics.", doi:10.1007/s00216-019-02021-8, Analytical and Bioanalytical Chemistry.
3. Si-Hung, L; Causon, TJ; Hann, S, (2017), "Comparison of fully wettable RPLC stationary phases for LC-MS-based cellular metabolomics.", 38(18): 2287–2295, doi: 10.1002/elps.201700157, ELECTROPHORESIS.
4. Ta, T.T; Le, S.H; Trinh, H.Q; Luu, T.N; Trinh, A.D, (2016), "Interpretation of anthropogenic impacts (agriculture and urbanization) on tropical deltaic river network through the spatio-temporal variation of stable (N, O) isotopes of NO(-)3,. 52(4-5):487-97. doi: 10.1080/10256016.2016.1142987, Isotopes in Environmental and Health Studies.
5. Pham, T.N.M; Nguyen, V.T; Trinh, T.T; Nguyen, K.H; Hoang, Q.A; Tran, M.T; Le, S.H; Dao, T.N; Vu, D.N; Nguyen, T.M.H; Nguyen, T.A.H; Duong, H.A; Nguyen, H.M; Tu, B.M, (2015), "Distribution, accumulation profile, and risk assessment of polybrominated diphenyl ethers in sediment from lake and river systems in Hanoi Metropolitan Area, Vietnam", doi:10.1007/s11356-015-5235-7, Environmental Science and Pollution Research.
6. Le, S.H; Ta, T.T; Nishida, K, "Analyses and assessment sources of heavy metals and nutrients in water and sediment environments in downstream of Cau River basin: a case study in Hai Duong province, Proceeding of International Young Researchers Workshop on River Basin Environment and Management, February 2014, Bangkok, Thailand.
7. Ta, T.T; Le, S.H; Do, T.N, (2014), "Isotope analysis of $\delta^{15}N$ and $\delta^{18}O$ for the assessment of source, exchange and transport of nutrients in water Environments in downstream of the Cau river basin: A case study in Hai Duong province", p.260, VNU Journal of Science.
8. Pham, T.N.M; Le, S.H; Phan, T.H, (2013), "Study on the photocatalytic activities of $ZnO_{1-x}S_x$ powder under visible light", Vol 51 2C, p.1008-1013, Vietnam Journal of Chemistry.

References

1. **Assoc. Prof. Dr. Stephan Hann**
   Institute of Analytical Chemistry, University of Natural Resources and Life Sciences,
   18 Muthgasse, Vienna, Austria
   Email: stephan.hann@boku.ac.at
   Phone: +43 1 47654-77186

2. **Ass. Prof. B.Sc. Ph.D. Tim Causon**
   Institute of Analytical Chemistry, University of Natural Resources and Life Sciences,
   18 Muthgasse, Vienna, Austria
   Email: tim.causon@boku.ac.at
   Phone: +43 1 47654-77187

3. **Assoc. Prof. Dr. Ta Thi Thao**
   Head of the Department of Analytical Chemistry, Faculty of Chemistry, University of Science, Hanoi National University, VNU, 19 Le Thanh Tong, Hanoi, Vietnam
   Email: thaott73@gmail.com

4. **Dr. Chu Dinh Binh**
   Department of Analytical Chemistry, School of Chemical Engineering, Hanoi University of Science and Technology, 1 Dai Co Viet, Hanoi, Vietnam
   Email: binh.chudinh@hust.edu.vn