Metabolomics relative quantitation with mass spectrometry using chemical derivatization and isotope labeling

Grace O’Maille a, Eden P. Go a, Linh Hoang a, Elizabeth J. Want a, Colin Smith a, Paul O’Maille b, Anders Nordström a, Hirotoshi Morita a, Chuan Qin a, Wilasinee Uritboonthai a, Junefredo Apon a, Richard Moore c, James Garrett c and Gary Siuzdak a,∗∗

a The Center for Mass Spectrometry and Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA
b Jack H. Skirball Center for Chemical Biology and Proteomics, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA
c Becton Dickinson, Sparks, MD 21152, USA

Abstract. Comprehensive detection and quantitation of metabolites from a biological source constitute the major challenges of current metabolomics research. Two chemical derivatization methodologies, butylation and amination, were applied to human serum for ionization enhancement of a broad spectrum of metabolite classes, including steroids and amino acids. LC-ESI-MS analysis of the derivatized serum samples provided a significant signal elevation across the total ion chromatogram to over a 100-fold increase in ionization efficiency. It was also demonstrated that derivatization combined with isotopically labeled reagents facilitated the relative quantitation of derivatized metabolites from individual as well as pooled samples.

Keywords: Metabolomics, human plasma, isotope labeling, quantitation, mass spectrometry

1. Introduction

The challenge of generating a global metabolite profile for discerning biochemical pathways, interactions or identifying markers [1–8] comes from the diversity of endogenous metabolites. Unlike proteins and nucleic acids, metabolites include a plethora of molecular species such as simple amino acids, nucleic acids, steroids, carbohydrates and lipids, of which some are difficult to analyze with current electrospray ionization mass spectrometry methods [9]. Among the requirements for a comprehensive profiling approach is that it should impose little analyte discrimination and allow for quantitative analysis, both of which are currently impossible to achieve. For example, gas chromatography mass spectrometry (GC-MS) is still the most extensively used for metabolite profiling [10–12], yet is limited to nonpolar and volatile molecules. A partial solution to this limitation is the use of chemical derivatization prior to GC-MS analysis [13]. Standard GC-MS derivatization schemes often involve the replacement of labile
hydrogens with trimethylsilyl (TMS) groups, introducing an additional 72 Da for each hydrogen replaced. Unfortunately, the respective increase in mass through derivatization often places the modified molecule above the typical range useful for GC-MS [14], therefore, limiting the utility of GC-MS for endogenous metabolite profiling.

Liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS) offers a viable alternative to GC-MS for both polar and higher molecular weight metabolites [15]. LC-ESI-MS for metabolite profiling investigations was initially explored by Tolstikov and Fiehn [16], and further investigated with micro-LC-ESI-MS [6,16–19], nano-ESI-MS [5,20–22], UPLC ESI-MS (ultra high pressure liquid chromatography) [23,24], as well as ESI alone with direct infusion [25–27]. However, a wide range of nonpolar compounds, compounds with only weakly acidic or basic groups and a wide range of natural compounds such as terpenes, sugars, alcohols, aromatic compounds and vitamins, which are not detected without derivatization or display poor sensitivity with ESI [28–30]. To address this shortcoming, several strategies for chemical derivatization of analytes prior to ESI-MS to improve the sensitivity of analytes have been developed by a number of research groups [28–45]. Recently, the use of chemical derivatization with isotope labeling has been reported to measure the quantitative metabolite differences between metabolome samples [43,45].

The rationale for implementing derivatization as part of an LC-ESI-MS metabolomics effort is to enable more comprehensive coverage of the metabolome. Another important aspect of a metabolomics investigation is to be able to perform relative metabolite quantitation between the different sets of samples investigated. A simple approach presented by Wang et al. demonstrated that without the use of isotopic labeling, single value normalization and counting generated a median coefficient of variation of 23.8% [22]. In another study, analytical reproducibility from 2 to 12% for selected metabolites was achieved by Fiehn et al. through the addition of stable isotope labeled standards prior to analysis [11]. The same approach was also used by Jonsson et al. for normalization of raw data in order to compensate for differences in recovery and injection volumes [46]. While stable isotope techniques have proven to be reliable, they suffer from the requirement of a labeled standard. Additionally, stable isotope standards are very expensive, imposing further constraints on this approach. An alternative approach is to employ relative quantitation as developed by Mashego et al. [47] in monitoring the metabolite level changes in yeast. In their study, cells were grown in 13C labeled media, thereby obtaining a stable isotope labeled metabolome to which an aliquot could be added, and co-extracted with yeast cells grown in 12C media. With this strategy, stable isotope labeled analogues of 80% of the glycolytic metabolites were observed. Similar strategies have been employed in subsequent studies with yeast [48,49] and plant cell culture [50].

While growing the system in labeled media is perhaps one of the best approaches to monitor metabolic changes, unfortunately, it is not always possible. For quantitative proteomics, the use of isotope-coded affinity tags (ICAT) [51] has gained popularity. The ICAT approach involves labeling one of the two samples with an affinity tagged molecule and the other sample with a stable isotope labeled analogue of the same tag. The samples are subsequently pooled and subject to proteolysis and affinity purification before LC-ESI-MS/MS analysis. The relative intensity of any labeled peptide to its stable isotope labeled counterpart will reveal the relative concentration of the particular protein.

In this manuscript, we present the development and application of derivatization with unlabeled and stable isotope labeled reagents for LC-ESI-MS analysis of human plasma metabolite extracts. Among the objectives are: (1) to broaden the detection of metabolites through the use of derivatization, thereby enhancing the predictive capacity of the data [38], and (2) to relatively quantitate without the addition of stable isotope labeled standards [39,40].
Since carboxylic and hydroxyl groups are common functional groups in metabolites, we chose to proceed with two well established derivatization procedures, namely butylation [42] for carboxyl groups and amination [44] for hydroxyl groups, in this proof of concept study. To achieve relative quantitation, isotopically labeled forms of the reagents were synthesized. This derivatization strategy was also applied to the metabolomics study of a set of clinical samples of sepsis, which is one of the leading causes of death in the intensive care unit [52]. The LC-ESI-MS data of plasma obtained from patients with sepsis were compared to the data of a control set of plasma obtained from patients with systemic inflammatory response syndrome (SIRS) [52–54]. The LC-ESI-MS data were obtained for underivatized as well as derivatized plasma metabolites, and were subsequently analyzed by XCMS, an in-house LC-MS data analysis software for non-linear retention time correction and alignment of LC-ESI-MS data [55].

2. Experimental section

2.1. Reagents and chemicals

Phenylalanine, citrulline, oxidized glutathione, cholesterol, succinic anhydride, dimethylaminoethanol, 1,1′-carbonyldiimidazole, and butanol were obtained from Sigma Chemical Company (St. Louis, MO, USA). $^{13}$C$_4$ labeled succinic acid and D$_9$ labeled butanol were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA, USA). All solvents for the LC-ESI-MS studies were HPLC grade (J.T. Baker, Phillipsburg, NJ, USA). All other solvents were Optima grade from Fisher Scientific (Los Angeles, CA, USA).

2.2. Human serum and plasma samples

Human serum (male, H-1388) purchased from Sigma Chemical Company (St. Louis, MO, USA) was used for method development and ionization enhancement studies. For clinical analysis, human plasma collected from intensive care unit patients diagnosed with sepsis ($n=27$) and Systemic Inflammatory Response Syndrome (SIRS) ($n=34$) were assayed. Detailed information about the gender, age, race, admission cause, medical history, and current medical conditions as well as current medications were collected. Plasma samples used in the current study were collected on the day of diagnosis of the sepsis or SIRS disease state and stored at $-80^\circ$C until analysis. Metabolite extraction, derivatization and LC-ESI-MS analysis for plasma samples were performed consecutively. The total time from the point of metabolite extraction to obtaining XCMS analyzed LC-ESI-MS data for each sample was about 8 h. Additional details for each procedure are provided in the following sections.

2.3. Serum/plasma metabolite extraction

A methanol protein precipitation extraction, as described by Want et al. [19], was used for all serum and plasma metabolite extractions. The procedure described here has been tested and was found suitable for both serum and plasma preparations; no significant differences were observed. Briefly, 100 µl of ice-cold methanol was added to a starting volume of 50 µl serum or plasma. The mixture was vortexed immediately for 20 s and incubated at $-20^\circ$C for 20 min, followed by centrifugation at 12,000 rpm (13,400 × g) for 10 min. The supernatant was retained and dried in a SpeedVac. The dried serum/plasma metabolite extract was reconstituted in 50 µl H$_2$O/acetonitrile (95/5, v/v) prior to LC-ESI-MS analysis. The same sample extraction method described here was used for the preparation of underivatized
serum/plasma. For derivatized samples, dried serum/plasma metabolite extracts were further derivatized with the respective reagents as described below.

2.4. Synthesis of amination reagent

The amination reagent used in this study was adapted from the work of Johnson et al. [44]. The synthesis of the amination reagent was slightly modified. Succinic anhydride (1.0 mmol) and dimethylaminoethanol (1.2 mmol) were mixed and refluxed in 20 ml acetone for 2.5 h. The resulting reaction mixture was then cooled to room temperature and filtered. The white precipitate was washed three times with 20 ml of cold acetone and subsequently dried to give monodimethylaminoethylsuccinate (MDMAES). MDMAES was found to be stable for up to 3 months when stored under refrigeration (at ~4°C) and away from light and moisture. The final amination reagent was prepared by first mixing MDMAES (0.1 mmol) and 1,1′-carbonyldiimidazole (0.1 mmol) in dichloromethane (0.2 ml) in a glass container and vortexed for 10 s at 2 min intervals until all solids were dissolved. An additional 0.8 ml of dichloromethane was then added to prepare a 100 mM solution of the amination reagent [44]. The final amination reagent must be kept away from moisture since it is highly reactive. Therefore, this reagent must be made fresh daily by mixing MDMAES and 1,1′-carbonyldiimidazole prior to use. NMR and high accuracy ESI-TOF mass spectral data analysis were performed to confirm the chemical structures of MDMAES and the amination reagent.

2.5. Synthesis of isotopically labeled amination reagent

In order to prepare the isotopically labeled amination reagent, 13C4 labeled succinic anhydride was used as a starting material and then follow the amination reagent synthesis described above. The synthesis of the 13C4 labeled succinic anhydride was adapted from the work of Zhang and Regnier [56]. Briefly, acetyl chloride and 13C4 labeled succinic acid were mixed at a molar ratio of 300 to 1 and refluxed for 1.5 h. The clear solution was cooled and kept at 0°C overnight. The crystals formed were collected and washed with ethyl ether and dried under vacuum to give 13C4 labeled succinic anhydride. The 13C4 labeled MDMAES and subsequently 13C4 labeled amination reagent were synthesized using the 13C4 labeled succinic anhydride. NMR and high accuracy ESI-TOF mass spectral data analysis were performed to confirm the chemical structures and mass of the resulting 13C4 labeled amination reagent.

2.6. Amination derivatization procedure

A 100 µl volume of a freshly prepared 100 mM amination reagent solution was added to each of the dried serum/plasma metabolite extract in a glass vial with a Teflon-lined cap. The vial was tightly capped and sonicated for 1 min prior to incubation at 65°C for 4 h. After incubation, the residual solvent was dried in a SpeedVac. The aminated metabolite residue was reconstituted in 50 µl H2O/acetonitrile (95/5, v/v) and sonicated for 1 min prior to LC-ESI-MS analysis. For isotopically labeled derivatization of human serum/plasma, a 13C4 labeled amination reagent solution was used.

2.7. Butylation derivatization procedure

The butylation reagent solution was prepared by mixing butanol and concentrated HCl in 3–1 ratio (v/v). A 100 µl volume of the butylation reagent solution was added to each of the dried serum/plasma
metabolite extract [42]. The reaction vial was tightly capped and vortexed for 30 s, followed by incubation at 65°C for 4 h [57]. The reaction mixture was then dried in a SpeedVac. The butylated metabolite residue was reconstituted in 50 µl H2O/acetonitrile (95/5, v/v) and sonicated for 1 min prior to LC-ESI-MS analysis. For isotopically labeled butylation of human serum/plasma, deuterated (D₉) butanol was used.

2.8. Chromatography and MS data acquisition

Reverse phase (RP) chromatography and mass detection of the serum metabolite extracts was performed on an Agilent 1100 LC/MSD SL system. Due to possible ionization drifts of the electrospray, each extracted sample was analyzed in triplicate, where the three injections were arranged in a random order among the run list of other serum/plasma samples to assess re-injection reproducibility. Additionally, a blank sample (consisted of the injection of 10 µl methanol) was placed between each serum/plasma sample in the run sequence to minimize potential carry-over. For each study sample run, 10 µl of serum metabolite extract was injected onto a C18 column (Symmetry® Column, 2.1 × 100 mm, 3.5 µm; Waters, Milford, MA, USA) and eluted at a flow rate of 250 µl/min under gradient conditions of 5% B to 90% B over 60 min. Mobile phase A consisted of water/acetonitrile/formic acid (95/5/0.1 (v/v/v)), and B consisted of acetonitrile/formic acid (100/0.1 (v/v)). Mass spectral data from 100–1000 m/z were collected in the positive ionization mode.

In order to minimize the contamination of the C18 column during the chromatographic separation of aminated samples, two additional steps were employed: (1) a 10-min holding at 5% B was added at the beginning of the gradient. (2) A precolumn and column switching step was added. All aminated samples were initially loaded onto a precolumn (Zorbax SB-C18, Rapid Resolution Cartridge Column, 2.1 × 30 mm, 3.5 µm; Agilent Technologies, Palo Alto, CA, USA). During the initial 10-min holding at 5% B, the solvent flow after the precolumn was directed to waste to flush out the amination reagent. At the 10-min time point of the gradient, a column switching procedure was activated so that the solvent flow from the precolumn was directed to the C18 separation column.

The sample throughput was about 14–16 samples per day for a 60- to 70-min LC separation and a 20-min blank sample run between sample injections. For each plasma sample derivatization procedure, the time between the collection of plasma to the complete acquisition of the LC-ESI-MS data was estimated to be about 10 h.

MS/MS fragmentation data for selected ions of interest were acquired using the data dependent MS/MS functionality of the MicroMass Q-TOF Micro instrument (Waters, Millford, MA). For each precursor ion selected, collision energy ranging from 25 to 45 eV was applied to achieve the optimal fragmentation pattern.

2.9. LC-ESI-MS data analysis

All LC-ESI-MS data were processed using XCMS [55], an open source software (http://metlin.scripps.edu/download) for metabolomics studies. The name XCMS originated from the design of a software capable of analyzing data from multiple types of chromatographically (GC and LC) separated samples giving the XC part of the name, and the MS part of the name originated from mass spectrometry data. XCMS was used to perform peak picking, peak integration, non-linear retention time correction and matching of peaks between samples. These XCMS functionalities were designed to compare the same
components (defined as “metabolite features”, or simply “features”, with a unique set of $m/z$ and retention time values) in all samples despite any retention time shifts. A data table and an EIC (extracted ion chromatogram) folder were generated for each XCMS analysis.

Each XCMS generated data table consisted of column headings showing $m/z$, retention time, and sample names (for displaying the intensity area values for each feature in all samples) where values corresponding to each heading were displayed in consecutive rows. A typical XCMS data table generated from the LC-ESI-MS analysis of underivatized human serum metabolite extract consisted of 2000 rows, which corresponded to the detection of 2000 metabolite features, with each feature representing a unique set of $m/z$ and retention time. In a previous study, Want and colleagues collected LC-ESI-MS data from human plasma metabolites extracted using various organic solvents and acids [19]. The LC-ESI-MS data analyzed by XCMS were presented in table formats in the supporting material of that manuscript.

In the previous study [19] as well as in the current study, isotopic ions, adducts and fragments as well as multiple peaks arising from multiple charged ions were treated as separate metabolite features; therefore, we noted that one molecular species may give rise to multiple features. For the purpose of further evaluations, the XCMS data tables were generated in tab-delimited format and the data were exported to other statistical software of choice. The statistical analysis section described in the following section made use of these metabolite features data output in the XCMS data tables.

In addition to the data table, a second XCMS generated output consisted of a folder containing images of superimposed aligned EICs (extracted ion chromatograms) of individual features shown as intensity versus retention time plots. This visual presentation enabled the examination of individual feature peak shape as well as a visual examination and comparison of the peak heights of the same metabolite feature detected in various samples. Additionally, different colors were manually assigned by the user to denote specific groups or sets of samples to be compared, which greatly enhanced the effectiveness of the manual analysis.

2.10. Statistical analysis

Due to the large number of metabolite features from the large number of samples detected by LC-ESI-MS and captured by XCMS, further statistical analyses were performed by exporting the XCMS data table outputs into MATLAB for more effective evaluation of the data from the clinical samples. The XCMS outputs consisted of the $m/z$, retention time and peak intensity areas for all metabolite features detected in each sample. Due to the XCMS retention time correction function, the metabolite features with the same $m/z$ eluted at the same retention time can be compared between samples.

3. Results and discussion

3.1. Enhancement of ionization efficiency

One of the goals of this study was to expand the utility of LC-ESI-MS to non-targeted global metabolomics studies by detecting low abundance and poorly ionizable metabolites. These objectives were accomplished through chemical derivatization of human serum metabolites by employing derivatization techniques, namely amination [44] and butylation [42].

Results of the LC-ESI-MS experiments clearly showed that both derivatization procedures increased the number of components, as observed by the marked increase of chromatographic peaks in the total ion chromatograms (TIC). Representative TIC of the underivatized, aminated and butylated serum
metabolites were compared in parallel as shown in Figs 1 and 2. Derivatization through amination introduced a covalently bound positively charged tertiary amine group to the derivatized compounds, which enhanced ionization efficiency either in solution or during the ESI process, and enabled detection. To verify whether the amination procedure was successful, MS/MS analysis was performed on the peak that eluted at 41 min. This peak, indicated by an asterisk symbol in Fig. 1, corresponded to the aminated cholesterol with $m/z$ 558. Fragmentation data of this peak showed a neutral loss of 189 Da, giving the prominent product ion at $m/z$ 369, which corresponded to the loss of a neutral acid at the ester linkage conjugating the amination tag to cholesterol [44] at the C-3 position (Fig. 1).

In addition to the detection of aminated cholesterol, aminated phospholipid ions were identified, and these were detected as the prominent peaks eluting between 10 and 15 min of the aminated TIC in (Fig. 1). For example, the underivatized phospholipid with $m/z$ 496 (lysophosphatidyl choline with a 16-carbon fatty acid chain [19,38]), typically eluting at 30 min, was detected at $m/z$ 668 in its aminated form, eluting at 16 min in the aminated sample. In the case of the 16-carbon phospholipid, the hydroxyl
group on the glycerol backbone was aminated, resulting in the aminated phospholipid with an increase of $m/z$ 172 corresponding to the amination tag conjugated through an ester linkage. Due to the incomplete amination, multiple ion peaks were detected consisting of the underivatized and aminated phospholipid ions.

An enhanced detection of metabolites was also observed in the butylated sample, where negatively charged carboxylic acids were derivatized to form neutral butyl esters. The enhanced detection upon butylation occurred mainly in the early eluting region (1–15 min) of the reversed phase LC-ESI-MS run (Fig. 2). This suggested an increase in the detectable fraction of polar metabolites, supporting the observation that the derivative of phenylalanine eluted in the early segment of the run. MS/MS data (obtained through separate set up using a MicroMass Q-TOF MS instrument) further confirmed the butylated phenylalanine product ($m/z$ 222) [42] with fragment ions at $m/z$ of 166 and 120 (Fig. 2).

In summary, chemical derivatization allowed for the detection of aminated cholesterol and butylated phenylalanine, whereas their underivatized forms were not detectable.
3.2. Validation and optimization of derivatization

Butylation and amination reaction conditions were studied to determine suitable incubation temperature and time duration. Using plasma samples spiked with known additional cholesterol (0.10, 0.50, 1.0, 2.0 and 5.0 mM) and phenylalanine (0.10, 1.0, 5.0, 10, 50 and 100 µM) as quality controls for amimation and butylation, respectively, the rate of formation of the derivatized products was monitored at 65°C for 1, 10 and 30 min as well as 1, 2, 4, 12, 24 and 48 h. Shorter incubation periods (i.e., <4 h) resulted in lower peak intensity of derivatized product ion. While, incubation duration at extended periods such as 12 h or longer were found to be undesirable since the degradation of many major products was observed; additionally, derivatization reproducibility also decreased resulting in high %CV values of >40%. It was determined that the incubation time duration of 4 h at 65°C was most suitable for both derivatization reactions, with butylation reaction >95% complete for phenylalanine, and the amination reaction >90% complete for cholesterol. By examining the peak intensity areas generated using XCMS, the derivatization reproducibility (n = 5) between preparations was found to be similar to the re-injection reproducibility (n = 3); the overall %CV was ~25%.

With amination, reaction solvent systems consisting of acetonitrile, acetone, chloroform, dichloromethane, DMSO and DMF as well as solvent mixtures were explored. Results showed that the amination reaction efficiency and reproducibility was most suitable when dichloromethane was used, giving the best amination reaction reproducibility. Three quality control serum/plasma extract samples: water, 3M HCl, and dichloromethane were incubated under the same reaction conditions. Additionally, control blank samples containing only the respective derivatization reagents were also prepared and incubated under the same reaction conditions. Changes in the detected LC-ESI-MS XCMS metabolite feature list were examined to ensure that the derivatization products were not resulted from the breakdown of the reagent, solvent impurity reactions, acid hydrolysis alone, or from elevated temperature alone.

Since the derivatization efficiencies of these derivatization agents have been previously studied, the focus of the current work was to ensure suitable derivatization conditions for metabolomics applications were used, as well as acceptable derivatization reproducibility, derivatized product stability, and re-injection reproducibility of the derivatization method. From our experience in LC-ESI-MS analysis of underivatized human plasma extracts, run to run (including both intra-day and inter-day) variability in peak intensity areas is typically ~25%. We found that with the addition of the derivatization procedure, the inter-assay variability in peak intensity area count remained within 25% variability, indicating that the derivatization methods were highly reproducible.

3.3. Derivatized metabolite analysis of control human plasma

To further examine the effects of derivatization, XCMS data tables were generated by comparing LC-ESI-MS data from underivatized and derivatized samples (see Section 2 for details). Comparison of the LC-ESI-MS data before and after derivatization revealed an increase in the number of detected features: from ~2000 to ~5000 for each of the amination and butylation derivatization procedures. It should be noted that the number of metabolite features does not directly correspond to the number of metabolites, since isotopic, adduct, fragment, and multiple charged ions were all considered separate features in XCMS analyses. In addition, it is also known that derivatization reactions inherently introduce complexity to the sample mixtures; for example, the amination reagent reacts with other nucleophiles including thiol and amine groups. Some ions contain both ionizable group(s) as well as targeted derivatization functional group(s); when the derivatization is not 100% complete, such ions were detected in the underivatized sample as well as in the derivatized sample giving rise to multiple ion peaks, an example is
the phospholipid ions described earlier in Section 3.1. Furthermore, multiple products (in the case where multiple derivatized chemical groups are present, as in the case for butylated glutathione) as well as breakdown products can arise from a single metabolite when subjected to the derivatization conditions.

As described in Section 3.2, we have examined reaction controls, with the reagent alone, with acid and/or heat treatment to account for false positives resulting from impurities in reagents, and from products of acid hydrolysis and heat treatment alone. For example, phospholipid molecules detected prominently in the TIC of underivatized plasma (the 5 most intense peaks, eluting between 25 and 35 min [19]) were absent in the TIC from butylated plasma due to acid hydrolysis (Fig. 2). To counteract these issues, XCMS datasets obtained from a series of control experiments were analyzed where the known artifacts, i.e., the unique features arising only from the described control experiments, were removed from the metabolite feature XCMS output. By discounting features that were considered known artifacts, an estimated 50% increase in the number of newly detected features was achieved with each of the derivatization methods. In other words, for each of the amination and butylation reactions, ~1000 new features were detected.

The intensity increase of newly detected XCMS features after amination and butylation procedures are illustrated in the 3D contour plots in Figs 1 and 2. Previous work has suggested that an increase in hydrophobicity of compounds results in increased ESI response [37,58,59]. Such an increase in hydrophobicity will also increase retention in a reversed phase system thereby reducing ionization suppression caused by salts and extensive co-elution of polar metabolites present in serum, giving favorable derivatization results, which is the case for the addition of butyl group and the amination conjugate.

3.4. Relative quantitation of isotopically labeled metabolites

In previous sections, we have demonstrated the advantage of amination and butylation derivatizations in enhancing the detection of an additional ~1000 metabolite features per LC-ESI-MS run. While most of the underivatized and derivatized metabolite features have not been structurally characterized, in an untargeted search of significant features, a typical approach is to compare sets of samples, i.e. samples from different disease states to detect a difference. In such comparative experiments, it is important to be able to obtain accurate ratios of the same metabolite features between the data from different sample sets. Chemical derivatization strategies combined with isotope labeling offer the potential for increased detection and the simultaneous direct relative quantitation of derivatized metabolites. The principal idea of differential isotope labeling for the current metabolomics study is similar to the first steps of ICAT [51] strategy used in comparative proteomics, except that there are no affinity capture of the reagent tags in the current work. The labeling approach presented here consists of two steps: (1) parallel derivatization of samples with various isotopically labeled forms of the chemical derivatization reagent, and (2) mixing the differentially derivatized products together and analyzed in a single LC-ESI-MS run.

Since the same features originated from different sample sets were differentially derivatized, the resulting data provided the relative abundance of that particular metabolite feature between the various sample sets, where the $m/z$ difference reflected the $m/z$ difference between the unlabeled and labeled forms of the derivatization reagent. Excellent $R$-square ($R^2$) values (>0.9977) were obtained from the differential derivatization using D0 and D9 forms of butylation reagent in a study mixture of citrulline, phenylalanine, and oxidized glutathione, demonstrating the utility for relative quantitation. The corresponding isotopic ion pairs ($m/z$ values for D0 and D9 forms) were 232/241, 222/231 and 419/437, respectively (Fig. 3). The $m/z$ difference of 9 observed for the citrulline and phenylalanine isotopic ion pairs originated from the replacement of 9 hydrogen atoms of butylation reagent with 9 deuterium atoms in the isotopically labeled butylation reagent.
Fig. 3. Differential isotopic labeling. An equal molar mixture of three analytes: citrulline, phenylalanine, and oxidized glutathione, was serially diluted (giving 0.1, 1, 5, 10, 50, 100 mM) followed by derivatization with unlabeled (D₀) butylation reagent. A separate equal molar solution mixture of these three metabolites, derivatized with labeled (D₉) butylation reagent at a single concentration (1 mM), was used as the internal standard. Equal volumes of labeled and unlabeled analyte mixtures were combined (each mixture prepared in triplicate) and each analyzed in triplicate in an LC-ESI-MS run resulting in a total of 54 analyses. (Upper right) Ratios of isotopic integrated peak areas (denoted as D₀/D₉) were plotted against ratios of concentration values (denoted as [D₀]/[D₉]) for each analyte. As shown are linear relationships between instrument response and analyte concentration for each analyte, and the calculated $R^2$ values are also listed. The overall %CV for the integrated isotopic peak areas ranged from 8 to 12% for the three metabolites each prepared at the six concentration levels. (Lower left to lower half) The chromatograms shown are those of D₀ and D₉ phenylalanine at concentration ratios of 10:1, 5:1, 1:1 and 0.1:1.

In the case of glutathione, however, an $m/z$ difference of 18 was observed due to the fact that all four carboxylic acids have been butylated and the doubly charged ions of the quadruple labeled oxidized glutathione were detected. The triply labeled doubly charged ions of glutathione (391/404) were also detected and their concentration ratios gave an $R^2$ value of $>0.9977$. The fact that both derivatized forms of oxidized glutathione gave the same linear correlation with concentration demonstrated that it is not crucial for derivatization reactions to go to completion. Provided that the derivatization reaction conditions are identical, relative quantities for the products of differentially labeled samples can be determined. Additionally, by examining the peak area ratios of the isotopic pairs, the precision greatly increased giving an overall %CV of $<12\%$ (as compared to the $\sim25\%$ in individual run analysis). This increase in assay precision is an important advantage when comparing the same metabolites between two or more samples as compared to individual analysis.

By combining the differentially derivatized metabolite mixtures in one LC-ESI-MS run, quick and precise relative quantitation of the individual metabolites was achieved. However, a potential drawback of the differential isotopic labeling approach is when ions with similar $m/z$ values may result in the
false increase of intensity for one of the ions in the ion pair. This problem can be minimized by using high accuracy and high resolution mass analyzers where ions of similar masses can be differentiated. Another problem might be chemical in nature wherein reactive metabolites may form adducts with each other. In this case, the difference between sample sets may still be observed when the metabolite profiles are examined, but the identification process would be more difficult.

When evaluating the effectiveness of derivatization approaches, additional benefits should be noted. For example, the fact that multiple ions could potentially arise from the same compound (such as phospholipids) is not an issue, rather, the multiple ions resulted will all be compared between sample sets, and the effectiveness of the comparison actually increases since the multiple ions will help validate the presence and ratio changes of the same metabolite originating from different sample sets. Additional benefits include enhanced structural elucidation when \( m/z \) differences between isotopic ion pairs point to multiple labeling of a metabolite. This was demonstrated in the triple and quadruple labeling of oxidized glutathione due to the presence of multiple reactive carboxylic acids. The phenomenon is potentially beneficial in identifying the number of reactive chemical groups in the unknown of interest. Furthermore, variations of the isotope labeling approach for quantification are limitless. For example, if multiple isotopically labeled forms of the derivatization reagent are available, e.g. containing 4, 8 and 12 deuterium labeled forms of the reagent, then the simultaneous direct relative quantitation of multiple sets of samples can be achieved with one single LC-ESI-MS analysis.

3.5. Clinical application for biomarker discovery

To demonstrate the utility of isotopically labeled derivatization with clinical samples, in addition to SIRS and sepsis samples individually analyzed as underivatized samples as well as individually derivatized with butylation reagent and with amination reagent, pooled plasma samples derivatized with either the \( D_0 \) and \( D_9 \) butylation reagent pair or the \( ^{13}C_0 \) and \( ^{13}C_4 \) amination reagent pair were also prepared and assayed. The pooled sepsis sample was derivatized with the unlabeled reagent while the pooled SIRS sample was derivatized with the isotopically labeled reagent. To illustrate the difference in data analysis approach of pooled and individually derivatized sepsis and SIRS samples, an example of the \( ^{13}C_0 \) and \( ^{13}C_4 \) amination ion pair peak intensities is shown in the upper portion of Fig. 4. The lower portion of Fig. 4 shows the XCMS EIC output plots of three potential sepsis biomarkers identified by individual sample derivatization.

In the pooled sepsis and SIRS differentially derivatized sample analysis, the intensity areas of the molecular ion peaks were integrated using XCMS to provide relative quantitative information regarding the two disease states. By comparing the corresponding metabolite feature intensity areas (generated by XCMS) between the two disease states in the aminated set and butylated set of samples using a \( t \)-test, and sort the entire metabolite feature list according to probability (\( p \)-value), the top 100 most significant features (ranked by \( t \)-test) from the pooled studies were compared with the significant features found in the individually derivatized samples.

The individual derivatized sepsis and SIRS samples are important due to potential individual differences between patients, which were not detected in pooled sample analyses. It is very likely that the pooled sample may be affected by a few members with exceptionally high or low abundance of a particular metabolite feature, and, therefore, the results of pooled samples may be biased. In the individually derivatized set of sepsis and SIRS samples that are butylated or aminated, four parameters of the XCMS output and raw LC-ESI-MS data were examined: (1) results of a \( t \)-test (i.e. evaluating the \( p \)-value), (2) fold change in peak intensity area, (3) EIC plot and (4) raw LC-ESI-MS data. The criteria used were
Fig. 4. Pooled versus individual analysis of the sepsis disease state. (Top) samples from sepsis and SIRS patients were first pooled by disease group, generating one pooled sepsis sample and one pooled SIRS sample, then followed by chemical derivatization with unlabeled (13C0) and labeled (13C4) amination reagents, respectively. The two derivatized and pooled sepsis and SIRS samples were subsequently combined in equal volumes and analyzed in a single LC-ESI-MS run. The chromatogram showing the ion peaks of a differentially isotopically labeled metabolite present in both pooled sepsis and pooled SIRS samples were detected simultaneously and quantitated using the respective integrated peak areas. (Below) Individual clinical samples from sepsis (n = 27) and SIRS (n = 34) patients were derivatized by amination reagent in one setup, and butylation reagent in another setup. The derivatized sepsis and SIRS samples resulted from the same derivatization method were analyzed and compared using XCMS. Alignment of three separate extracted ion chromatograms (EICs) generated using XCMS of individual LC-ESI-MS runs of all sepsis and SIRS samples are shown. Results are color coded according to the disease state: sepsis in red and SIRS in black. Research into the structural characterization of the molecules is still ongoing.

(1) $p$-value $> 0.001$ (99.9% confidence), (2) fold change $>1.5$ or $<0.5$, i.e., more abundant in sepsis or more abundant in SIRS, respectively, (3) normal peak shapes in the XCMS EIC plot, and (4) confirmation of the presence of identified peaks in the raw LC-ESI-MS data. Based on the criteria used, the data analysis of underivatized, butylated and aminated sepsis and SIRS samples resulted in 215, 186 and 105 respective reproducible significant metabolite features that are currently being further characterized and evaluated.

At least 5 of the significant features from the pooled analysis were found to be the same as those observed with the individually analyzed samples, illustrating the utility to pool samples from sample sets enabling differential labeling and direct relative quantitation. The current approach produced encouraging and reproducible data that are being further pursued in the areas of MS/MS fragmentation data analysis, accurate mass analysis, elementary composition calculation, mass spectral database search, fraction collection and concentration for other analytical techniques such as capillary NMR, and syn-
thesis of candidate compounds for validation purposes. The structural characterization and biological relevance of the significant features is currently on-going and will be presented in a follow-up study.

4. Conclusions

This study demonstrated the use of chemical derivatization for LC-ESI-MS in a metabolomics approach. Butylation and amination of serum samples displayed enrichment in detected metabolic components. This was illustrated by (1) chromatograms richer in peaks, (2) detection of selected “type”-compounds, phenylalanine and cholesterol, which were undetected in the non-derivatized serum. The >50% increase in detected metabolite features after butylation derivatization and amination derivatization compared to the underivatized sample is a significant result, as a large number of metabolites present in plasma were not detected by LC-ESI-MS without derivatization. When reproducible chemical derivatization techniques are used to routinely assess human biofluids such as human serum, the results may be used to gain insights and further understanding of human metabolism. While two mechanistically different chemical derivatization reagents were used in the current study to maximize the detection of different metabolite features, it would be of great interest to use other derivatization reagents targeting the same chemical species to compare and further confirm the results of derivatized human serum extracts; for example ferrocenyl azide [33] and tris(trimethoxyphenyl)phosphonium reagent (TMPP) [33,41] are very promising reagents that have been shown to modify hydroxyl group and significantly enhance ionization of previously undetected metabolites. The use of derivatization reagents targeting other function groups are also expected to provide additional useful data for metabolomics research. Furthermore, a different ionization enhancing strategy involving solvent composition [60] is yet to be explored to achieve even higher number of metabolite features to be detected.

The future goals of this study include validating the increase or decrease in ion intensities (i.e., concentration) of the significant metabolite feature(s), structurally characterizing these metabolite(s) or disease biomarker(s), synthesizing isotopically labeled internal standard(s) of the disease biomarker(s), and spiking the internal standard(s) into clinical samples to more accurately quantitate the disease biomarker(s). In an ideal situation, all the metabolite features detected in underivatized as well as all derivatized samples will be structurally characterized, and a complete metabolic map can be constructed to expand our understanding of human metabolism.

The combination of derivatization with stable isotope labeling offered the possibility of performing simultaneous relative quantitation of detected components while affording improved precision via peak ratio comparisons. The comparative analysis of SIRS and sepsis clinical samples also showed encouraging results that are being further investigated. While this work demonstrated the effective use of chemical derivatization for metabolomics studies, it also lays the groundwork for future studies aimed at achieving even more comprehensive and quantitative metabolite profiles.

Acknowledgements

A.N. is supported by a postdoctoral fellowship from The Swedish Research Council (VR). The Scripps authors appreciate the support of grants from the NIH (P30 MN062261) and DOE (DE-AC02-05CH11231) in the development of this technology. We would like to also thank Dr. Meng-Lin Tsao and Dr. Juncai Meng for technical assistance.
References


J.K. Kim, K. Harada, T. Bamba, E. Fukusaki and A. Kobayashi, Stable isotope dilution-based accurate comparative quan-
Submit your manuscripts at http://www.hindawi.com