analytical chemistry

Article

pubs.acs.org/ac

Development of a Liquid Chromatography–High Resolution Mass Spectrometry Metabolomics Method with High Specificity for Metabolite Identification Using All Ion Fragmentation Acquisition

Shama Naz,[†,⊥] Hector Gallart-Ayala,[†,⊥] Stacey N. Reinke,[†] Caroline Mathon,[†] Richard Blankley,[‡] Romanas Chaleckis,[†,§] and Craig E. Wheelock*[,†,§]

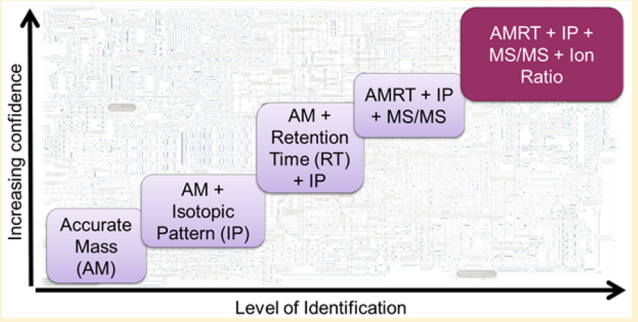
[†]Division of Physiological Chemistry 2, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm SE 17177, Sweden

[‡]Agilent Technologies, Cheadle, Cheshire U.K.

[§]Gunma University Initiative for Advanced Research (GIAR), Gunma University, Gunma, Japan

Supporting Information

ABSTRACT: High-resolution mass spectrometry (HRMS)-based metabolomics approaches have made significant advances. However, metabolite identification is still a major challenge with significant bottleneck in translating metabolomics data into biological context. In the current study, a liquid chromatography (LC)–HRMS metabolomics method was developed using an all ion fragmentation (AIF) acquisition approach. To increase the specificity in metabolite annotation, four criteria were considered: (i) accurate mass (AM), (ii) retention time (RT), (iii) MS/MS spectrum, and (iv) product/precursor ion intensity ratios. We constructed an in-house mass spectral library of 408 metabolites containing AMRT and MS/MS spectra information at four collision energies. The percent relative standard deviations between ion ratios of a metabolite in an analytical standard vs sample matrix were used as an additional metric for establishing metabolite identity. A data processing method for targeted metabolite screening was then created, merging *m/z*, RT, MS/MS, and ion ratio information for each of the 413 metabolites. In the data processing method, the precursor ion and product ion were considered as the quantifier and qualifier ion, respectively. We also included a scheme to distinguish coeluting isobaric compounds by selecting a specific product ion as the quantifier ion instead of the precursor ion. An advantage of the current AIF approach is the concurrent collection of full scan data, enabling identification of metabolites not included in the database. Our data acquisition strategy enables a simultaneous mixture of database-dependent targeted and nontargeted metabolomics in combination with improved accuracy in metabolite identification, increasing the quality of the biological information acquired in a metabolomics experiment.



High-resolution mass spectrometry (HRMS)-based metabolomics has become an integral method for understanding health and disease.[1,2] Metabolomics has been used to obtain insight into multiple biochemical processes, including biomarker discovery, food safety, and nutrition,[1,3,4] and is considered an important component of precision medicine initiatives.[5] These studies are often performed without a prior knowledge of metabolite identity, and compound identification is frequently based upon database searches (e.g., HMDB, Metlin, and KEGG).[6−8] The accurate identification/annotation of metabolites is a vital component of transferring HRMS data into biological information; however, it remains a significant bottleneck in nontargeted metabolomics to correctly annotate the biological identity of a detected feature.[9−12] Several guidelines have been proposed for metabolite identification to aid in the ability to directly compare data from different studies and laboratories,[13−16] following upon the criteria proposed by the Metabolite Standard Initiative (MSI).[17]

The MSI defined four levels of metabolite identification.[17] The highest level (level 1) is based on matching two or more orthogonal properties [e.g., accurate mass (AM), retention time (RT)/index, isotopic pattern, MS/MS spectrum] of an authentic reference standard analyzed under the same condition as the metabolite of interest. This level of structural information provides a high level of confidence in metabolite identity but is resource intensive. Attempts have also been made to develop targeted screening methods for a large number of metabolites.[18−20] However, while useful, these screening methods are by definition limited to a selected group of targeted metabolites, which restricts discovery efforts to those metabolites included in the targeted list. While generating MS/MS data for all metabolites is challenging, a number of

Received: March 13, 2017
Accepted: June 22, 2017
Published: June 22, 2017

DOI: 10.1021/acs.analchem.7b00925
Anal. Chem. 2017, 89, 7933−7942

tools have been proposed for in silico MS/MS-based metabolite identification.[21−23] However, even this level of annotation does not address the issue of identifying coeluting isobaric compounds, which traditionally requires chromatographic resolution. To improve metabolite identification and reduce the requirement for multiple analytical runs for structural confirmation, two different MS/MS strategies have been implemented to date in nontargeted metabolomics: (i) with selection of the precursor ion (data dependent acquisition), and (ii) without selection of the precursor ion [all ion fragmentation (AIF), data independent acquisition (DIA), and MS^E].[24−31] DIA-based MS generates MS/MS spectra that contain a mixed population of product ions together with their precursor ions and the extracted ion chromatogram (EIC) of each product ion needs to be mapped to its parent compound. This can be a challenging process; however, recent software developments have addressed some of these issues.[32,33] The AIF and MS^E approaches have been successfully used to conduct multiple fragmentation experiments in a single acquisition.[30,31,34,35] The difficulty in identifying coeluting isobaric compounds has been suggested to be solved using a DIA-based approach in combination with software deconvolution algorithms that merge precursor ions from low energy experiments and product ions from high energy experiments.[32] This approach has been successfully applied in lipidomics.[35,36]

The aim of the current work was to establish a comprehensive analytical workflow for the application of liquid chromatography−mass spectrometry (LC-MS) to nontargeted metabolomics with a high level of accuracy in metabolite identification, employing the all ion fragmentation (AIF) approach. To accomplish this aim, we developed an HRMS-based metabolomics method coupled to both reversed phase (RP) and hydrophilic interaction liquid chromatography (HILIC) for metabolite screening. The applied AIF mode includes 3 sequential full scans at 0, 10, and 30 eV collision energies. In the subsequent data analysis, EIC from any precursor or associated product ions of interest can be extracted from the low or high energy scan data. One EIC is chosen for relative quantification (the quantifier ion) of the metabolite, and further product ions from the same compound are used as qualifier ions. The ratios of qualifier/quantifier ion intensities are established from authentic analytical standards and should be preserved when measured in a biological sample, increasing the accuracy of the identification. The same acquired data (0 eV) can be used in parallel for a global metabolite profiling workflow, enabling a combined database-dependent targeted and nontargeted metabolomics experiment. The combination of the AIF-based data acquisition with the ion ratio confirmation and deconvoluted coeluting isobaric pairs provides a useful method for increasing the accuracy of metabolite identification in a metabolomics experiment.

MATERIALS AND METHODS

Reagents and Chemicals. LC-MS grade water and formic acid were purchased from Sigma-Aldrich (St. Louis, MO, United States). Acetonitrile (Optima, LC/MS), methanol (Optima, LC/MS) and 2-propanol (Optima, LC/MS) were purchased from Fisher-Scientific (Loughborough, U.K.). The internal lock masses (purine and HP-0921) and tune mix for calibrating the TOF-MS (ESI-low concentration tuning mix) were purchased from Agilent Technologies (Santa Clara, CA, United States). The analytical standards used to construct the compound spectral database as well as the internal standards (Tables S1−S5) were purchased from Sigma-Aldrich, Cayman Chemical Company (Ann Arbor, MI, United States), Toronto Research Chemicals (Ontario, Canada), Zhejiang Ontores Biotechnologies Co., Ltd. (Zhejiang, China), and Avanti Polar Lipids, Inc. (Alabaster, AL, United States) depending upon availability. The internal standards and standards were prepared at 1 mM concentrations in the appropriate solvent for dissolution, stored at −20 °C, and diluted appropriately on the day of analysis.

LC-HRMS Instrumentation. All experiments used a 1290 Infinity II ultrahigh performance liquid chromatography (UHPLC) system coupled to a 6550 iFunnel quadrupole-time-of-flight (Q-TOF) mass spectrometer equipped with a dual AJS electrospray ionization source (Agilent Technologies).

Metabolite separation was performed with two complementary stationary phases. Polar metabolites were separated on a HILIC SeQuant ZIC-HILIC (Merck, Darmstadt, Germany) column 100 Å (100 × 2.1 mm, 3.5 μm particle size) coupled to a guard column (2.1 × 2 mm, 3.5 μm particle size) and an inline-filter. Sample analysis in both positive and negative ionization mode was performed using water with 0.1% formic acid (solvent A) and acetonitrile with 0.1% formic acid (solvent B). The elution gradient used was as follows: isocratic step at 95% B for 1.5 min, 95 to 40% B in 12 min, maintained at 40% B for 2 min, then decreasing to 25% B at 14.2 min, maintained for 2.8 min, then returned to initial conditions over 1 min, and then the column was equilibrated at initial conditions for 7 min. The flow rate was 0.3 mL/min; injection volume was 2 μL, and the column oven was maintained at 25 °C.

Nonpolar metabolites were separated on a RP Zorbax Eclipse Plus C18, RRHD (Agilent Technologies) (100 × 2.1 mm, 1.8 μm particle size) column coupled to a guard column (5 × 2 mm, 1.8 μm particle size) and an inline-filter. Sample analysis in both positive and negative mode was performed using water with 0.1% formic acid (solvent A) and 2-propanol:acetonitrile (90:10, v/v) with 0.1% formic acid (solvent B). The gradient elution was set as follows: isocratic step of 5% B for 3 min, 5 to 30% B in 2 min, then B was increased to 98% in 13.5 min, maintained at 98% B for 1.5 min, returned to initial conditions over 0.5 min, and then held for a further 4.5 min. The flow rate was 0.4 mL/min, injection volume was 2 μL, and the column oven was maintained at 50 °C.

Both chromatographic separations were coupled to an Agilent 6550 Q-TOF-MS system. The system was calibrated and tuned according to the protocols recommended by the manufacturer. Nitrogen (purity >99.999%) was used as a sheath gas and drying gas at a flow of 8 and 15 L/min, respectively. The drying and sheath gas temperature was set at 250 °C with the nebulizer pressure at 35 psig and voltage 3000 V (± for positive and negative ionization mode). The fragmentor voltage was set at 380 V. The acquisition was obtained with a mass range of 40−1200 m/z for HILIC and 50−1200 m/z for RP in AIF mode, where a single high-resolution full scan is acquired, including three sequential experiments at three alternating collision energies (one full scan at 0 eV, followed by one MS/MS scan at 10 eV, and then followed by one MS/MS scan at 30 eV). The data acquisition rate was 6 scans/s.

An internal lock mass mixture (Agilent Technologies) was prepared at a final concentration of 2 μM purine ($C_5H_4N_4$) and 2.5 μM HP-0921 ($C_{18}H_{18}O_6N_3P_3F_{24}$) in acetonitrile:water (19:1, v/v). The internal lock mass mixture was constantly infused at a flow rate of 1 mL/min (split 1:100) using an isocratic pump together with the LC eluent for constant mass

DOI: 10.1021/acs.analchem.7b00925
Anal. Chem. 2017, 89, 7933−7942

correction [positive ionization mode: purine (m/z 121.0509), HP-0921 (m/z 922.0098); negative ionization mode: purine (m/z 119.0363), HP-0921 (m/z 966.0007, HP-0921 + formate adduct)]. Although observed mass accuracy will depend upon the resolution, potential metabolite coelution, and isobaric compounds, a mass accuracy of <5 ppm was expected.

The Agilent 6550 Q-TOF-MS system was calibrated both in positive and negative ionization mode with ESI low concentration tuning mix (Agilent Technologies) at 2 GHz extended dynamic range mode over a mass range of 0−1700 m/z. The mass resolution obtained at full width half maximum (fwhm) for each calibrant mass is as follows: m/z 118.0863, resolution ∼14 000; m/z 322.0481, resolution ∼21 000; m/z 622.0210, resolution ∼26 000; m/z 922.0098, resolution ∼29 000; m/z 1221.9906, resolution ∼30 000; and m/z 1521.9715, resolution ∼31 000).

Analytical Standard Characterization for in-House Database Construction. For metabolite database construction, 408 analytical standards and 21 internal standards were diluted appropriately (∼10 to 50-fold according to signal intensity) and characterized both in full scan and MS/MS mode at four different collision energies (10, 20, 30, and 40 eV) using the same mobile phase conditions optimized for the chromatographic separation, applying flow injection analysis. Metabolites were separated into 3 different conditions in RP depending upon elution time: (1) eluting very early (<3 min), (2) eluting from 3 to 7 min, and (3) eluting late (>7 min). Different mobile phase conditions were used for the 3 conditions (90% water, 50% water, and 10% water, respectively) for MS and MS/MS characterization. Metabolites were separated according to previous experience and log P values. A similar approach was used for metabolites in the HILIC database (10% water, 50% water, and 90% water). Following characterization, it was decided that the combination of 10 and 30 eV provided the optimal fragmentation for the range of metabolites included in the in-house database. The ionization efficiency of individual metabolites can vary greatly. Accordingly, to analyze diverse metabolites simultaneously with high sensitivity, all metabolites included in our in-house database were analyzed in both positive and negative ionization mode. During characterization, multiple compounds evidenced adduct formation and/or in-source fragmentation among the protonated or deprotonated molecule. The MS/MS acquisition was therefore performed on the precursor ion as well as on adducts or in-source product ions. Precursor ions were assigned manually based upon the molecular formulas of all metabolites, and MS spectra were interrogated considering the most common ions that appeared in both positive and negative ionization mode. Metabolite characterization was performed individually for each compound to avoid incorrect assignments due to the presence of multiple metabolites. In addition, all standards and internal standards were analyzed using the appropriate LC method for RT determination. The measured RTs, acquired MS [protonated (+H), deprotonated (−H), adducts (sodium, ammonium, potassium, and formate) and in-source fragmentation (NH_3, H_2O, CO_2, CO_2H, etc.)] and MS/MS spectra (at four different collision energies) were used to create a personal compound database library (PCDL) using the Agilent PCDL software (version B.07.00, Agilent Technologies) (Tables S1−S5). The metabolites were split into two separate PCDLs based upon their chromatographic retention in either HILIC or RP. This approach enabled the selective identification of metabolites based upon their optimal chromatographic method (either HILIC or RP). Generally, polar metabolites were in the HILIC database (n = 194), and nonpolar metabolites were in the RP database (n = 214).

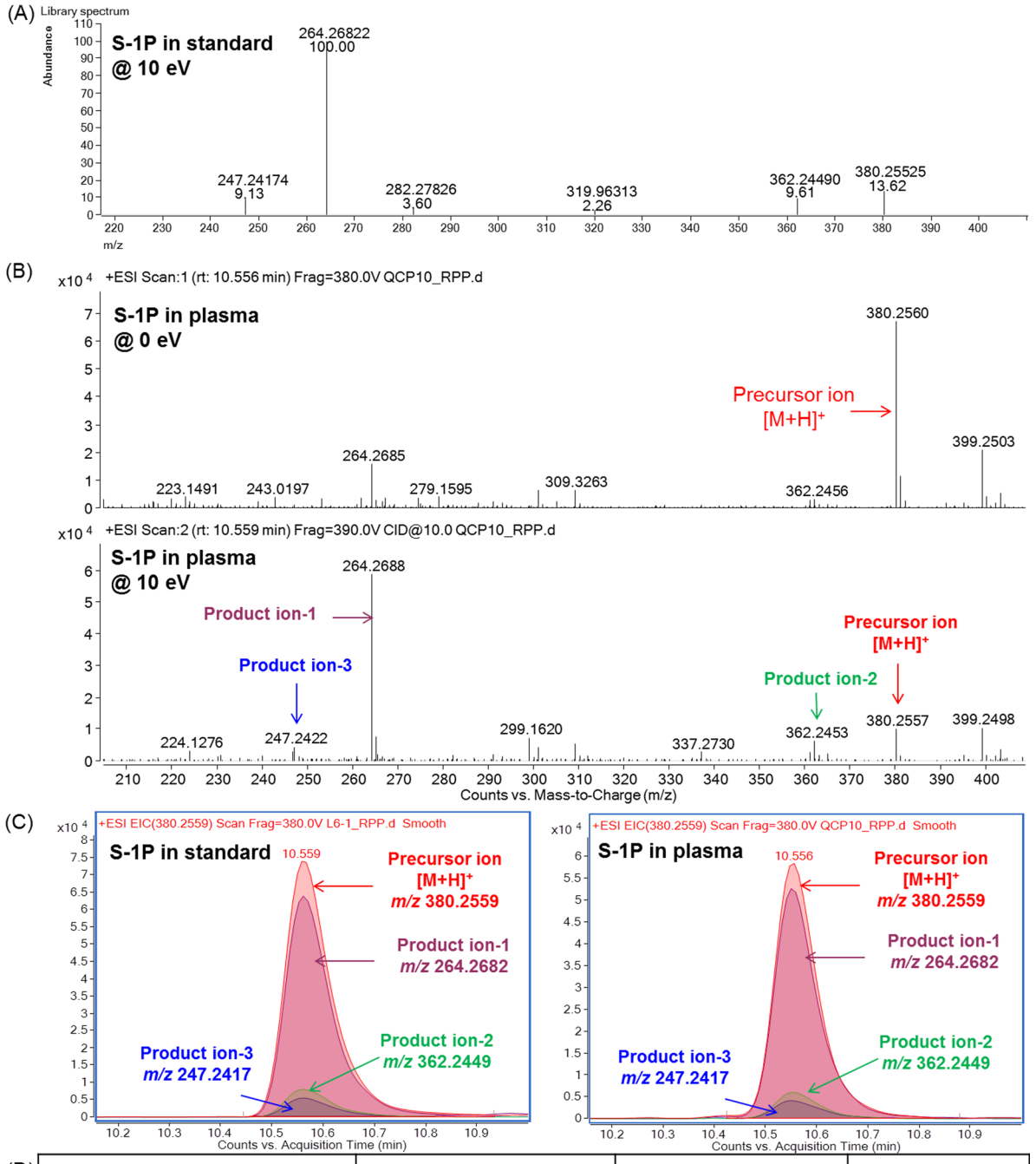
Data Processing Method for Targeted Metabolite Screening. A data processing method for targeted metabolite screening was constructed in the Agilent TOF-Quant software (version B.07.00, Agilent Technologies) using precursor and product ion information. Two or more product ions (from 10 and/or 30 eV) were then chosen depending on the fragmentation pattern for each metabolite in the following order of priority: (i) the most abundant and (ii) the most unique. For each metabolite, a quantifier ion (precursor ion), qualifier ion(s) (product ions), their relative ion ratios, and RTs were merged in the data processing method (Tables S2−S5). Similar to the PCDL libraries, the data processing methods were also separately constructed for HILIC (n = 194) and RP (n = 214). An individual metabolite is then identified based upon MS/MS of the known selected product ions as opposed to MS/MS spectral matching, which is a significant challenge in DIA-MS.

Sample Preparation. Plasma (n = 76) was previously collected[37] and stored at −80 °C. On the day of analysis, samples were thawed on ice, and protein precipitation was performed by adding 300 μL of ice-cold LC-MS grade methanol containing a mixture of internal standards (Table S1) to 100 μL of plasma. Samples were vortexed for 20 s and left to stand on ice for 10 min before centrifuging (Eppendorf Centrifuge 5430 R, Hamburg, Germany) at 15 000g for 10 min at 4 °C. An aliquot of 60 μL of the supernatant was transferred to 4 separate LC vials (for analysis in HILIC positive, HILIC negative, RP positive, and RP negative ionization modes), and extracts were evaporated to dryness (miVac Duo Concentrator, Genevac Ltd., Ipswich, U.K.) and stored at −80 °C. On the day of analysis, dried extracts were reconstituted in 50 μL acetonitrile:water (8:2, v/v) for HILIC and in 50 μL methanol for RP metabolomics analyses.

Urine (n = 51) was previously collected[38] and stored at −80 °C. On the day of the analysis, the specific gravity was measured (UG-α Digital Specific Gravity Refractometer, ATAGO, Tokyo, Japan), and all samples were normalized to the sample with the lowest measured specific gravity using LC-MS grade water.[39] Metabolite extraction was performed by adding 180 μL of acetonitrile (containing a mixture of internal standards; Table S1) in 20 μL of specific gravity normalized urine and centrifuged (13 000g, 15 min, 4 °C). An aliquot of 40 μL of the supernatant was transferred to 4 separate LC vials (for analysis in HILIC positive, HILIC negative, RP positive, and RP negative mode) for LC-MS analysis.

HCT116 cells (n = 24) were cultured as previously described[40] and harvested with ice-cold methanol as previously reported.[41] The resulting cell pellets were dried (miVac Duo Concentrator, Genevac Ltd., Ipswich, U.K.) and 100 μL of methanol was added to the cell pellet, followed by sonication for 2 min and vortexing for 30 s. Samples were then centrifuged at 10 000g for 10 min at 4 °C. Supernatants were filtered using 0.1 μm membrane spin filters (Merck Millipore, Germany) and centrifuged for 3.5 min at 5000g at 4 °C. Eighty microliters of supernatant was transferred to four separate LC vials (for analysis in HILIC positive, HILIC negative, RP positive, and RP negative mode) for LC-MS analysis.

DOI: 10.1021/acs.analchem.7b00925
Anal. Chem. 2017, 89, 7933−7942



(D)

Ion ratio	Standard (based on peak area)	Plasma (based on peak area)	Relative error (%)
Product ion-1/ precursor ion	0.84	0.91	-7.5
Product ion-2/ precursor ion	0.11	0.10	6.4
Product ion-3/ precursor ion	0.08	0.07	4.2

Figure 1. Metabolite identification with ion ratio confirmation using data independent acquisition (AIF) data. (A) Sphingosine-1-Phosphate (d18:1) (S-1P) MS/MS library spectrum at 10 eV collision energy, obtained from an analytical standard. (B) The upper panel shows the spectrum at 0 eV (full scan), and the lower panel shows the MS/MS spectrum at 10 eV for the S-1P chromatographic peak in plasma (data acquired in RP LC-HRMS positive ionization AIF mode). Product ions were matched with the library MS/MS spectrum (1A). (C) Extracted ion chromatograms of S-1P, for its precursor ion (0 eV, red), and product ions (10 eV, product ion-1 = maroon, product ion-2 = green, and product ion-3 = blue) from an analytical standard (left panel) and plasma (right panel). Overlaying the precursor and fragment ion chromatograms shows that they have the same chromatographic apex. (D) The table displays the ion ratio calculations and the percent relative error of the plasma ion ratio compared to S-1P analytical standard ion ratio.

DOI: 10.1021/acs.analchem.7b00925
Anal. Chem. 2017, 89, 7933−7942

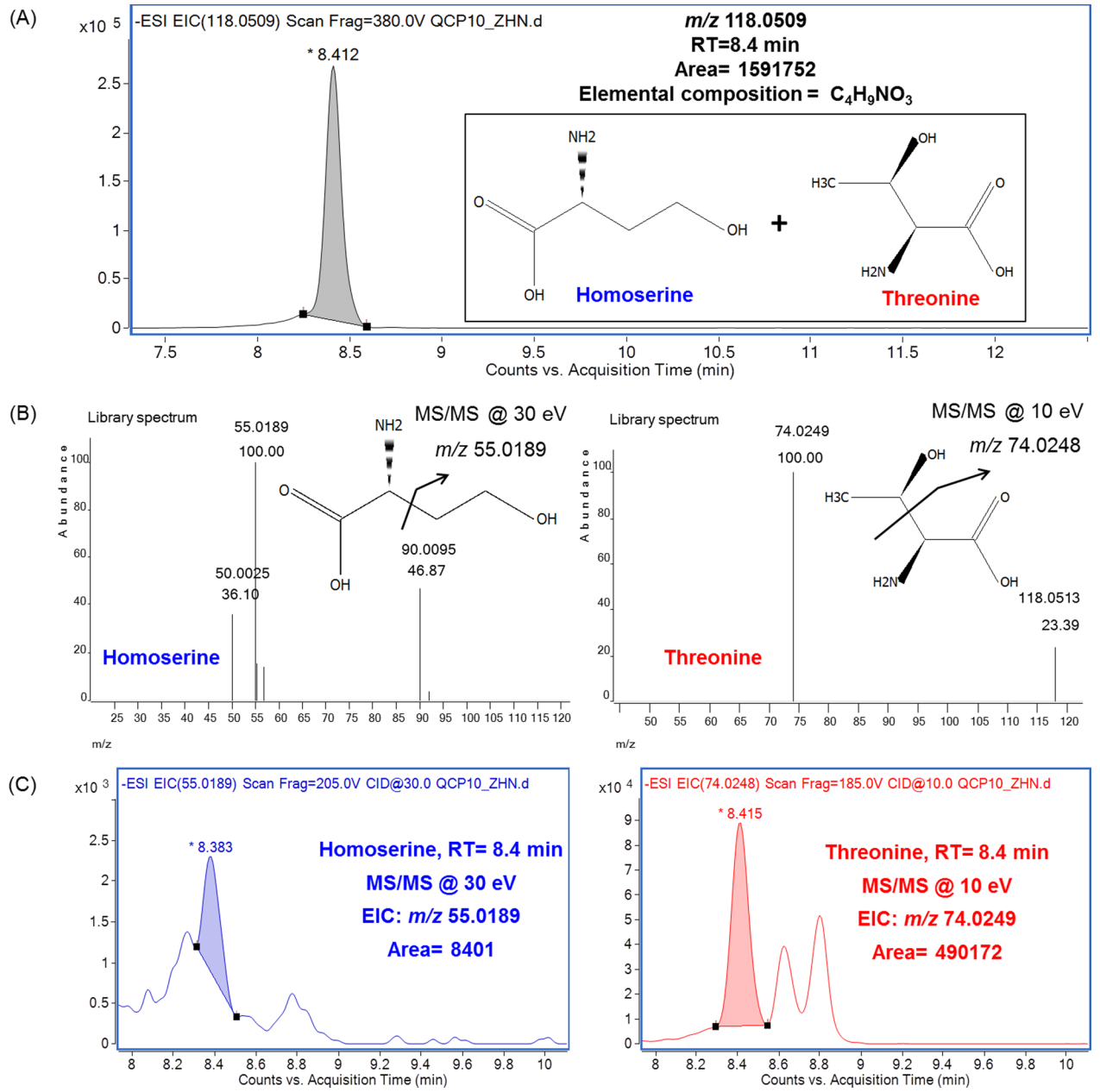


Figure 2. Improving selectivity for coeluting isobaric compounds using the AIF approach. (A) This panel displays the HILIC LC-HRMS negative ionization mode chromatographic peak at 8.4 min in plasma with putative metabolite identification as homoserine and threonine based upon an AMRT search against our in-house database (m/z 118.0509, 0 eV, full scan). (B) The left panel shows the MS/MS spectrum of homoserine at 30 eV, displaying a homoserine-specific peak using the product ion m/z 55.0189. The right panel shows the MS/MS spectrum of threonine at 10 eV, displaying a threonine-specific peak using the product ion m/z 74.0248. (C) The left panel shows the extracted ion chromatogram at m/z 55.0189 (30 eV), which is selective for the monitoring ion of homoserine. The right panel shows the extracted ion chromatogram at m/z 74.0249 (10 eV), which is selective for the monitoring ion of threonine. Use of a specific product ion from the AIF data enables the accurate quantification (relative) of each coeluting isobaric metabolite.

RESULTS AND DISCUSSION

All Ion Fragmentation. An issue with many LC-HRMS-based metabolomics methods is that metabolite identification is based upon AM followed by RT confirmation, provided that the analytical standards are available. However, structurally similar compounds can elute in the same chromatographic region or shift in RT depending upon the sample matrix. Thus, there is a potential for obtaining both false negatives and false positives by applying only AMRT-based identification criteria. Accordingly, there is a need to increase the specificity of metabolite identification. The AIF acquisition mode can be useful in this regard because the simultaneous acquisition of both precursor and product ions increases the accuracy in the identification of unknown metabolites.[24,25,30,31] In the case of AIF-based approaches, samples are often analyzed in full scan mode for metabolite profiling, while for the purpose of metabolite identification, a pooled quality control sample is analyzed in AIF mode at different collision energies.[42] A recent lipidomics-based study successfully combined full scan and AIF into a single analytical run.[36] We adopted and further optimized a similar approach for metabolomics, where in a single run

DOI: 10.1021/acs.analchem.7b00925
Anal. Chem. 2017, 89, 7933−7942

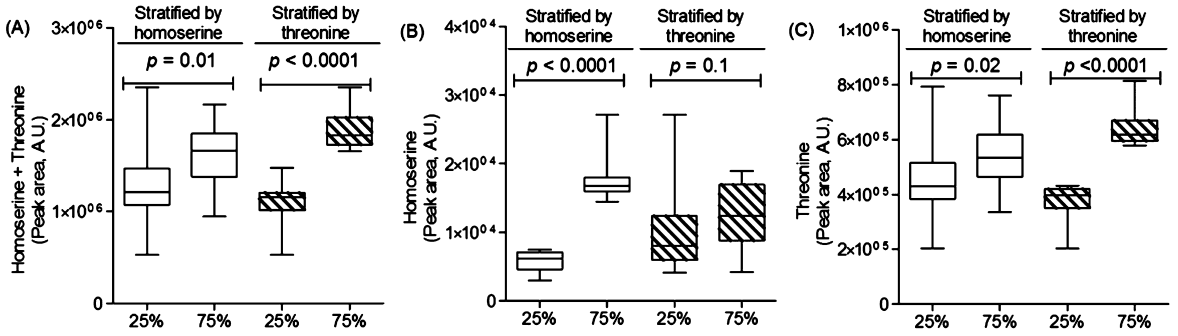


Figure 3. Specificity in metabolite identification for the coeluting isobaric pairs in plasma. Abundance of integrated peaks based upon (A) combined homoserine + threonine (m/z 118.0509 at 0 eV), (B) homoserine-specific (homoserine, m/z 55.0189 at 30 eV), and (C) threonine-specific (threonine, m/z 74.0248 at 10 eV). Samples were stratified into the 25% (n = 20) and 75% (n = 20) quantiles based upon the homoserine-specific peak (white box) and then similarly by the threonine-specific peak (dashed box). Following stratification, the significance level of the two quantiles was tested. Stratification based on the homoserine-specific value showed that the homoserine + threonine combined peak (p = 0.01) as well as homoserine (p < 0.0001) and threonine specific peak (p = 0.02) integration remained significant. With stratification by the threonine-specific value, only the combined (p < 0.0001) and threonine specific peak (p < 0.0001) remained significant. The percent relative standard deviation also decreased for the threonine- and homoserine-specific peak integration relative to that of the combined integrated peak. Significance was determined by the Mann−Whitney test (GraphPad Prism, version 5.02, La Jolla, CA, United States). A.U. = arbitrary units.

three consecutive full scans are acquired at three different collision energies: 0 eV to obtain precursor ion spectra (full scan), and 10 and 30 eV to acquire product ions. The MS/MS spectra of the 408 metabolites in our in-house database were obtained from 10, 20, 30, and 40 eV collision energies. Following characterization, it was observed that small metabolites in our in-house database were well-fragmented at low collision energy (e.g., 10 eV), and large and stable molecules in our in-house database were well-fragmented at a relatively high energy (e.g., 30 eV), providing optimal intensity of both the precursor ion and the corresponding selected product ions. Thus, a low (10 eV) and a high (30 eV) collision energy was selected based on the fragmentation pattern of the in-house database metabolites. The developed AIF method therefore obtains information related to the precursor and product ions across several metabolite classes in a single analytical run.

Ion Ratio Approach. Ion ratios, which are not based on a previous precursor ion selection, have been used for compound identification in gas chromatography coupled to electron ionization-MS.[43] We implemented this approach in our data processing method to increase the specificity of metabolite identification. The inclusion of the ion ratio (product ion/precursor ion) calculation in the data processing method for targeted metabolite screening adds an extra metabolite identification point in addition to the MSI suggested identification criteria.[17] For example, an analytical standard of sphingosine-1-phosphate (d18:1) (S-1P, molecular formula $C_{18}H_{38}NO_5P$) was characterized in RP positive mode with a precursor ion $[M + H]^+$ of m/z 380.2555 at 10.55 min, and its 3 most abundant product ions at 10 eV were m/z 264.2682, m/z 362.2449, and m/z 247.2417 (Figure 1A and 1B). However, in a biological matrix (e.g., plasma), the challenge of identification increases due to coelution and/or the presence of metabolite information on similar m/z and molecular formula. A search of the m/z 380.2555 feature in HMDB ($[M + H]^+$, ±5 ppm) generates two hits with the same molecular formula: S-1P and *N*-palmitoyl-phosphoethanolamine (Figure S1). A comparison of the RT to the analytical standard suggests the identity as S-1P; however, AM and RT are insufficient to conclusively identify the metabolite.[17] Overlaid EICs of S-1P for its precursor ion (full scan) and product ions (10 eV) in plasma show that they have the same chromatographic apex as the S-1P analytical standard (Figure 1C). To increase confidence in the identification of S-1P, the ion ratio was calculated in plasma and compared with the ion ratio of the S-1P analytical standard (Figure 1D). The ratios of the precursor ion with the three product ions were calculated separately, both in the S-1P analytical standard and in plasma, as well as the percentage of relative error in comparison to the analytical standard. According to European Commission guidelines, an error of ±25% is considered as an acceptable range to determine if the compound is ion ratio confirmed.[43] For S-1P, all three product ion ratios passed this criterion, increasing the confidence in S-1P identification. The ion ratios for the 413 standards in our in-house database are presented in Tables S2−S5.

Improving Selectivity in Metabolite Identification. Even after applying the AIF method in combination with a metabolite database of analytical standards, the selective analysis of isobaric compounds is still a major concern. For coeluting metabolites, the combination of product ions with precursor ions for structurally similar metabolites can be challenging. As recently reviewed, this limitation can be overcome by (i) improving chromatographic resolution and (ii) using software deconvolution algorithms that merge precursor ions from low energy experiments and product ions from high energy experiments.[44] The current data processing method for metabolite screening can also improve the selectivity for coeluting isobaric compounds by using a product ion as the diagnostic ion instead of the precursor ion. Accordingly, by applying this customized data processing method, the selectivity for coeluting isobaric compounds is increased. For example, in the analysis of plasma by HILIC negative ionization mode, an intense chromatographic peak was observed at 8.4 min (m/z 118.0509, full scan). On the basis of m/z and mass error <5 ppm, a search in our in-house database of the $[M - H]^-$ feature generated a possible compound identification as homoserine and threonine assigned to the elemental composition of $C_4H_9NO_3$. The equivalent search in HMDB and KEGG produces 5 and 8 compounds, respectively: threonine, hydroxyethyl glycine (drug metabolite), 4-amino-3-hydroxybutyrate, homoserine, allothreonine, 2-methylglycine, and (−)-erythro-(2*R*,3*R*)-dihydroxybutylamide. On the basis of

DOI: 10.1021/acs.analchem.7b00925
Anal. Chem. 2017, 89, 7933−7942

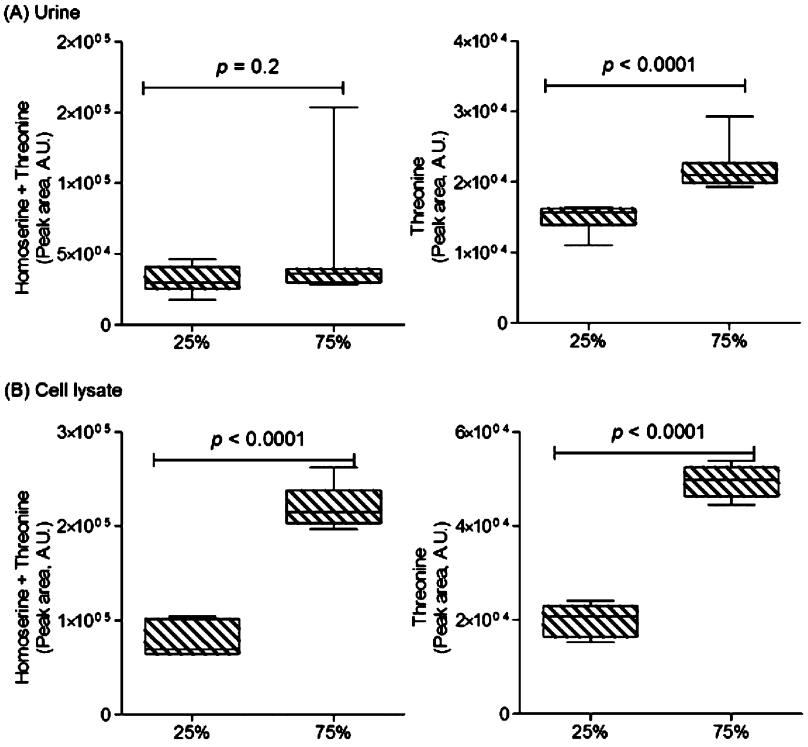


Figure 4. Specificity in metabolite identification in coeluting isobaric pairs in urine and cell extracts. Abundance of integrated peaks based upon combined homoserine + threonine (m/z 118.0509 at 0 eV), homoserine-specific (homoserine, m/z 55.0189 at 30 eV), and threonine-specific (threonine, m/z 74.0248 at 10 eV) integration. A homoserine-specific peak was not detected in either matrix. Samples were stratified into the 25% (urine: n = 13 and cell extracts: n = 6) and 75% (urine: n = 13 and cell extracts: n = 6) quantiles based upon the threonine-specific peak (dashed box). (A) In urine, stratification by the threonine-specific value showed that the combined peak is not significant between quantiles (p = 0.2), whereas the threonine-specific integrated peak was highly significant (p < 0.0001). B) In cell extracts, stratification by the threonine-specific value showed that the combined (p < 0.0001) and the threonine-specific peak were both significant (p < 0.0001). However, the percent relative standard deviation in the 25% quantile improved from 23.8 to 10.6% and from 16.9 to 7.0% in the 75% quantile with the threonine-specific integrated peak. Significance was determined by the Mann−Whitney test (GraphPad Prism, version 5.02, La Jolla, CA, United States).

the fragmentation pattern and chromatographic retention of similar metabolites, the most probable annotation for the chromatographic peak observed at 8.4 min and m/z 118.0509 corresponds to homoserine/threonine/allothreonine. Allothreonine is a plant metabolite that appears in the human diet; however, it is not often reported. This situation results in a putative metabolite annotation of both compounds (homoserine and/or threonine) based upon AM and RT because they coelute in our chromatographic conditions (Figure 2A). However, the MS/MS spectrum of homoserine at 30 eV displays a specific product ion at m/z 55.0189, and the MS/MS spectrum of threonine at 10 eV displays a specific product ion m/z 74.0248, which are unique ions for each metabolite (Figure 2B). Therefore, instead of using the precursor ion as the diagnostic ion, exchanging with this unique product ion provides each metabolite's specific contribution to the peak. Subsequently, monitoring of a selective product ion as a quantifier from the AIF experiments enables us to resolve the coeluting threonine and homoserine (Figure 2C) as well as other isobaric compounds (e.g., paraxanthine and theophylline; data not shown).

Ion-specific peak integration increases the precision in statistical analysis, which subsequently enhances biological interpretation. This strength of the current method was demonstrated in three different matrixes (plasma, urine, and cell lysate) using the homoserine and threonine isobaric pair as an example (Figures 3 and 4). The homoserine and threonine peak areas were each integrated separately using their specific product ion as a quantifier ion (homoserine, m/z 55.0189 at 30 eV; threonine, m/z 74.0248 at 10 eV), and their combined peak area (homoserine + threonine, m/z 118.0509 at 0 eV) was also integrated. The sample sets for each matrix were then stratified by the abundance of the threonine value (obtained from threonine-specific peak integration) and the homoserine value (obtained from homoserine-specific peak integration). Following stratification, the 25 and 75% quantile of each sample set were selected in an extreme value approach to test for significance. Figure 3 shows that for plasma in either stratification, the major contribution in the combined (homoserine + threonine) peak is coming from threonine. All three differentially integrated peaks were significant between the two quantiles when stratified by homoserine. However, if the stratification was based upon the threonine values, only the combined peak and threonine-specific integrated peak were significant between the 2 quantiles (p < 0.0001). For the urine and cell lysate, the homoserine-specific product ion peak could not be integrated due to low intensity (Figures 4A and B). However, the effect of the metabolite-specific peak contribution was demonstrated for urine (Figure 4A). The combined homoserine + threonine integrated peak was not significant (p = 0.2), but the threonine-specific peak was significantly different between the 25 and 75% quantile (p < 0.0001). In addition, the percent relative standard deviations (RSD) decreased in the homoserine- and threonine-specific integrated peaks between the two quantiles relative to the combined peak integration (Figure 4), increasing the precision of the

DOI: 10.1021/acs.analchem.7b00925
Anal. Chem. 2017, 89, 7933−7942

measurement. This metabolite- and matrix-dependent deconvolution example demonstrates the strength of the current method for increasing the accuracy of metabolite annotation by targeted ion selection, which can have a significant effect upon the observed biological shifts.

Data Processing Method for Metabolite Screening Including Ion Ratio Information. The AIF approach along with metabolite deconvolution algorithms aids in metabolite identification; however, it generally requires an additional laboratory and/or computational step for metabolite identification.[32] Thus, there remains an unmet need to have a flexible, customizable data processing method in which metabolite identification can be performed without a need to reanalyze samples. Toward that end, we developed a data processing method using our in-house metabolite database in the Agilent Quant software. This method extracts MS and MS/MS (e.g., from the AIF acquired data) information in a single data processing step, while simultaneously increasing the confidence of metabolite identification using ion ratio information in combination with deconvolution of isobaric compounds that can be applied to all samples. In the developed processing method, data are presented such that for each compound in each sample the "quantifier" is the high resolution EIC of the precursor ion derived from the 0 eV experiment, and "qualifier" ions are EICs of associated product ions extracted from the 10 and/or 30 eV experiments (this information was extracted and assigned manually). Two or more product ions were then chosen depending on the fragmentation pattern for each metabolite in the following order of priority: (i) the most abundant and (ii) the most unique. All of these chromatograms are overlaid so the coelution of all precursor and product ions can be verified. Additionally, the observed intensity ratios of the quantifier (precursor) and qualifier (product) ions are measured and displayed in relation to the expected ratios (Figure 1, Tables S2−S5). Where a precursor or product ion EIC shows interference or low signal-to-noise resulting from the sample matrix, the user can select a more selective ion and reprocess the data accordingly. In the case of isobaric compounds, specific product ions (upon availability) were selected as a quantifier, while precursor and other product ions were selected for qualifying the compound, and their ion ratios were also calculated accordingly. In compliance with the European Commission guidelines,[43] to qualify a compound to be ion ratio confirmed, a cutoff value of ±25% of the relative error (against the corresponding analytical standard ion ratios) was used. Our metabolite screening method gives three different rankings of accuracy in terms of metabolite identification: Rank 1, the metabolite is AMRT confirmed; Rank 2, the metabolite is AMRT and MS/MS confirmed; and Rank 3, the metabolite is AMRT, MS/MS, and ion ratio confirmed. Accordingly, our data acquisition and processing method provides high specificity for metabolite identification and relative quantification.

The ability of the developed method to detect metabolites in plasma is illustrated in Tables S2−S5. All four LC-MS methods (HILIC positive, HILIC negative, RP positive, and RP negative) were applied, and the acquired data were processed with our method. In total, 223 metabolites were identified in plasma (112 and 111 metabolites were detected from the HILIC and RP platforms, respectively). According to the identification criteria, 40 were AMRT confirmed (Rank 1), 83 were AMRT and MS/MS confirmed (Rank 2), and 100 were AMRT and MS/MS and ion ratio confirmed (Rank 3). The 100 highest-ranking metabolites (Rank 3) exhibit increased confidence in the accuracy of the metabolite identification. The metabolite distribution from each platform is presented in Figure S2.

This AIF-based method for metabolite screening offers a number of advantages over conventional nontargeted metabolomics approaches; however, there are limitations that should be taken into account. First, the method relies on characterizing analytical standards to generate the compound database, which is time-consuming and costly. However, commercial suppliers are starting to manufacture metabolomics standards kits,[45,46] and libraries can be exchanged between laboratories. Second, while applying AIF can increase the confidence in metabolite identification, the use of three energies requires faster acquisition rates, resulting in reduced signal intensity. This can become a significant issue for low abundance peaks, which are either not detectable via AIF and/or too low abundance for MS/MS acquisition. Finally, the current database dependent metabolomics method identifies metabolites with high accuracy based upon an in-house database; however, there is a trade-off in that this approach limits discovery work for reporting novel compounds. This issue can be addressed with the AIF approach (which includes a full scan at 0 eV), because the same acquired data can be used in parallel for a nontargeted (global profiling) workflow. It should be noted that, while the method performs well for the compounds in the in-house database, it remains to demonstrate the applicability of this approach to a wider swath of metabolites.

While the AIF approach to performing a metabolomics experiment is useful, it can be argued that a broad triple quadrupole (QqQ)-based method would perform better for a focused group of metabolites. This approach to metabolomics has indeed been successfully adopted by a number of groups.[18−20] The advantages of increased accuracy in the quantification as well as linear response of the instrument need to be weighed against the difficulty to perform discovery work. The primary advantage of the current AIF approach is the ability to discover new metabolites of interest that are not included in the in-house database (or presumably in the QqQ-targeted method). It accordingly comes down to a question of the experimental goals. One can envision a scenario in which an initial representative set of samples is selected for AIF analysis, followed by selection of a targeted set of metabolites of interest to build a targeted method on a QqQ system. This approach represents a logical compromise for designing a metabolomics experiment. These nontargeted data can be processed, and putative metabolites can be annotated with database searches. For compounds of interest, chemical standards can be acquired, characterized, and added to the appropriate PCDL. The original data files can then be reprocessed, eliminating the necessity of reanalyzing samples (Figure S3). The ability to use parallel but complementary data analysis strategies enables the rapid and high quality identification of common metabolites contained in the database while still allowing for unbiased profiling.

■ CONCLUSIONS

In the current study, a LC-HRMS/MS method using the AIF approach and a data processing method for targeted metabolite screening were developed to produce a metabolomics analysis workflow (Figure S3). The combination of AIF and custom data processing method enables the identification of selected metabolites in a single analytical run. The inclusion of the

DOI: 10.1021/acs.analchem.7b00925
Anal. Chem. 2017, 89, 7933−7942

product/precursor ion ratio confirmation provides increased accuracy for metabolite identification. An additional step to deconvolute coeluting isobaric pairs further increases the precision of the method for integrating peaks of specific metabolites. This method is flexible, expandable, and customizable according to study needs and enables the retrospective analysis of data to confirm the identities of novel metabolites. The current AIF-based workflow offers the advantage of increased accuracy in metabolite identification, making it a useful platform for investigating the metabolome.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.7b00925.

Supporting figures, list of internal standards, and proposed metabolomics workflow (PDF)

Calculated ion ratios for metabolites in the database, Table S2 (XLSX)

Calculated ion ratios for metabolites in the database, Table S3 (XLSX)

Calculated ion ratios for metabolites in the database, Table S4 (XLSX)

Calculated ion ratios for metabolites in the database, Table S5 (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: craig.wheelock@ki.se; Phone: +46 8 524 87630; Fax: +46 8 736 0439.

ORCID

Craig E. Wheelock: 0000-0002-8113-0653

Author Contributions

⊥S.N. and H.G.-A. contributed equally to the manuscript.

Notes

The authors declare the following competing financial interest(s): Dr. Richard Blankley is an employee of Agilent Technologies, Inc.

■ ACKNOWLEDGMENTS

We thank Dr. Antonio Checa for discussions on the method and assistance with the figures. S.N.R. was supported by a Canadian Institutes of Health Research Fellowship (MFE-161244). C.E.W. was supported by the Swedish Heart Lung Foundation (HLF 20150640). We acknowledge the support of the Swedish Heart Lung Foundation (HLF 20140469), the Swedish Research Council (2016-02798), the Karolinska Institutet and AstraZeneca Joint Research Program in Translational Science, and the Novo Nordisk Foundation (TrIC NNF15CC0018486 and MSAM NNF15CC0018346).

■ REFERENCES

(1) Dunn, W. B.; Broadhurst, D. I.; Atherton, H. J.; Goodacre, R.; Griffin, J. L. *Chem. Soc. Rev.* **2011**, *40*, 387−426.

(2) Kell, D. B.; Oliver, S. G. *Metabolomics* **2016**, *12*, 148.

(3) Johnson, C. H.; Ivanisevic, J.; Siuzdak, G. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 451−459.

(4) Brennan, L. *Essays Biochem.* **2016**, *60*, 451−458.

(5) Beger, R. D.; Dunn, W.; Schmidt, M. A.; Gross, S. S.; Kirwan, J. A.; Cascante, M.; Brennan, L.; Wishart, D. S.; Oresic, M.; Hankemeier, T.; Broadhurst, D. I.; Lane, A. N.; Suhre, K.; Kastenmuller, G.; Sumner, S. J.; Thiele, I.; Fiehn, O.; Kaddurah-Daouk, R. *Metabolomics* **2016**, *12*, 149.

(6) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorndahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A. *Nucleic Acids Res.* **2013**, *41*, D801−807.

(7) Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G. *Ther. Drug Monit.* **2005**, *27*, 747−751.

(8) Kanehisa, M.; Goto, S.; Sato, Y.; Furumichi, M.; Tanabe, M. *Nucleic Acids Res.* **2012**, *40*, D109−114.

(9) Wishart, D. S. *Bioanalysis* **2011**, *3*, 1769−1782.

(10) Viant, M. R.; Kurland, I. J.; Jones, M. R.; Dunn, W. B. *Curr. Opin. Chem. Biol.* **2017**, *36*, 64−69.

(11) Allard, P. M.; Genta-Jouve, G.; Wolfender, J. L. *Curr. Opin. Chem. Biol.* **2017**, *36*, 40−49.

(12) Dias, D. A.; Jones, O. A.; Beale, D. J.; Boughton, B. A.; Benheim, D.; Kouremenos, K. A.; Wolfender, J. L.; Wishart, D. S. *Metabolites* **2016**, *6*, 46.

(13) Salek, R. M.; Steinbeck, C.; Viant, M. R.; Goodacre, R.; Dunn, W. B. *GigaScience* **2013**, *2*, 13.

(14) Steinbeck, C.; Conesa, P.; Haug, K.; Mahendraker, T.; Williams, M.; Maguire, E.; Rocca-Serra, P.; Sansone, S. A.; Salek, R. M.; Griffin, J. L. *Metabolomics* **2012**, *8*, 757−760.

(15) Haug, K.; Salek, R. M.; Conesa, P.; Hastings, J.; de Matos, P.; Rijnbeek, M.; Mahendraker, T.; Williams, M.; Neumann, S.; Rocca-Serra, P.; Maguire, E.; Gonzalez-Beltran, A.; Sansone, S. A.; Griffin, J. L.; Steinbeck, C. *Nucleic Acids Res.* **2013**, *41*, D781−786.

(16) Dunn, W. B.; Erban, A.; Weber, R. J. M.; Creek, D. J.; Brown, M.; Breitling, R.; Hankemeier, T.; Goodacre, R.; Neumann, S.; Kopka, J.; Viant, M. R. *Metabolomics* **2013**, *9*, S44−S66.

(17) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. *Metabolomics* **2007**, *3*, 211−221.

(18) Virgiliou, C.; Sampsonidis, I.; Gika, H. G.; Raikos, N.; Theodoridis, G. A. *Electrophoresis* **2015**, *36*, 2215−2225.

(19) West, J. A.; Beqqali, A.; Ament, Z.; Elliott, P.; Pinto, Y. M.; Arbustini, E.; Griffin, J. L. *Metabolomics* **2016**, *12*, 59.

(20) Siskos, A. P.; Jain, P.; Romisch-Margl, W.; Bennett, M.; Achaintre, D.; Asad, Y.; Marney, L.; Richardson, L.; Koulman, A.; Griffin, J. L.; Raynaud, F.; Scalbert, A.; Adamski, J.; Prehn, C.; Keun, H. C. *Anal. Chem.* **2017**, *89*, 656−665.

(21) Huan, T.; Tang, C.; Li, R.; Shi, Y.; Lin, G.; Li, L. *Anal. Chem.* **2015**, *87*, 10619−10626.

(22) Aguilar-Mogas, A.; Sales-Pardo, M.; Navarro, M.; Guimera, R.; Yanes, O. *Anal. Chem.* **2017**, *89*, 3474−3482.

(23) Edmands, W. M.; Petrick, L. M.; Barupal, D. K.; Scalbert, A.; Wilson, M.; Wickliffe, J.; Rappaport, S. M. *Anal. Chem.* **2017**, *89*, 3919−3928.

(24) Zhu, X.; Chen, Y.; Subramanian, R. *Anal. Chem.* **2014**, *86*, 1202−1209.

(25) Mullard, G.; Allwood, J. W.; Weber, R.; Brown, M.; Begley, P.; Hollywood, K. A.; Jones, M.; Unwin, R. D.; Bishop, P. N.; Cooper, G. J. S.; Dunn, W. B. *Metabolomics* **2015**, *11*, 1068−1080.

(26) Hoffmann, T.; Krug, D.; Huttel, S.; Muller, R. *Anal. Chem.* **2014**, *86*, 10780−10788.

(27) Neumann, N. K. N.; Lehner, S. M.; Kluger, B.; Bueschl, C.; Sedelmaier, K.; Lemmens, M.; Krska, R.; Schuhmacher, R. *Anal. Chem.* **2014**, *86*, 7320−7327.

(28) Calderon-Santiago, M.; Priego-Capote, F.; Luque de Castro, M. D. *Anal. Chem.* **2014**, *86*, 7558−7565.

(29) Rost, H. L.; Rosenberger, G.; Navarro, P.; Gillet, L.; Miladinovic, S. M.; Schubert, O. T.; Wolski, W.; Collins, B. C.; Malmstrom, J.; Malmstrom, L.; Aebersold, R. *Nat. Biotechnol.* **2014**, *32*, 219−223.

DOI: 10.1021/acs.analchem.7b00925
Anal. Chem. 2017, 89, 7933−7942

(30) Geiger, T.; Cox, J.; Mann, M. *Mol. Cell. Proteomics* **2010**, *9*, 2252−2261.
(31) Silva, J. C.; Gorenstein, M. V.; Li, G. Z.; Vissers, J. P.; Geromanos, S. J. *Mol. Cell. Proteomics* **2006**, *5*, 144−156.
(32) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. *Nat. Methods* **2015**, *12*, 523−526.
(33) Li, H.; Cai, Y.; Guo, Y.; Chen, F.; Zhu, Z. J. *Anal. Chem.* **2016**, *88*, 8757−8764.
(34) Renaud, J. B.; Sabourin, L.; Topp, E.; Sumarah, M. W. *Anal. Chem.* **2017**, *89*, 2747−2754.
(35) Castro-Perez, J. M.; Kamphorst, J.; DeGroot, J.; Lafeber, F.; Goshawk, J.; Yu, K.; Shockcor, J. P.; Vreeken, R. J.; Hankemeier, T. *J. Proteome Res.* **2010**, *9*, 2377−2389.
(36) Gallart-Ayala, H.; Courant, F.; Severe, S.; Antignac, J. P.; Morio, F.; Abadie, J.; Le Bizec, B. *Anal. Chim. Acta* **2013**, *796*, 75−83.
(37) Malinovschi, A.; Janson, C.; Borres, M.; Alving, K. *J. Allergy Clin. Immunol.* **2016**, *138*, 1301−1308.
(38) Daham, K.; James, A.; Balgoma, D.; Kupczyk, M.; Billing, B.; Lindeberg, A.; Henriksson, E.; FitzGerald, G. A.; Wheelock, C. E.; Dahlen, S. E.; Dahlen, B. *J. Allergy Clin. Immunol.* **2014**, *134*, 306−313.
(39) Edmands, W. M.; Ferrari, P.; Scalbert, A. *Anal. Chem.* **2014**, *86*, 10925−10931.
(40) Gad, H.; Koolmeister, T.; Jemth, A. S.; Eshtad, S.; Jacques, S. A.; Strom, C. E.; Svensson, L. M.; Schultz, N.; Lundback, T.; Einarsdottir, B. O.; Saleh, A.; Gokturk, C.; Baranczewski, P.; Svensson, R.; Berntsson, R. P.; Gustafsson, R.; Stromberg, K.; Sanjiv, K.; Jacques-Cordonnier, M. C.; Desroses, M.; Gustavsson, A. L.; Olofsson, R.; Johansson, F.; Homan, E. J.; Loseva, O.; Brautigam, L.; Johansson, L.; Hoglund, A.; Hagenkort, A.; Pham, T.; Altun, M.; Gaugaz, F. Z.; Vikingsson, S.; Evers, B.; Henriksson, M.; Vallin, K. S.; Wallner, O. A.; Hammarstrom, L. G.; Wiita, E.; Almlof, I.; Kalderen, C.; Axelsson, H.; Djureinovic, T.; Puigvert, J. C.; Haggblad, M.; Jeppsson, F.; Martens, U.; Lundin, C.; Lundgren, B.; Granelli, I.; Jensen, A. J.; Artursson, P.; Nilsson, J. A.; Stenmark, P.; Scobie, M.; Berglund, U. W.; Helleday, T. *Nature* **2014**, *508*, 215−221.
(41) Yuan, M.; Breitkopf, S. B.; Yang, X.; Asara, J. M. *Nat. Protoc.* **2012**, *7*, 872−881.
(42) Bird, S. S.; Marur, V. R.; Sniatynski, M. J.; Greenberg, H. K.; Kristal, B. S. *Anal. Chem.* **2011**, *83*, 940−949.
(43) SANCO/11915/2015. Analytical quality control and method validation procedures for pesticide residues analysis in food and feed. https://ec.europa.eu/food/sites/food/files/plant/docs/pesticides_mrl_guidelines_wrkdoc_11945.pdf (accessed February, 15, 2016).
(44) Cajka, T.; Fiehn, O. *Anal. Chem.* **2016**, *88*, 524−545.
(45) Mass Spectrometry Metabolite Library of Standards (MSMLS). http://www.iroatech.com/userfiles/file/MSMLS%20information%20sheet_2017.pdf (accessed July 5, 2017).
(46) Building an UHPLC MSMS library for untargeted metabolomics using an Agilent 1290-6550 LC-QTOF instrument. http://mslibrary.barupal.org/ (accessed February 15, 2016).

DOI: 10.1021/acs.analchem.7b00925
Anal. Chem. 2017, 89, 7933−7942