analytical chemistry

Editors' Highlight

pubs.acs.org/ac

isoMETLIN: A Database for Isotope-Based Metabolomics

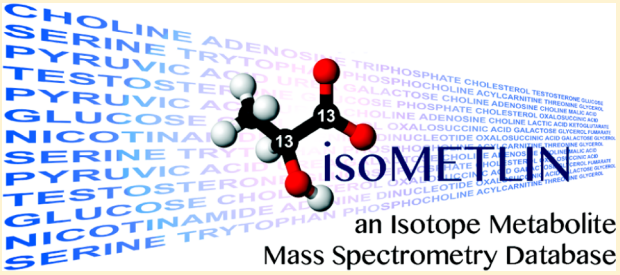
Kevin Cho,[†,‡] Nathaniel Mahieu,[†,‡] Julijana Ivanisevic,[§] Winnie Uritboonthai,[§] Ying-Jr Chen,[†,‡] Gary Siuzdak,*[,§] and Gary J. Patti*[,†,‡]

[†]Department of Chemistry, Washington University in St. Louis, St. Louis, Missouri 63130, United States

[‡]Departments of Genetics and Medicine, Washington University School of Medicine, St. Louis, Missouri 63110, United States

[§]Scripps Center for Metabolomics and Mass Spectrometry, The Scripps Research Institute, La Jolla, California 92037, United States

ABSTRACT: The METLIN metabolite database has become one of the most widely used resources in metabolomics for making metabolite identifications. However, METLIN is not designed to identify metabolites that have been isotopically labeled. As a result, unbiasedly tracking the transformation of labeled metabolites with isotope-based metabolomics is a challenge. Here, we introduce a new database, called isoMETLIN (http://isometlin.scripps.edu/), that has been developed specifically to identify metabolites incorporating isotopic labels. isoMETLIN enables users to search all computed isotopologues derived from METLIN on the basis of mass-to-charge values and specified isotopes of interest, such as ^{13}C or ^{15}N. Additionally, isoMETLIN contains experimental MS/MS data on hundreds of isotopomers. These data assist in localizing the position of isotopic labels within a metabolite. From these experimental MS/MS isotopomer spectra, precursor atoms can be mapped to fragments. The MS/MS spectra of additional isotopomers can then be computationally generated and included within isoMETLIN. Given that isobaric isotopomers cannot be separated chromatographically or by mass but are likely to occur simultaneously in a biological system, we have also implemented a spectral-mixing function in isoMETLIN. This functionality allows users to combine MS/MS spectra from various isotopomers in different ratios to obtain a theoretical MS/MS spectrum that matches the MS/MS spectrum from a biological sample. Thus, by searching MS and MS/MS experimental data, isoMETLIN facilitates the identification of isotopologues as well as isotopomers from biological samples and provides a platform to drive the next generation of isotope-based metabolomic studies.



Conventionally, mass spectrometry-based metabolomic technologies have been applied to compare the levels of hundreds to thousands of small molecules between sample groups.[1] For this standard application of metabolomics, isotopic labels are not required and therefore generally are not used. Recently, however, a variation of the standard metabolomic workflow has been introduced that is referred to as isotope-based metabolomics.[2,3] Here, a stable isotope is introduced into a biological system and metabolomic technologies are used to track its fate unbiasedly. It is important to note that, while both approaches use similar technologies to profile small molecules at the global level, the objective of each approach is different. In the standard application of metabolomics, the aim is to identify metabolites whose concentrations vary across samples. In isotope-based metabolomics, the objective is to determine the metabolic fates of a stable isotope introduced into a biological system.

In contrast to metabolic flux approaches in which isotope incorporation is measured in a small number of targeted metabolites, isotope-based metabolomics is comprehensive and untargeted in scope. As a result, metabolites incorporating label must be identified de novo. The power of such an experimental paradigm is that it facilitates the discovery of unexpected molecular transformations as well as alterations in pathways that are independent of metabolite pool sizes.[4−6]

To identify a metabolite that has incorporated a stable isotope requires special consideration. In the standard metabolomic workflow that does not involve isotopic labels, the mass-to-charge values of ions of interest are searched in metabolite databases such as METLIN.[7,8] The search returns putative matches on the basis of accurate mass. When a metabolite has been isotopically labeled, however, searching its mass-to-charge value in metabolite databases does not provide any valid matches because they do not include labeled metabolites. Thus, a comprehensive database of isotopically labeled metabolites is needed for isotope-based metabolomic studies.

Searching the accurate mass of an isotopically enriched compound in a database containing labeled metabolites has the potential to resolve isotopologues, but accurate mass alone is insufficient to distinguish isotopomers. Isotopologues are compounds that have different numbers of isotopic labels and therefore have unique mass-to-charge values. Isotopomers are compounds that have the same number of isotopic labels, but the position of each isotopic label within the structure is

Received: August 4, 2014
Accepted: August 28, 2014
Published: August 28, 2014

   dx.doi.org/10.1021/ac5029177 | *Anal. Chem.* 2014, 86, 9358−9361

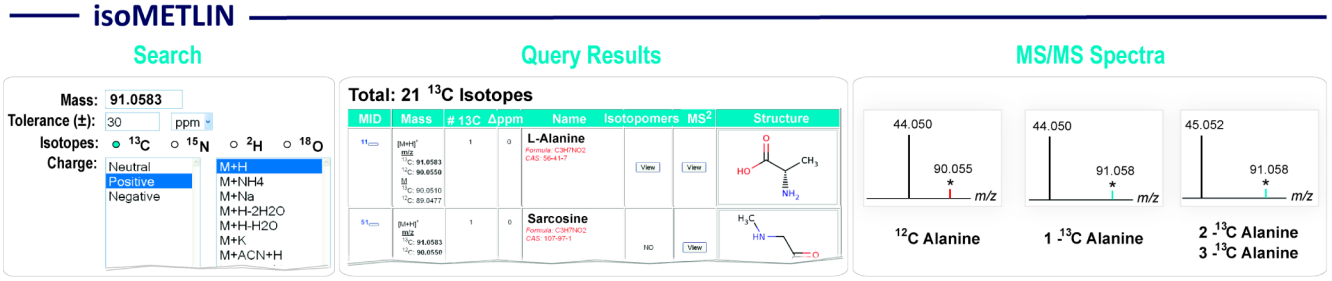


Figure 1. Screenshots of isoMETLIN interface. The isoMETLIN interface is analogous to METLIN. Users input accurate mass and specify isotopes of interest prior to searching. A list of putative matches is returned as shown. MS/MS data can be viewed when available.

unique. While identifying isotopologues can be biochemically informative, differentiating isotopomers generally provides greater insight into metabolic pathways.[9] When cells are labeled with 2,5-^{13}C-glucose, for example, citrate is labeled in the 5 position when pyruvate dehydrogenase is active and in the 3 position when pyruvate carboxylase is active.[10]

Indeed, similar types of isotopomer analyses have been applied with gas chromatography/mass spectrometry-based studies for flux modeling.[11] A major challenge in extending these modeling approaches to electrospray ionization-mass spectrometry (ESI-MS) is that the fragmentation patterns of small molecules analyzed by ESI-MS are poorly understood. Although the METLIN database has fragmentation patterns for over 12 000 metabolites, structural annotations for each detected fragment have only been predicted.[8,12] The reliability of the predictions has not been experimentally tested, and the mapping of parent atoms into fragments is largely unexplored. Generating theoretical MS/MS spectra for isotopically labeled metabolites with only the current data in METLIN therefore has major practical limitations.

Given the difficulty of predicting and interpreting the MS/MS spectra of metabolites analyzed by ESI-MS, the manual approach of cataloguing the experimental MS/MS spectra of metabolites in the METLIN database has fueled the development of the field of metabolomics. In the same sense, experimental analysis and cataloguing of metabolite isotopomers will be invaluable to the development of isotope-based metabolomics. Additionally, the experimental analysis of isotopomers will enable the annotation of MS/MS fragments and therefore will improve our understanding of metabolite fragmentation patterns in ESI-MS. In this work, we introduce isoMETLIN, which is a new database specifically designed to facilitate the identification of metabolites incorporating isotopic labels. While METLIN is optimized for standard metabolomics, isoMETLIN is the analogous comprehensive resource for isotope-based metabolomics. isoMETLIN enables the searching of ^{13}C, ^{15}N, ^{2}H, and ^{18}O isotopologues on the basis of accurate mass and also includes experimental MS/MS data for several hundred isotopically labeled metabolites that will facilitate the identification of isotopomers.

EXPERIMENTAL SECTION

Materials. High-performance liquid chromatography (HPLC)-grade solvents were obtained from Sigma-Aldrich (St. Louis, MO, USA). All isotopes were obtained from Cambridge Isotope Laboratories (Tewksbury, MA, USA).

Mass Spectrometry Analysis. Isotope standards were diluted to approximately 1 μg/mL in water, methanol, or a combination of each depending on the compound. Each standard was analyzed by using a HPLC system (1200 series, Agilent Technologies) interfaced to an ESI quadrupole time-of-flight (QTOF) mass spectrometer (6520 QTOF, Agilent Technologies). The standards were analyzed by flow injection using an isocratic 50% water and 50% acetonitrile mobile-phase composition at a flow rate of 20 μL/min. When analyzing the standards in positive mode, 0.1% formic acid was added to each mobile phase.

ESI source parameters were set as follows: gas temperature, 325 °C; drying gas, 5 L/min; nebulizer, 15 psi; fragmentor, 125 V; skimmer, 65 V; and capillary voltage, 3500 or −2500 V in positive or negative mode, respectively. A narrow isolation window (1.3 Da) was used to acquire MS/MS data for each standard at collision energies of 0, 10, 20, and 40 V.

Data Import. We generated isotopomer structures by using Marvin software 5.8.2. Raw data were processed by using an in-house R script for extracting and normalizing each MS/MS spectrum. Fragments with relative intensities less than 0.5% of the most intense fragment were excluded. The fragment intensities shown in isoMETLIN result from averaging the data over all scans acquired.

RESULTS AND DISCUSSION

Currently, there are no databases optimized to identify metabolites in liquid chromatography/mass spectrometry (LC/MS)-based metabolomic analyses that have been isotopically labeled. Yet, unbiased tracking of metabolic transformations with isotopes by LC/MS (i.e., isotope-based metabolomics) is an attractive experimental design that would greatly benefit from such a resource.[3] isoMETLIN, which is an analog of the widely used METLIN database, is designed to facilitate the identification of both isotopologues and isotopomers. This resource is meant to power studies in which the metabolites that are isotopically enriched are unknown.

Identifying Isotopologues. In isotope-based metabolomic studies, isotopologues represent metabolites that contain different numbers of stable isotopes. If a biological system is enriched with ^{13}C, for example, the number of isotopologues for any metabolite is equal to the number of carbon atoms that the metabolite has plus 1 (e.g., alanine $C_3H_5O_2$ has 4 carbon isotopologues). Given that each isotopologue has a unique isotopic composition, isotopologues can be resolved by nominal mass. With mass spectrometry, isotopologues are simpler to differentiate than isotopomers. To identify isotopologues with isoMETLIN, users can search in simple or batch mode. The isoMETLIN interface is analogous to that of METLIN. Briefly, the user inputs the accurate mass(es) of the ion(s) of interest in addition to specifying several other parameters (Figure 1). The error tolerance and polarity should be selected on the basis of the instrument in which the data

dx.doi.org/10.1021/ac5029177 | Anal. Chem. 2014, 86, 9358−9361

were acquired. The user can also choose which adducts they expect to detect (e.g., $[M + H]^+$, $[M + Na]^+$, etc.). For isotopic studies in particular, each adduct has the potential to greatly increase the number of putative hits. Thus, it is recommended that users pick only those adducts that are most likely to occur. The most important parameter to specify during the search is the stable isotope being used. At this time, isoMETLIN supports searches of ^{13}C, ^{15}N, ^{2}H, ^{18}O, or any combination of these atoms. isoMETLIN was populated by calculating the masses of all ^{13}C, ^{15}N, ^{2}H, and ^{18}O isotopologues for each compound in METLIN. The isoMETLIN search returns the METLIN ID, the mass of both the labeled and unlabeled metabolite, the number of stable isotopes included in the metabolite, the ppm error, the name of the metabolite, the metabolite's structure, and MS/MS data when available.

Identifying Isotopomers. The objective of isotope-based metabolomics is to determine the pathway through which a precursor is metabolized. In many cases, discriminating between alternative pathways can be accomplished by determining isotopologue distributions.[13] Other times, however, different pathways may produce the same isotopologue distribution but different isotopomer distributions.[14] For this reason, isotopomer distributions have been referred to as the ultimate amount of information on metabolite labeling.[9] Because the electron impact fragmentation patterns of central carbon metabolites are well characterized, isotopomer distributions are already routinely used in gas chromatography/mass spectrometry (GC/MS)-based studies to model flux through central carbon metabolism.[15] In contrast, investigations into isotopomer distributions with LC/MS have been limited due to our incomplete understanding of the fragmentation patterns produced from metabolites analyzed by ESI-MS. A major goal of our work here is to extend the analysis of isotopomer distributions to LC/MS, which is a platform applied extensively for metabolomics and is likely to provide additional unanticipated applications.

Unlike isotopologues, accurate mass is insufficient to resolve isotopomers. In isotope-based metabolomic studies, isotopomers represent metabolites that have the same number of isotopic labels but the arrangement of the labels within the metabolite's structure is different. As an example, 1-^{13}C-alanine, 2-^{13}C-alanine, and 3-^{13}C-alanine are isotopomers (Figure 1). Each has a single ^{13}C, but the location of the label within alanine is unique. Our approach for distinguishing isotopomers in LC/MS-based metabolomics is to use fragmentation patterns. By observing shifts in the mass of fragments, isotopic labels can be localized to specific carbon atoms within the structure of the metabolite. It should be noted that some isotopomers of interest may not be distinguished by standard MS/MS analysis since fragments typically contain more than a single atom. As shown in Figure 1, 2-^{13}C alanine and 3-^{13}C alanine have the same MS/MS fragments when analyzed in positive mode with a collision energy of 10 V. A combination of additional experiments involving negative-mode, various collision energies, and MS^n analysis may therefore be required to distinguish the positions of some isotopic labels. Positive- and negative-mode data as well as fragmentation patterns from various collision energies are already included in isoMETLIN. At this time, these data are available for hundreds of compounds through the same interface as METLIN. Investigators can support their identification of isotopomers by matching the fragmentation patterns of standards in isoMETLIN with those fragmentation patterns that they observe in research samples. Further improving the differentiation of isotopomers is the focus of future work.

Mapping Precursor Atoms to Fragments. A major challenge in constructing a database of experimental MS/MS data for metabolite isotopomers is that the number of isotopomers increases exponentially with carbon number. Moreover, many isotopomers are not commercially available. Thus, we examined strategies to accurately predict MS/MS isotopomer patterns on the basis of experimental MS/MS isotopomer data. To date, our efforts have focused largely on ^{13}C-labeled compounds because they are widely used and relevant to most biochemical pathways. However, a comparable approach is also being extended to other isotopic labels. The objective is to use experimental MS/MS data from ^{13}C-labeled isotopomers to map carbons in the parent ion to fragment ions. All precursor carbons can be tracked into fragments by analysis of each singly ^{13}C-labeled isotopomer. Once the precursor carbons are mapped to fragments for a given metabolite, then additional isotopomer MS/MS spectra can be computationally generated. For example, we analyzed each singly ^{13}C-labeled isotopomer of alanine. These include 1-^{13}C alanine, 2-^{13}C alanine, and 3-^{13}C alanine. From these MS/MS data, it was determined that when alanine is analyzed in positive mode the fragment occurring at m/z 44.050 contains carbon 2 and carbon 3 of alanine (Figure 1). Having measured only these isotopomers of alanine, we could then generate the MS/MS patterns of additional alanine isotopomers (Figure 2, top). As shown by comparing the calculated MS/MS spectra to those from experimental standards, this approach reliably generated MS/MS data for 2,3-^{13}C alanine and 1,2,3-^{13}C alanine. In addition to the hundreds of isotopomer standards for which we

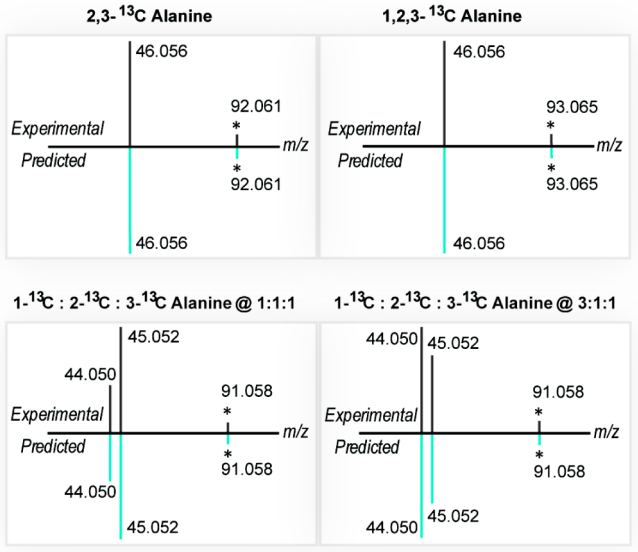


Figure 2. Calculating MS/MS data for isoMETLIN. All data are from positive mode at a collision energy of 10 V. (Top) Once precursor atoms are mapped to detected fragments by using isotopomer standards, additional isotopomer MS/MS patterns can be generated as shown for 2,3-^{13}C alanine and 1,2,3-^{13}C alanine. (Bottom) MS/MS data for pure isotopomers in isoMETLIN can be combined in different proportions to match experimentally acquired MS/MS data from a biological sample with multiple isotopomers. Physically mixing alanine isotopomer standards at different concentrations provided experimental MS/MS data that matched those data obtained from averaging pure isotopomer MS/MS spectra at the same ratios.

dx.doi.org/10.1021/ac5029177 | Anal. Chem. 2014, 86, 9358−9361

have already acquired experimental MS/MS data, isoMETLIN also contains computationally generated MS/MS isotopomer patterns based on this logic when possible. These MS/MS spectra are labeled as such within isoMETLIN. For metabolites other than alanine, depending on their fragmentation patterns, the minimum number of singly ^{13}C-labeled isotopomers that are needed to predict all of the MS/MS isotopomer spectra may be fewer than the number of carbons in the molecule.

Resolving Isotopomers in Biological Systems. In the standard metabolomic workflow, metabolite identifications are supported by comparing MS/MS data from the compound of interest in the biological sample to MS/MS data from a model compound.[16] The presence of additional fragments in the spectra of the research sample that are not detected in the spectra of the model compound prevents confirmation of the metabolite identity. These analyses can lead to false negatives when the metabolite of interest is not isolated purely in the collision cell and the MS/MS patterns therefore contain contaminating fragments from different molecules.[17] To limit the frequency of contaminated MS/MS spectra, compounds of interest are separated both chromatographically and by mass within the instrument. A challenge in analyzing labeled metabolites with the same isotopic composition is that these compounds cannot be separated by chromatography or by mass. For example, 1-^{13}C alanine, 2-^{13}C alanine, and 3-^{13}C alanine cannot be easily separated. Thus, if all three of these isotopomers were present in a sample, the experimental MS/MS data obtained would not match any of the MS/MS data in isoMETLIN for each pure isotopomer. Yet, it is likely that multiple isobaric isotopomers will be present in a sample. Determining the contributions of these isotopomers can be important to understanding a system's biochemical phenotype.

To facilitate distinguishing the contributions of metabolites that have the same isotopic composition, isoMETLIN has a mixing function. isoMETLIN allows users to average the MS/MS data from pure isotopomers to generate a mixed MS/MS spectrum that can be matched against the MS/MS spectrum of the sample. The MS/MS data from each pure isotopomer can be averaged together with different coefficients until the computationally mixed spectrum matches the experimental spectrum of the sample. As an example, we experimentally mixed isotopomer standards of 1-^{13}C alanine, 2-^{13}C alanine, and 3-^{13}C alanine in ratios of 1:1:1 and 3:1:1. This mixture of isotopomers is representative of what might be present in a biological sample. Each was measured experimentally, and the acquired MS/MS data were then compared to the MS/MS data computationally determined by averaging the MS/MS spectra of 1-^{13}C alanine, 2-^{13}C alanine, and 3-^{13}C alanine at 1:1:1 or 3:1:1 ratios (Figure 2, bottom). These data indicate that the ratios used to average the isoMETLIN spectra from pure standards approximate the ratios of the isotopomers in a sample.

■ CONCLUSIONS

Technologies to identify isotopically labeled metabolites unbiasedly have immense biochemical value. Currently, however, isotope-based metabolomic studies are limited because no database exists to search isotopologues and isotopomers. Here, we describe a new database called isoMETLIN to address these challenges. isoMETLIN, which is an analog of the widely used METLIN database, enables users to search ^{13}C, ^{15}N, ^{2}H, and ^{18}O isotopologues of all of the compounds in METLIN. isoMETLIN also contains experimental MS/MS spectra for hundreds of isotopomers at this time. From these experimental data, additional isotopomer MS/MS spectra are calculated, including averaged spectra that can be used to identify isotopomers that cannot be experimentally separated. In addition to facilitating the identification of labeled compounds in isotope-based metabolomic studies, the MS/MS data in isoMETLIN will help extend flux modeling to LC/MS analyses and ultimately will improve our overall understanding of the fragmentation of metabolites analyzed by ESI-MS.

■ AUTHOR INFORMATION

Corresponding Authors

*Phone: 314-935-3512. E-mail: gjpattij@wustl.edu.

*Phone: 858-784-9415. E-mail: siuzdak@scripps.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge Ron Trolard and Krista Backiel at Cambridge Isotope Laboratories for their generous support. Graphics were created by the authors. Copyright Patti Lab and Siuzdak Lab.

■ REFERENCES

(1) Patti, G. J.; Yanes, O.; Siuzdak, G. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 263−269.

(2) Huang, X.; Chen, Y. J.; Cho, K.; Nikolskiy, I.; Crawford, P. A.; Patti, G. J. *Anal. Chem.* **2014**, *86*, 1632−1639.

(3) Chen, Y. J.; Huang, X.; Mahieu, N. G.; Cho, K.; Schaefer, J.; Patti, G. J. *Biochemistry* **2014**, *53*, 4755−4757.

(4) Lane, A. N.; Fan, T. W.; Higashi, R. M.; Tan, J.; Bousamra, M.; Miller, D. M. *Exp. Mol. Pathol.* **2009**, *86*, 165−173.

(5) DeBerardinis, R. J.; Thompson, C. B. *Cell* **2012**, *148*, 1132−1144.

(6) Darpolor, M. M.; Basu, S. S.; Worth, A.; Nelson, D. S.; Clarke-Katzenberg, R. H.; Glickson, J. D.; Kaplan, D. E.; Blair, I. A. *NMR Biomed.* **2014**, *27*, 381−389.

(7) Kind, T.; Liu, K. H.; Lee do, Y.; DeFelice, B.; Meissen, J. K.; Fiehn, O. *Nat. Methods* **2013**, *10*, 755−758.

(8) Tautenhahn, R.; Cho, K.; Uritboonthai, W.; Zhu, Z.; Patti, G. J.; Siuzdak, G. *Nat. Biotechnol.* **2012**, *30*, 826−828.

(9) Christensen, B.; Nielsen, J. *Metab. Eng.* **1999**, *1*, 282−290.

(10) Yang, C.; Harrison, C.; Jin, E. S.; Chuang, D. T.; Sherry, A. D.; Malloy, C. R.; Merritt, M. E.; DeBerardinis, R. J. *J. Biol. Chem.* **2014**, *289*, 6212−6224.

(11) Zamboni, N.; Fendt, S. M.; Ruhl, M.; Sauer, U. *Nat. Protoc.* **2009**, *4*, 878−892.

(12) Wolf, S.; Schmidt, S.; Muller-Hannemann, M.; Neumann, S. *BMC Bioinf.* **2010**, *11*, 148.

(13) Shlomi, T.; Fan, J.; Tang, B.; Kruger, W. D.; Rabinowitz, J. D. *Anal. Chem.* **2014**, *86*, 1583−1591.

(14) Wiechert, W. *Metab. Eng.* **2001**, *3*, 195−206.

(15) Zamboni, N. *Curr. Opin. Biotechnol.* **2011**, *22*, 103−108.

(16) Want, E. J.; Masson, P.; Michopoulos, F.; Wilson, I. D.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Loftus, N.; Holmes, E.; Nicholson, J. K. *Nat. Protoc.* **2013**, *8*, 17−32.

(17) Nikolskiy, I.; Mahieu, N. G.; Chen, Y. J.; Tautenhahn, R.; Patti, G. J. *Anal. Chem.* **2013**, *85*, 7713−7719.