analytical chemistry

Article

pubs.acs.org/ac

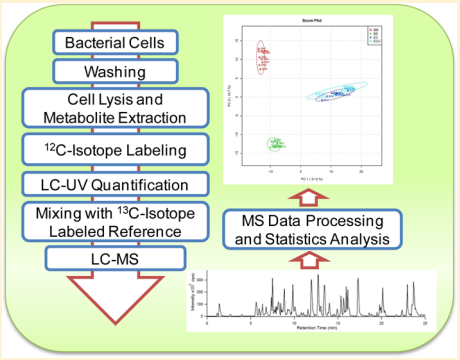
Development of Isotope Labeling Liquid Chromatography–Mass Spectrometry for Metabolic Profiling of Bacterial Cells and Its Application for Bacterial Differentiation

Yiman Wu and Liang Li*

Department of Chemistry, University of Alberta Edmonton, Alberta T6G 2G2, Canada

Supporting Information

ABSTRACT: Quantitative and comprehensive profiling of cellular metabolites is currently a challenging task in bacterial metabolomics. In this work, a simple and robust method for profiling the amine- and phenol-containing metabolome of bacterial cells is described. The overall workflow consists of methanol-based cell lysis and metabolite extraction with ultrasonication, differential isotope dansylation labeling of cellular metabolites, and analysis of the labeled metabolites by liquid chromatography–mass spectrometry (LC–MS). Over a thousand peak pairs or putative metabolites can be detected from bacterial cells in a 25 min LC–MS run and near 2500 putative metabolites can be found in one bacterium from combined results of multiple analyses. After careful examination and optimization of the sample preparation process, this method is shown to be effective for both Gram-positive and Gram-negative bacteria. An idea of applying LC–ultraviolet (UV) detection to quantify the total amount of labeled metabolites is shown to be effective for normalizing the amounts of metabolites present in different samples for metabolome comparison. The use of differential isotopic labeling allows relative quantification of each individual metabolite, which facilitates comparative metabolomics studies and the generation of a metabolic fingerprint of a bacterium. Finally, this method is demonstrated to be useful for the differentiation of three bacterial species in cultured media and spiked human urine samples.



Metabolomics can provide valuable and complementary information to the genomics, transcriptomics, and proteomics data.[1] In recent years, the study of microbial metabolomics has received growing research interest because of its potential applications in a wide range of microbial research fields, including metabolic engineering.[2] There are at least two major research areas in microbial metabolomics.[3] The first one is to investigate changes of the metabolic profile under different environmental conditions or look for the key metabolic changes in a mutant strain.[4] The second one is to generate a metabolic fingerprint of various bacterial species for the purpose of microorganism identification or differentiation.[3,5−7] In both cases, a robust and sensitive method capable of detecting and quantifying a large number of metabolites is desirable.

Among the various analytical platforms used for metabolome analysis,[8] liquid chromatography–mass spectrometry (LC–MS), particularly reversed phase LC–MS (RPLC–MS) has been widely used. However, RPLC–MS is not suitable for detecting very polar and ionic metabolites. To address this problem, our group has reported a $^{12}C/^{13}C$-dansylation-labeling technique for the analysis of the amine- and phenol-containing submetabolome; dansyl labeling allows separation of polar and ionic metabolites on RPLC while providing a signal enhancement of 10- to 1000-fold.[9] This method enables detection of hundreds of metabolites to over a thousand using one-dimensional LC–MS and has been shown to be useful for metabolic profiling of biofluids such as urine,[9] cerebrospinal fluid,[10] and saliva.[11]

However, differing from biofluid analysis, cellular metabolome profiling is a much more challenging task. Prior to metabolite analysis by NMR or MS, several sample preparation processes are needed, including (1) separation of cells from the growth medium; (2) a washing procedure to remove any interfering compounds from the cell surface; (3) rapid quenching to stop cellular activity; (4) extraction of intracellular metabolites, and (5) a disruption procedure to enhance the extraction process.[2] A number of studies have been reported with a focus on evaluating the sample preparation methods for cellular metabolome analysis.[3,12−21] However, the analytical performance of each method is very likely dependent on the type of metabolites analyzed and the detection method used. In this work, we describe a method based on differential isotope dansylation-labeling LC–MS, in combination with a fast step-gradient LC–UV quantification method for sample amount normalization,[22] for microbial metabolome profiling. This method allows the detection and quantification of thousands of putative metabolites from a bacterium with high precision and accuracy, compared to current methods of detecting less than 300 putative metabolites in general. As an

Received: February 4, 2013
Accepted: March 15, 2013

   dx.doi.org/10.1021/ac400330z | *Anal. Chem.* XXXX, XXX, XXX–XXX

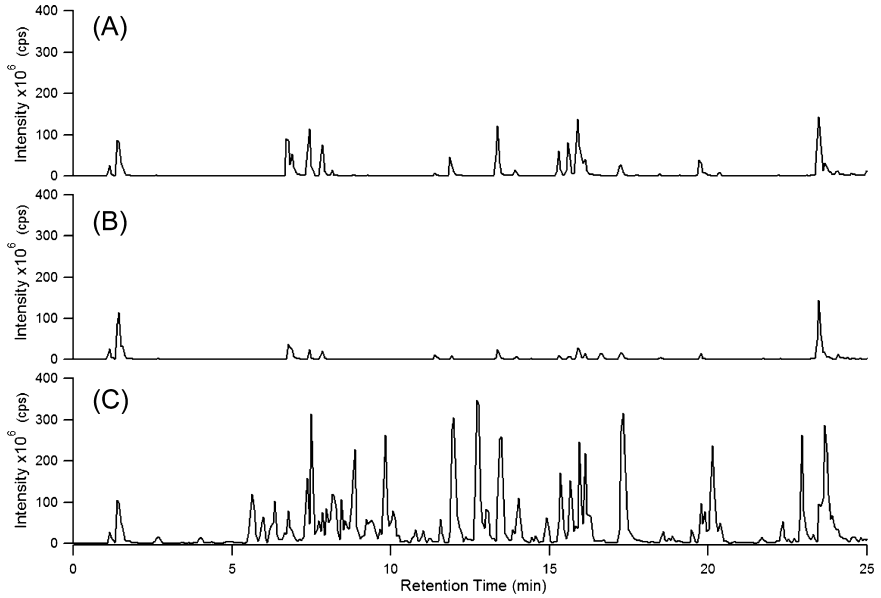


Figure 1. Base peak chromatograms of (A) the first wash solution, (B) the second wash solution, and (C) the metabolites extracted from *E. coli*.

example of applications, the possibility of employing this isotope labeling LC−MS method for identifying bacteria in clinical samples is assessed by analyzing bacteria spiked in human urine samples. To our knowledge, most reported studies of using metabolomics for bacterial identification were based on the use of GC−MS.[6,7,23]

EXPERIMENTAL SECTION

The Supplemental Note N1 of the Supporting Information provides information on chemicals and reagents used, dansylation-labeling chemistry, LC−UV quantification of labeled metabolites, and LC−MS experimental conditions.

Cell Culture and Harvest Conditions. For the method optimization work, *Escherichia coli* (ATCC 47076) cells were grown in nutrient broth (0.3% beef extract, 0.5% peptone) at 37 °C and 225 rpm in a shaking incubator for ∼24 h. All cultures were harvested at OD600 of 1.5 and were spun at 4640*g* at 4 °C for 10 min. The pellets were resuspended in 1 mL of ice-cold 0.9% NaCl, and spun in an Eppendorf 5415C microcentrifuge at 16000*g* at 4 °C for 1 min. The cell pellets were flash frozen in a dry ice/ethanol bath and stored at −80 °C until further use. For the cell differentiation study, *E. coli* (ATCC 47076) cells were grown at 37 °C on nutrient agar plates (0.3% beef extract, 0.5% peptone, 1.8% agar) for ∼24 h. *Bacillus subtilis* (ATCC 9372) and *Bacillus megaterium* (ATCC 14581) cells were grown at 30 °C on the same nutrient agar plates for ∼24 h. For the urine experiments, *E. coli* was grown overnight at 37 °C in LB medium (1% tryptone, 0.5% yeast extract) to an OD600 of ∼4 and then diluted to 1×10^5 cells/mL in each urine sample. Ten microliters of the spiked urine was diluted to 100 μL in water and then spread onto the Nutrient Agar plate and incubated overnight at 37 °C. Cells from the plates were scraped into 1 mL of 0.9% NaCl and rapidly centrifuged in an Eppendorf 5415C microcentrifuge at 16000*g* for 1 min. The cell pellets were resuspended in 1 mL of 0.9% NaCl and spun again. The final cell pellets were stored at −80 °C until further use.

Metabolite Extraction. The performance of three solvent systems was evaluated in this work: 50/50 MeOH/H_2O (MeOH), 50/50 ACN/H_2O (ACN), and 40/40/20 MeOH/ACN/H_2O (MAW). Each solvent extraction experiment was carried out in triplicate. For each extraction, the cell pellets were resuspended in 1 mL of the corresponding solvent system (0 °C for ACN and −20 °C for MeOH and MAW), and disrupted using ultrasonication, as described below. The resulting suspensions were centrifuged at 16000*g* for 10 min. The supernatants were dried using a SpeedVac and resuspended in 250 mL water. The resulting solutions were used for the labeling step.

Three cell disruption methods were compared, together with a control experiment (i.e., no disruption, CT). They were ultrasonication (SN), microwave (Mic), and freeze−thaw cycle (FT). Each disruption method, as well as the control experiment, was carried out in triplicate. Prior to disruption, the cell pellets were resuspended in 1 mL of 50/50 MeOH/H_2O. For the ultrasonication-assisted extraction, the cell suspensions were placed in a Branson ultrasonic cleaner 1510-MT (Branson Ultrasonics Corporation, Danbury, CT) with an ice bath for 10 min. Microwave-assisted extraction was performed for 10 min at 240 W power with a 1200 W microwave oven (Panasonic, Toronto, ON, Canada). For freeze−thaw cycle extraction, the cell suspensions were rapidly dipped into a liquid nitrogen bath for 30 s and thawed on ice for 1 min. This procedure was repeated 3 times. All of the resulting suspensions were centrifuged at 16 000*g* for 10 min. The supernatants were dried and resuspended in 250 mL water.

LC−MS and Data Analysis. The extracted metabolites were labeled by dansylation chemistry,[9] quantified by LC−UV,[22] and analyzed by LC−MS using a Bruker 9.4 T Apex-Qe Fourier transform ion-cyclotron resonance (FTICR) mass spectrometer (see the Supplemental Note N1 of the Supporting Information). The extracted peak-pair data from LC−MS were aligned by retention time and accurate mass, and only those peak-pair features shared by no less than 50% of the samples were retained for multivariate analysis. Heatmap comparison, ANOVA, and multivariate analysis were performed by Metaboanalyst[24] (www.metaboanalyst.ca), and the data were mean-centered and autoscaled (unit variance) prior to analysis.

dx.doi.org/10.1021/ac400330z | *Anal. Chem.* XXXX, XXX, XXX−XXX

RESULTS AND DISCUSSION

Effect of Cell Washing. Since we are interested in intracellular metabolite profiling, it is important to ensure that the metabolites detected are from inside of the cells rather than from the growth medium. In this work, after the cells were scraped off the dish, they were washed twice with 0.9% NaCl solution to remove any extracellular compound potentially stuck to the surface of the cells. The isotonic 0.9% NaCl solution was selected because it would not cause significant leakage[3] and was reported to be effective for quenching the cell metabolism.[19]

Figure 1 shows a comparison between the base peak chromatograms of the washing solutions and the *E. coli* cell extract. By comparing the first and second wash solutions (panels A and B in Figure 1), we can see a significant decrease in signal intensity for most of the peaks, suggesting that the washing step is effective at removing the extracellular metabolites from the cells while not causing observable cell lysis. A comparison between the second washing solution (Figure 1B) and the cell extracts (Figure 1C) reveals that there are fewer peaks detected in the washing solution with much lower intensities. Figure 1C shows that many metabolite peaks are observed and distributed across the entire gradient elution window. At a given retention time, a number of peak pairs from the differentially labeled metabolites can be observed in a mass spectrum (see Supplemental Note N2 of the Supporting Information for an example). In this particular case, 943 peak pairs were detected from the cell extract, while 65 peak pairs were found in the second washing solution. Among them, only 16 ion pairs have overlaps, which is less than 1.7% of the detectable peak pairs from the cell extracts and the intensities of these 16 ion pairs found in the washing solution are much lower than those from the cell extracts. These results indicate that almost all detected metabolites in the ion chromatograms shown in Figure 1C should have originated from the cells.

Metabolism Quenching. Although a quenching step is usually recommended before carrying out the extraction in order to stop further metabolism, the use of a quenching solution (typically 60% methanol) can often lead to metabolite leakage that can be as large as 60%.[16] The quenching procedure is more critical for studying metabolites with fast turnover rates, such as those involved in the energy metabolism (e.g., ATP) and glycolic pathways (e.g., glucose-6-phosphate).[25] Since our study focuses more on the general metabolic profile of the bacterial cells, rather than studying their metabolite fluxes, we decided not to include any additional quenching solutions in our experiment. Nevertheless, as noted before, we used a cold 0.9% NaCl solution to rapidly wash the cells, which could at the same time serve as the quenching step. Although the extent of metabolism quenching during the cell washing step is unknown, as will be shown below, the relative quantities of most of the putative metabolites detected in biological replicate experiments are reproducible, suggesting that any further metabolism during the sample handling process, if present at all, does not significantly affect the overall metabolic profiles. This may be due to the fact that, prior to metabolite extraction, the amine- and phenol-containing metabolites profiled by dansylation LC−MS do not undergo further metabolism extensively.

Comparison of Extraction Solvents. As both methanol (MeOH) and acetonitrile (ACN) have been reported to be the optimal solvent for extraction of intracellular metabolites for different types of cells,[19,20] we have compared these two solvents (1:1 organic:water) to see which one works better for extracting the amine- and phenol-containing submetabolome from bacterial cells, using *E. coli* as the model system. In addition, a combination of the two solvents in water (2:2:1 MeOH:ACN:H_2O, MAW) was also compared to see if a mixture of solvents can perform better than the use of a single solvent, with an expectation that a pure solvent and a solvent mixture may have different metabolite extraction and solubility properties. The performance of the extraction solvents was evaluated according to three criteria,[13] namely the number of peak pairs detected, the relative intensity of each peak pair, and the reproducibility of each extraction method. In this experiment, each individual sample was labeled with ^{12}C-dansyl chloride and a pooled sample was labeled with ^{13}C-dansyl chloride to serve as the internal standard.

The number of peak pairs detected in each extraction solvent is plotted in Figure 2A. More than 1000 peak pairs were detected for each extraction solvent, which is much more than around 300 peaks detectable in similar cellular samples reported.[12,13] Since only the labeled amine- and phenol-containing compounds can be picked up as peak pairs, the dansylation isotope-labeling LC−MS detection scheme eliminates artifacts from the instrument[26] and other interferences (e.g., impurities leached from the plastic container). Moreover, we have also used a built-in function in the peak extraction software to filter the peak pairs found in the method blanks, thereby eliminating contributions from dansyl products of solvents, reagents, and any impurities present therein. Therefore, each peak pair should represent a true metabolite. Since many of them were not identified, each peak pair detected is considered to be from a putative metabolite. Figure 2A shows that ACN and MeOH extractions gave a similar number of peak pairs. The use of solvent combination (MAW), on the other hand, was not capable of extracting as many metabolites as the use of ACN or MeOH.

The relative intensity of each peak pair was calculated by taking the ^{12}C/^{13}C ratio (i.e., ratio of each individual sample to the pooled internal standard). Only those peaks commonly detected across all three extraction solvents were used for comparing the relative intensity. Figure S1 of the Supporting Information shows the number distributions of the peak pairs detected within a solvent (triplicate) and among the three solvent systems. From the combined results of all 9 runs, 2381 unique peak pairs or putative metabolites were detected. Among them, 851 peak pairs were found in all three extraction solvents. Figure 2B shows the relative intensity of 10 selected amino acids with different types of side-chains (i.e., hydrophobic, polar, acidic, and basic), and Figure 2C compares the relative intensity of 10 other compounds encompassing a variety of classes (e.g., purine and pyrimidine derivatives, amino acid derivatives, dipeptide, amino sugar). These metabolites were identified based on the accurate mass and retention time matches with those of the dansyl-labeled standard compounds. It can be seen that in most cases, the results from these three solvents were comparable, with MeOH performing slightly better on average. For some compounds, the performance of MeOH extraction was significantly better than ACN or MAW (e.g., pyridoxal 5′-phosphate, adenosine), but it is difficult to observe specific trends on which class of compounds are more favorable in each solvent.

In addition to comparing the relative intensity of representative compounds, we have also carried out a heatmap

dx.doi.org/10.1021/ac400330z | *Anal. Chem.* XXXX, XXX, XXX−XXX

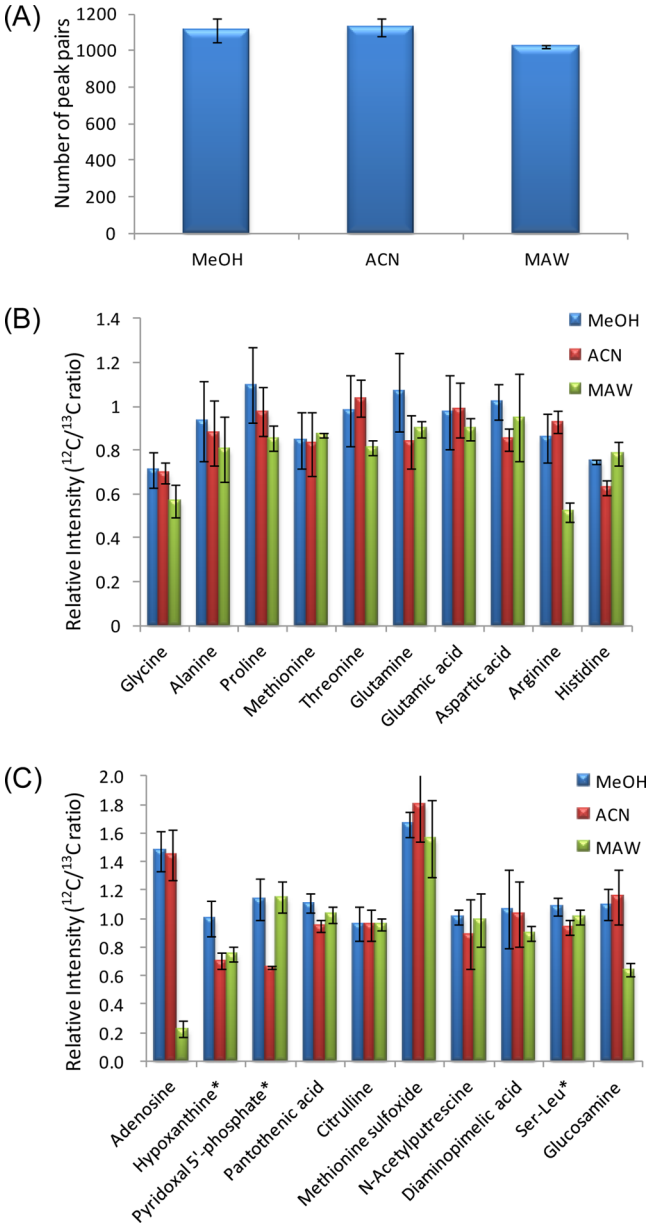


Figure 2. Comparison of the (A) average number of peak pairs detected, (B) relative intensities of 10 amino acids, and (C) relative intensities of 10 other selected compounds extracted by the three solvent systems (all ratios were referenced to the same pooled sample). The MS data was preprocessed to exclude peaks with intensities of less than 50000 counts (i.e., $S/N < 20$). Compounds labeled with an asterisk (*) indicates level 2 identification (see text). All the other compounds were definitively identified (level 1).

comparison of all commonly detected peak pairs in order to avoid any bias. The heatmap comparison results are shown in Figure S2A of the Supporting Information, where more red-colored features indicate higher signal intensities. It was found that MeOH extraction in general could give higher intensities for most of the peak pairs detected, compared to ACN or MAW extractions. These results indicate that MeOH is a better extraction solvent in terms of the relative intensity of the extracted metabolites.

To compare the reproducibility of different solvents, box plots showing the distributions of relative standard deviations of peak pairs were constructed and they are given in Figure S3A of the Supporting Information. For all three solvents, the relative standard deviations for the majority of peak pairs were below 20%, indicating a good reproducibility. The mean values (represented by the dots) were 12.6%, 10.2%, and 9.8% for MeOH, ACN, and MAW, respectively. The results from the *t* test of these values indicate that ACN and MAW were not significantly different, while the mean value of MeOH was different from those of ACN and MAW. Thus, ACN and MAW gave slightly better reproducibility than MeOH. However, the reproducibility of all of these methods should be sufficient for most metabolic-profiling applications.

The above results indicate that the three solvent systems studied detected a similar number of peak pair with a similar number reproducibility. In terms of signal reproducibility, ACN and MAW were slightly better than MeOH. However, MeOH was more effective in getting higher amounts of compounds. Since the purpose of the work is to do metabolic profiling of bacterial cells, it is desirable to get as high yields as possible from the extraction process. Therefore, MeOH was chosen as the extraction solvent for all the subsequent experiments.

Comparison of Sample Disruption Methods. In addition to the selection of an appropriate extraction solvent, it is also important to use an effective disruption method to facilitate metabolite extraction into the extraction solvent. In this work, four protocols were evaluated: no disruption (or control, CT), freeze–thaw cycle (FT), ultrasonication (SN), and microwave (Mic). In many of the cellular metabolomics studies reported,[12,18] either freeze–thaw cycle or no cell disruption was applied. Ultrasonication has also been used, but mostly with a sonicator containing a metal tip. However, this way of ultrasonication is very time-consuming as only one sample can be processed at a time and may cause cross-contamination if the tip is not washed thoroughly. In our work, we used an ultrasonic cleaner, which is capable of handling multiple samples without direct contact with the sample. The use of microwave has also been demonstrated to be efficient for extracting metabolites from biological samples.[27,28]

The same three criteria have been applied to evaluate the performance of the sample disruption methods. Figure 3A shows the number of peak pairs detected in each method. Among these four protocols, ultrasonication and microwave gave slightly higher numbers of peak pairs. Comparison of the relative peak intensity was also carried out on the commonly detected peak pairs (i.e., 834 out of a total of 2484 peak pairs from 12 runs). The relative intensities of 20 selected compounds are shown in panels B and C of Figure 3. The heatmap showing the comparison of relative intensities of all the peak pairs commonly detected in the four disruption methods is shown in Figure S2B of the Supporting Information. Note N3 of the Supporting Information describes the results of relative intensity comparisons. It is apparent that ultrasonication gave higher intensities for most of the compounds, followed by the microwave. The performance of the freeze–thaw cycle was very similar to the control, which is not surprising considering that the control samples were flash frozen and stored under −80 °C before the extraction, and therefore cell lysis is expected to occur to some extent through this freeze–thaw process. The heatmap shown in Figure S2B of the Supporting Information indicates that the extraction with no disruption or freeze–thaw gave low amounts for the majority of the peak pairs. Applying a disruption procedure appears to be essential in order to achieve a high efficiency of extraction. Our results also suggest that the commonly used

dx.doi.org/10.1021/ac400330z | *Anal. Chem.* XXXX, XXX, XXX–XXX

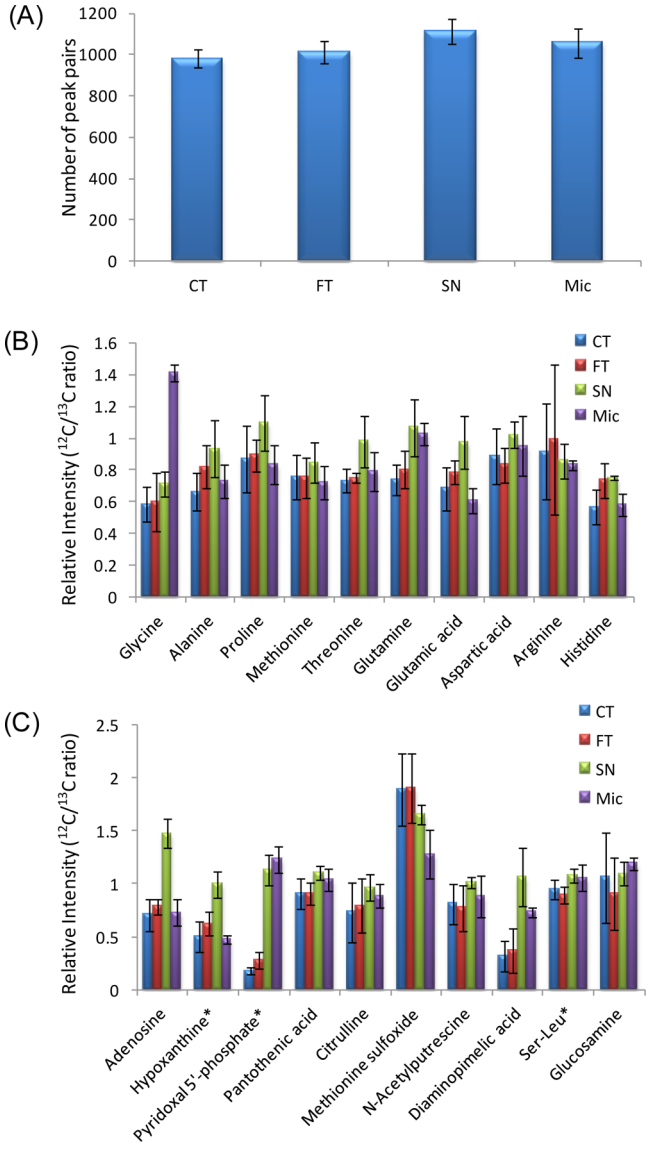


Figure 3. Comparison of the (A) average number of peak pairs detected, (B) relative intensities of 10 amino acids, and (C) relative intensities of 10 other selected compounds extracted in 1:1 MeOH:H_2O in combination with one of the four disruption methods (all ratios were referenced to the same pooled sample).

freeze−thaw cycle procedure is insufficient for metabolite extraction. The performance of ultrasonication and the microwave was similar, indicating that both methods can be used to effectively break the cells and facilitate metabolite extraction into the extracting solvent.

The reproducibility of the four disruption methods was also examined and the results are shown in Figure S3B of the Supporting Information. For most of the metabolites, the relative standard deviation of peak intensity from triplicate experiments was below 20%. The mean values for CT, FT, SN, and Mic were 15.2%, 16.2%, 12.9%, and 12.6%, respectively. Therefore, in terms of intensity reproducibility, SN and Mic were found to be better disruption methods than CT and FT. We also compared the reproducibility of the number of peak pairs detected among these methods (see Figure S4 of the Supporting Information). The average number of peak pairs detected was 981 ± 46, 1012 ± 50, 1111 ± 61, and 1058 ± 70 for CT, FT, SN, and Mic, respectively. There was no significant difference in terms of peak pair reproducibility. Based on the overall consideration of the three criteria examined, it can be concluded that the use of SN or Mic is a better choice to facilitate the metabolite extraction. Because SN is very convenient to do, compared to Mic, we used ultrasonication with an ultrasonic cleaner to perform solvent extraction for the subsequent experiments.

Bacterial Differentiation: Sample Amount Normalization. The application of the extraction-labeling method described above is demonstrated on the differentiation of three different bacterial cells, namely *E. coli* (EC), *B. subtilis* (BS) and *B. megaterium* (BM), based on their metabolic profiles. These bacteria were chosen as the model system because they include both Gram-positive (*B. subtilis* and *B. megaterium*) and Gram-negative (*E. coli*) species. Each cell was analyzed in a total of 9 biological replicates, including 3 interday replicates and 3 intraday replicates on each day. In addition to the biological replicates, one additional experimental replicate was carried out for each of the three bacteria on each day. These replicate experiments were designed to demonstrate the robustness of this method, which is an important prerequisite for studying the differences in metabolic profiles.

Since the cells were scraped off the plates, it is difficult to control as well as count the number of cells harvested. For fair comparison of the metabolic profiles of the same or different cells, sample amount normalization among the different comparative samples is needed. We recently reported a normalization method based on the use of dansylation-labeling LC−UV for determining the total amount of the labeled metabolites in a sample, followed by adjusting the sample volume of individual samples to mix with a control sample (i.e., a pooled sample from several individual samples).[22] In this work, we applied this method to normalize the metabolite amounts among the comparative cell samples.

The quantification results are shown in Table S1 of the Supporting Information, which illustrates that although the sample amount within each cell type was similar in most cases, a variation of as large as 2-fold could still be observed. Table S1 of the Supporting Information also shows that the amount of metabolites extracted from *B. subtilis* and *B. megaterium* was considerably larger than that extracted from *E. coli*. This is due to the larger number of cells harvested because *B. subtilis* and *B. megaterium* grow better in the growth medium (nutrient agar) than *E. coli*. In this work, to compare the metabolic profiles of different cell species, we took the same total amount of the labeled metabolites from each sample and assumed that any difference observed from their metabolic profiles solely came from differences in the abundances of individual metabolites. As will be discussed later, this assumption was proven to be valid for the purpose of differentiating cell types.

Bacterial Differentiation: Data Analysis. To produce a reference sample from which all the individual samples can be compared to, aliquots of individual extracts from the three different cells were mixed to generate a pooled sample, which was subsequently labeled by ^{13}C-dansyl chloride. An equal amount of ^{12}C-dansyl-labeled individual sample and ^{13}C-dansyl-labeled pooled reference sample was mixed, followed by LC−MS analysis. The peak ratios of individual peak pairs found in the mass spectra were calculated. Since the same reference sample was used, the peak ratios of an individual peak pair obtained from different samples reflected the concentration differences of the putative metabolite in these samples.

dx.doi.org/10.1021/ac400330z | *Anal. Chem.* XXXX, XXX, XXX−XXX

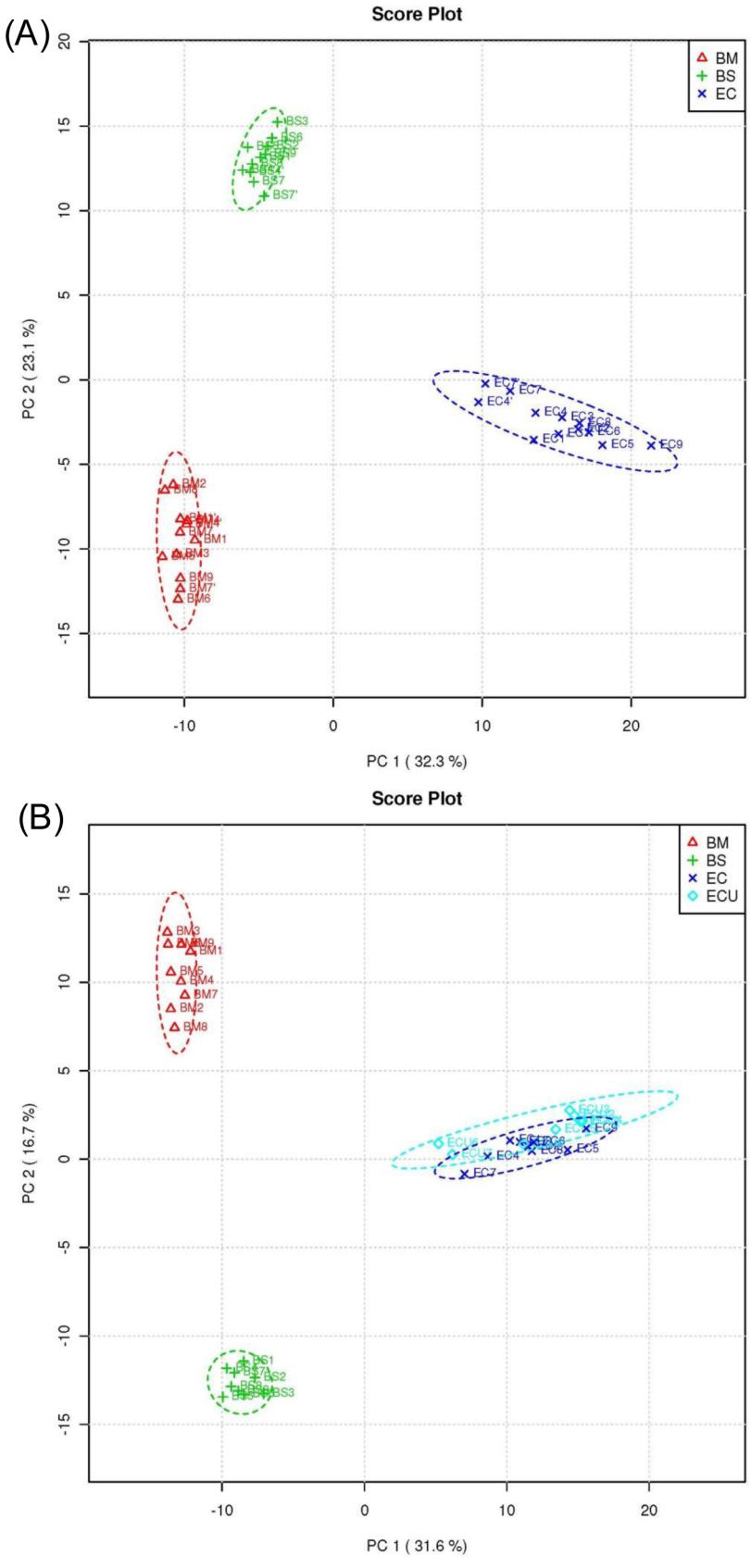


Figure 4. PCA score plots for (A) three standard bacterial cultures and (B) three standard bacterial cultures plus the urinary bacteria. *B. megaterium* (BM), in red; *B. subtilis* (BS), in green; *E. coli* (EC), in blue; and *E. coli* from urine (ECU), in light blue. For each species, nine biological replicates and three experimental replicates were presented in (A), while only the nine biological replicates were shown in (B).

dx.doi.org/10.1021/ac400330z | *Anal. Chem.* XXXX, XXX, XXX−XXX

Among the comparative samples, there are a total of 704 peak pairs or putative metabolites with ratios detectable in at least 50% of the samples. These ratios were used for the principal component analysis (PCA), and the resulting PCA score plot is shown in Figure 4A. The score plot clearly demonstrates that the three different bacteria can be well-separated. Although *B. subtilis* and *B. megaterium* belong to the same Bacilli genus, there are still substantial differences in their metabolic profiles, which are reflected by the second principle component. These results indicate the potential of this method for differentiating bacteria at the species level. Of course, the specificity and applicability of the method for differentiating a great variety of bacteria species and strains, including clinically relevant microorganisms, requires further investigation.

The distribution of the intensity ratio of the peak pairs detected from each sample that were used for the PCA analysis is shown in Figure S5 of the Supporting Information (only the data from biological replicates were plotted). It is interesting to note that the intensity ratio distributions of all nine *E. coli* samples were larger than those of *B. megaterium* and *B. subtilis*, while the distributions of *B. megaterium* and *B. subtilis* were very similar. The median of the intensity ratios for each sample ranges from 0.67 to 1.56 with a mean value of 1.06. Thus, many of the metabolites have similar ratios, which is expected, as they are likely involved in essential metabolisms common to different cell types. The most important variables that contribute to the differentiation were determined by PLS-DA analysis. Table T2 of the Supporting Information lists the top 20 metabolites (i.e., with the highest VIP scores), together with their *p* values obtained from ANOVA analysis. The metabolites were identified by matching the accurate mass and retention time with authentic standards (level 1) or our standard library[29] (level 2) or by searching the accurate mass against the human metabolome database (HMDB)[30] (level 3). Table T2 of the Supporting Information also shows the average $^{12}C/^{13}C$ ratio of these metabolites in each species, with the % RSD included in parentheses. We can see that the % RSD is generally in the range of 10−30%, indicating that a good reproducibility was obtained. However, we note that for some compounds (e.g., compound **16**), the % RSD is considerably large and a detailed examination of the data variations reveals that the large %RSD is mainly attributed to the biological variations (data not shown). Therefore, compounds like this one should not be selected as a discriminator, if one or several individual metabolites are used for differentiating the cell type.

From the above results of the median peak ratio analysis of individual metabolites and the PCA/PLS-DA analysis of the metabolic profiles, we can conclude that the proposed sample amount normalization method can be used for bacterial metabolic comparison to reveal subtle differences. We noticed that some of the metabolite distribution patterns among the three different cells can be rationalized. For example, diaminopimelic acid, a key component of the bacterial cell wall that is incorporated into the peptidoglycan structure of Gram-negative bacteria and Gram-positive bacilli,[31] has been detected in all three species, with the ratios in the two Gram-positive species being markedly larger than those in *E. coli* (see Figure S6 of the Supporting Information). This can be explained by considering that Gram-positive bacteria have thicker cell wall structures and that peptidoglycan is highly abundant in Gram-positive bacteria cell wall.[31,32] We also found that this ratio is larger in *B. megaterium* than in *B. subtilis*, possibly because of the larger size of *B. megaterium*. For many other compounds (e.g., glutamine), it is difficult to provide a simple explanation for the different amounts present in the three species because glutamine is involved in several metabolic pathways. However, as long as the amount found in each bacteria species is very consistent and that they are significantly different from the other species (as indicated by the *p* values), this compound can still serve as a good marker for bacterial identification. Thus, by matching the ratios of many of these marker metabolites obtained in an unknown sample with the data generated from a known species, we would be able to identify the unknown species with high confidence.

Of course, the applicability of this approach of bacterial identification depends on the presence of an informative library containing specific panels of metabolite-markers from a wide variety of bacteria species that are of interest to a practical application. For example, for diagnosis of bacterial infection, a metabolite-marker library of bacteria found in clinical samples may be constructed. This strategy is similar to the use of protein-markers for bacterial identification that is done mainly using matrix-assisted laser desorption ionization (MALDI) MS[33−35] and has been adapted in some clinical diagnosis laboratories.[36−38] Metabolite-based analysis may offer an alternative or complementary tool to the protein-based method for bacterial differentiation or identification.

Analysis of Bacteria in Human Urine. In addition to the presence of a library of metabolite markers, a reliable identification process also depends on the consistent detection of the bacterial metabolic profiles in various real samples. In order to investigate whether a single bacterial strain isolated from a real biological sample can give a similar metabolic profile as the standard cultures, we spiked 1×10^5 cells/mL of *E. coli* cultured in LB medium into 1 mL of human urine samples and spread the urine sample onto the nutrient agar plates to isolate pure *E. coli* strains, as a model system to mimic a real clinical sample (denoted as ECU). This was done with three healthy individuals' urine samples, each with three replicate experiments. The use of urine samples from different people can provide us an insight into whether differences in urine would have an effect on the metabolic profile of the bacteria. The isolated *E. coli* strains were then extracted, labeled, and analyzed in the same way as the standard bacteria cultures. The resulting data were compared with the data obtained from the three standard cultures using PCA, as shown in Figure 4B. It can be seen that the *E. coli* strains isolated from spiked urine samples can still be clustered close together with the standard *E. coli* cultures, and that they can be clearly separated from the two *Bacilli* "false strains". In addition, there is no significant distinction between the *E. coli* strains obtained from the three different urine samples, indicating that the composition of urine would not affect the bacterial metabolite profile.

The PCA score plot is useful for visual inspection of the data. However, in order to obtain a more confident conclusion on the similarity of the "real sample bacteria" to the standard bacteria cultures, a comparison of the ratios of all commonly detected peak pairs was carried out. For a total of 454 commonly detected peak pairs, we assigned a score of 1 to the closest match and a score of 0 to the other two nonmatches for each peak pair. This procedure gave a score of 83 for *B. megaterium*, 89 for *B. subtilis*, and 282 for *E. coli*. The much higher score for *E. coli* is another indication that this urinary bacteria can be differentiated from the "false strains" and thus be correctly identified.

dx.doi.org/10.1021/ac400330z | *Anal. Chem.* XXXX, XXX, XXX−XXX

As an example, Table T3 of the Supporting Information shows the ratios of 10 definitively identified metabolites (level 1) detected in the four bacteria strains, from which we can see that for all these 10 metabolites, the average ratios found in the urinary bacteria were closest to *E. coli* and were significantly different from that of *B. subtilis* and *B. megaterium*. The % RSDs for different urinary bacteria samples were all below 30%, indicating again that the metabolic profile of bacterial cells was largely independent of the urine composition.

The success of identifying the bacteria strain in the spiked urine samples is an important illustration on the prospect of our method to practical applications, which requires sample-independent, unbiased analyses. Our hypothesis was that regardless of the source of the bacteria, by growing the cells on the same medium under exactly the same conditions, similar metabolic profiles could be obtained. The results of the current experiment support this hypothesis, leading to the possibility of using this method for real sample analysis which we will undertake in the future.

Detection Sensitivity. In this study, we have not carried out the specific experiments to determine the sensitivity of this method, but we note that cells grown under the conditions indicated in the Experimental Section for a period of 24 h, which corresponds to the order of 1×10^9 cells, could generate sufficiently high signals (i.e., more than 1000 peak pairs detected with $S/N > 20$). We also noticed that although *E. coli* was used as the model species for method optimization, the peak intensity for the majority of peaks were on the same order of magnitude for all three bacteria species, indicating that this extraction-labeling method works equally well for Gram-positive and Gram-negative species.

CONCLUSION

A method for quantitative metabolome profiling of bacterial cells with relatively high metabolome coverage has been developed. Application of this method was successfully demonstrated on the differentiation of three different bacteria species, as well as identification of bacterial cells spiked in urine samples. The possibility of applying the current method in clinical applications has been discussed, which will be our future research focus. We envisage that another important application of this method is in the area of biological metabolomics, where specific changes of the metabolome within cells after exposing to different environmental conditions can be probed for functional studies of cellular metabolisms and networks.

ASSOCIATED CONTENT

Supporting Information

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: Liang.Li@ualberta.ca.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by Genome Canada, the Natural Sciences and Engineering Research Council of Canada, and the Canada Research Chairs program. We thank Gareth Lambkin (Biological Services, University of Alberta) for his assistance in cell culture. Y.M.W. also thanks the Alberta Ingenuity for the graduate scholarship.

REFERENCES

(1) Fiehn, O. *Comp. Funct. Genomics* **2001**, *2*, 155−168.
(2) Liebeke, M.; Dorries, K.; Meyer, H.; Lalk, M. In *Functional Genomics: Methods and Protocols*, 2nd ed.; Kaufmann, M., Klinger, C., Eds.; Humana Press Inc: Totowa, 2012; Vol. *815*, pp 377−398.
(3) Marcinowska, R.; Trygg, J.; Wolf-Watz, H.; Mortiz, T.; Surowiec, I. *J. Microbiol. Methods* **2011**, *87*, 24−31.
(4) Wachsmuth, C. J.; Almstetter, M. F.; Waldhier, M. C.; Gruber, M. A.; Nurnberger, N.; Oefner, P. J.; Dettmer, K. *Anal. Chem.* **2011**, *83*, 7514−7522.
(5) Fan, T. W. M.; Bird, J. A.; Brodie, E. L.; Lane, A. N. *Metabolomics* **2009**, *5*, 108−122.
(6) Olivier, I.; Loots, D. T. *J. Microbiol. Methods* **2012**, *88*, 419−426.
(7) Cevallos-Cevallos, J. M.; Danyluk, M. D.; Reyes-De-Corcuera, J. I. *J. Food Sci.* **2011**, *76*, M238−M246.
(8) Lenz, E. M.; Wilson, I. D. *J. Proteome Res.* **2007**, *6*, 443−458.
(9) Guo, K.; Li, L. *Anal. Chem.* **2009**, *81*, 3919−3932.
(10) Guo, K.; Bamforth, F.; Li, L. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 339−347.
(11) Zheng, J. M.; Dixon, R. A.; Li, L. *Anal. Chem.* **2012**, *84*, 10802−10811.
(12) Winder, C. L.; Dunn, W. B.; Schuler, S.; Broadhurst, D.; Jarvis, R.; Stephens, G. M.; Goodacre, R. *Anal. Chem.* **2008**, *80*, 2939−2948.
(13) Shin, M. H.; Lee, D. Y.; Liu, K. H.; Fiehn, O.; Kim, K. H. *Anal. Chem.* **2010**, *82*, 6660−6666.
(14) Yanes, O.; Tautenhahn, R.; Patti, G. J.; Siuzdak, G. *Anal. Chem.* **2011**, *83*, 2152−2161.
(15) Rabinowitz, J. D.; Kimball, E. *Anal. Chem.* **2007**, *79*, 6167−6173.
(16) Bolten, C. J.; Kiefer, P.; Letisse, F.; Portais, J. C.; Wittmann, C. *Anal. Chem.* **2007**, *79*, 3843−3849.
(17) Sellick, C. A.; Hansen, R.; Maqsood, A. R.; Dunn, W. B.; Stephens, G. M.; Goodacre, R.; Dickson, A. J. *Anal. Chem.* **2009**, *81*, 174−183.
(18) Villas-Boas, S. G.; Hojer-Pedersen, J.; Akesson, M.; Smedsgaard, J.; Nielsen, J. *Yeast* **2005**, *22*, 1155−1169.
(19) Dietmair, S.; Timmins, N. E.; Gray, P. P.; Nielsen, L. K.; Kromer, J. O. *Anal. Biochem.* **2010**, *404*, 155−164.
(20) Maharjan, R. P.; Ferenci, T. *Anal. Biochem.* **2003**, *313*, 145−154.
(21) Smart, K. F.; Aggio, R. B. M.; Van Houtte, J. R.; Villas-Boas, S. G. *Nat. Protoc.* **2010**, *5*, 1709−1729.
(22) Wu, Y. M.; Li, L. *Anal. Chem.* **2012**, *84*, 10723−10731.
(23) Havlicek, V.; Lemr, K.; Schug, K. A. *Anal. Chem.* **2013**, *85*, 790−797.
(24) Xia, J. G.; Psychogios, N.; Young, N.; Wishart, D. S. *Nucleic Acids Res.* **2009**, *37*, W652−W660.
(25) Weibel, K. E.; Mor, J. R.; Fiechter, A. *Anal. Biochem.* **1974**, *58*, 208−216.
(26) Mathur, R.; O'Connor, P. B. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 523−529.
(27) Stout, S. J.; daCunha, A. R.; Picard, G. L.; Safarpour, M. M. *J. Agric. Food Chem.* **1996**, *44*, 3548−3553.
(28) Young, J. C. *J. Agric. Food Chem.* **1995**, *43*, 2904−2910.
(29) Guo, K.; Peng, J.; Zhou, R. K.; Li, L. *J. Chromatogr., A* **2011**, *1218*, 3689−3694.
(30) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M. A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, D. D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; MacInnis, G. D.; Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L. *Nucleic Acids Res.* **2007**, *35*, D521−D526.
(31) Royet, J.; Dziarski, R. *Nat. Rev. Microbiol.* **2007**, *5*, 264−277.
(32) Van Heijenoort, J.; Elbaz, L.; Dezelee, P.; Petit, J. F.; Bricas, E.; Ghuysen, J. M. *Biochemistry* **1969**, *8*, 207−213.

(33) Fenselau, C.; Demirev, P. A. *Mass Spectrom. Rev.* **2001**, *20*, 157−171.
(34) Toh-Boyo, G. M.; Wulff, S. S.; Basile, F. *Anal. Chem.* **2012**, *84*, 9971−9980.
(35) Chen, W. J.; Tsai, P. J.; Chen, Y. C. *Anal. Chem.* **2008**, *80*, 9612−9621.
(36) Kok, J.; Chen, S. C. A.; Dwyer, D. E.; Iredell, J. R. *Pathology* **2013**, *45*, 4−17.
(37) Khot, P. D.; Couturier, M. R.; Wilson, A.; Croft, A.; Fisher, M. A. *J. Clin. Microbiol.* **2012**, *50*, 3845−3852.
(38) Tan, K. E.; Ellis, B. C.; Lee, R.; Stamper, P. D.; Zhang, S. X.; Carroll, K. C. *J. Clin. Microbiol.* **2012**, *50*, 3301−3308.