



RESEARCH ARTICLE

Qualitative Metabolome Analysis of Human Cerebrospinal Fluid by ^{13}C -/ ^{12}C -Isotope Dansylation Labeling Combined with Liquid Chromatography Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Kevin Guo,¹ Fiona Bamforth,² Liang Li¹

¹Department of Chemistry, University of Alberta, Edmonton, Alberta, T6G 2G2, Canada

²Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Alberta, Canada

Abstract

Metabolome analysis of human cerebrospinal fluid (CSF) is challenging because of low abundance of metabolites present in a small volume of sample. We describe and apply a sensitive isotope labeling LC-MS technique for qualitative analysis of the CSF metabolome. After a CSF sample is divided into two aliquots, they are labeled by ^{13}C -dansyl and ^{12}C -dansyl chloride, respectively. The differentially labeled aliquots are then mixed and subjected to LC-MS using Fourier-transform ion cyclotron resonance mass spectrometry (FTICR MS). Dansylation offers significant improvement in the performance of chromatography separation and detection sensitivity. Moreover, peaks detected in the mass spectra can be readily analyzed for ion pair recognition and database search based on accurate mass and/or retention time information. It is shown that about 14,000 features can be detected in a 25-min LC-FTICR MS run of a dansyl-labeled CSF sample, from which about 500 metabolites can be profiled. Results from four CSF samples are compared to gauge the detectability of metabolites by this method. About 261 metabolites are commonly detected in replicate runs of four samples. In total, 1132 unique metabolite ion pairs are detected and 347 pairs (31%) matched with at least one metabolite in the Human Metabolome Database. We also report a dansylation library of 220 standard compounds and, using this library, about 85 metabolites can be positively identified. Among them, 21 metabolites have never been reported to be associated with CSF. These results illustrate that the dansylation LC-FTICR MS method can be used to analyze the CSF metabolome in a more comprehensive manner.

Key words: Cerebrospinal fluid, Isotope labeling, Metabolite identification, Database search, Accurate mass measurement, LC-MS, Metabolomics

Electronic supplementary material The online version of this article (doi:10.1007/s13361-010-0033-4) contains supplementary material, which is available to authorized users.

Correspondence to: Liang Li; e-mail: Liang.Li@ualberta.ca

Introduction

Cerebrospinal fluid (CSF) is an important body fluid for discovery of potential biomarkers of diseases, particularly those related to malfunctions of central nervous system (CNS), because of its proximity to the neuropathol-

Received: 23 September 2010

Revised: 8 November 2010

Accepted: 9 November 2010

Published online: 27 January 2011

ogy site in the brain [1, 2]. Recently, proteomics and metabolomics tools have been applied for the analysis of proteome and metabolome of CSF [3, 4]. However, metabolome analysis of CSF remains to be an analytical challenge because of low concentrations of many metabolites present in CSF and the availability of only small volumes of samples. Through literature mining of more than 2000 books and journal articles, a total of 329 metabolites have been found to be detectable in various studies of human CSF (www.csfmetabolome.ca) [5]. Among them, only about 75 metabolites were reported to have concentrations of above 1 μM . Analyzing metabolites in CSF is mostly done by using NMR [6–10], GC-MS [10, 11], and LC-MS [10, 12–14]. In a report that compares different techniques for human CSF metabolome analysis [5], it was shown that the NMR technique could positively identify about 53 metabolites, GC-MS could identify 41 metabolites and LC-MS using Fourier-transform ion cyclotron resonance (FTICR) MS could identify 17 metabolites. The combination of the three techniques identified a total of 70 different metabolites.

Among different analytical techniques available for CSF metabolome analysis, LC-MS is a very promising technique for improving metabolite detectability, thereby expanding the metabolome coverage. However, it, presently, has some limitations. Low concentrations of metabolites (less than 1 μM) are generally difficult to analyze without special sample enrichment in a typical un-targeted metabolome profiling work. Another major problem is related to metabolite identification. Targeted analysis of certain known metabolites in CSF can be carried out by LC-MS using sensitive methods such as multiple reaction monitoring (MRM) or selected reaction monitoring (SRM) in a tandem mass spectrometer. For example, 23 metabolites of neurotransmitters were analyzed using MRM MS/MS [15]. However, there are only a few reports of untargeted metabolome analysis of CSF by LC-MS [12–14, 16–19]. Because of low abundance of metabolites present in CSF, less than 5000 features could be detected from LC-MS and many of the features were likely from non-metabolite signals. Only a few metabolites could be identified in a LC-MS run.

We have recently developed an isotope labeling chemistry based on dansylation reaction that tags amine- and phenol-containing metabolites [20]. Unlike other chemical derivatization schemes where the main purpose was to introduce a mass tag to the metabolites for quantitative metabolite or metabolome analysis [21–28], dansylation can not only effectively introduce an isotope tag for accurate metabolite quantification, but also improve the chromatographic retention of the metabolites in reversed-phase LC and enhance the efficiency of electrospray ionization (ESI). In our previous work, sensitive detection of metabolites from human urine samples was demonstrated. In particular, we observed a significant enhancement of detection signals (up to three orders of magnitude) in LC-MS after dansylation of metabolites. In this work, we report our studies of applying

this strategy for analyzing amine- and phenol-containing metabolites in a more challenging biofluid, CSF, and demonstrate that the method of isotope labeling via dansylation can be used for profiling the metabolome of CSF with more extended coverage than previously possible.

Experimental

Chemicals and Reagents

All chemicals and reagents were purchased from Sigma-Aldrich Canada (Markham, ON, Canada) except those otherwise noted. $^{13}\text{C}_2$ -dimethyl sulfate that was used to synthesize ^{13}C -dansyl chloride was purchased from Cambridge Isotope Laboratories (Andover, MA, USA). LC-MS grade of water, methanol, and acetonitrile (ACN) were purchased from ThermoFisher Scientific (Edmonton, AB, Canada).

Lumbar CSF samples were collected from patients screened for meningitis in accordance with guidelines established by the University of Alberta Health Research Ethics Board. As part of the disease screening procedure, CSF samples were required to be stored at 4°C for 2 d, after which they were placed in a freezer for long-term storage at –80°C. For this work, CSF samples from four individuals were processed by adding LC-MS grade acetonitrile in 1:1 (vol/vol) ratio, and then stored in a –80 °C freezer for further use.

Synthesis of ^{13}C -Dansyl Chloride

The synthesis of ^{13}C -dansyl chloride as a derivatizing reagent was reported in previous work [20]. The purity and confirmation of ^{13}C -dansyl chloride was tested against the commercial ^{12}C -dansyl chloride using LC-FTICR MS and LC-UV. The purity of ^{13}C -dansyl chloride is over 99% based on the LC-MS and LC-UV results. NMR was also used to further confirm the identity and purity of the synthesized ^{13}C -dansyl chloride.

Dansylation Labeling Reaction

About 50 μL of the CSF in acetonitrile were mixed with an equal volume of sodium carbonate/sodium bicarbonate buffer (0.5 mol/L, pH 9.4) in reaction vials. 50 μL of freshly prepared ^{12}C -dansyl chloride solution (20 mg/mL) (for light labeling) or ^{13}C -dansyl chloride (20 mg/mL) (for heavy labeling) were then added. The dansylation reaction was allowed to proceed for 60 min at 60 °C with shaking at 150 rpm. After 60 min, 20 μL of methylamine (0.5 mol/L) were added to the reaction mixture to consume the excess dansyl chloride and quench the dansylation reaction. After an additional 30 min of 60 °C incubation, the ^{13}C -labeled mixture was combined with the ^{12}C -labeled counterpart for LC-MS analysis.

LC-FTICR MS Measurement

The HPLC system was an Agilent 1100 series binary system (Agilent, Palo Alto, CA, USA) and was modified to reduce extra system solvent volume according to an Agilent protocol (Agilent publication number: 5988-2682EN). A reversed-phase Agilent Eclipse plus C_{18} column (2.1×100 mm, $1.8 \mu\text{m}$ particle size, 95 \AA pore size) was purchased from Agilent Canada (Mississauga, ON, Canada). LC Solvent A was 0.1% (vol/vol) LC-MS grade formic acid in 5% (vol/vol) LC-MS grade ACN, and Solvent B was 0.1% (vol/vol) LC-MS grade formic acid in LC-MS grade acetonitrile. The gradient elution profile was as follows: $t=0$ min, 20% B; $t=3.0$ min, 35% B; $t=16$ min, 65% B; $t=18.6$ min, 95% B; $t=21$ min, 95% B; $t=21.3$ min, 98% B; $t=23.0$ min, 98%B; $t=24.0$ min, 20%B. The flow rate was $150 \mu\text{L}/\text{min}$. The flow from RPLC was split in 1:3 and a $50 \mu\text{L}/\text{min}$ flow was loaded to a sample injector and the electrospray ionization (ESI) source of a Bruker 9.4-Tesla Apex-Qe FTICR mass spectrometer (Bruker, Billerica, MA, USA) or an Applied Biosystems, QStar Pulsar i mass spectrometer while the rest of flow was delivered to the waste. All MS spectra were obtained in the positive ion mode. The QStar Pulsar i LC-MS system was only used for detecting individual dansylated standards for the development of the dansylation library. For each standard, the system was used to generate the MS spectra of the labeled product to assess the purity of the product and labeling efficiency. All of the mass spectral data presented in this work were obtained using the 9.4-Tesla FTICR mass spectrometer.

Results and Discussion

Figure 1 shows the work flow for generating a qualitative metabolome profile from a CSF sample. Briefly, a sample is divided into two equal aliquots, one labeled with the heavy or ^{13}C -dansyl-chloride reagent and another one labeled with the light or ^{12}C -reagent. The labeled aliquots are mixed and then injected into LC-FTICR MS for analysis. FTICR MS offers high resolution and high accuracy measurement of the metabolite masses [29]. Thus, the ^{13}C -/ ^{12}C -labeled ion pairs can be picked based on their characteristic mass differences as well as perfect co-elution of the pairs due to the absence of an isotope effect on chromatography separation. The detected ion pairs can be searched against a database for putative metabolite identification or a library of standards for definitive identification. In this work, putative identification is based on the accurate mass matches of the dansylated metabolites with the human endogenous metabolites found in the Human Metabolome Database (HMDB) (about 8000 metabolite entries) [30]. Definitive identification is based on matches of accurate masses and retention times to a ^{13}C -/ ^{12}C -labeled authentic standards library (220 standards; see below).

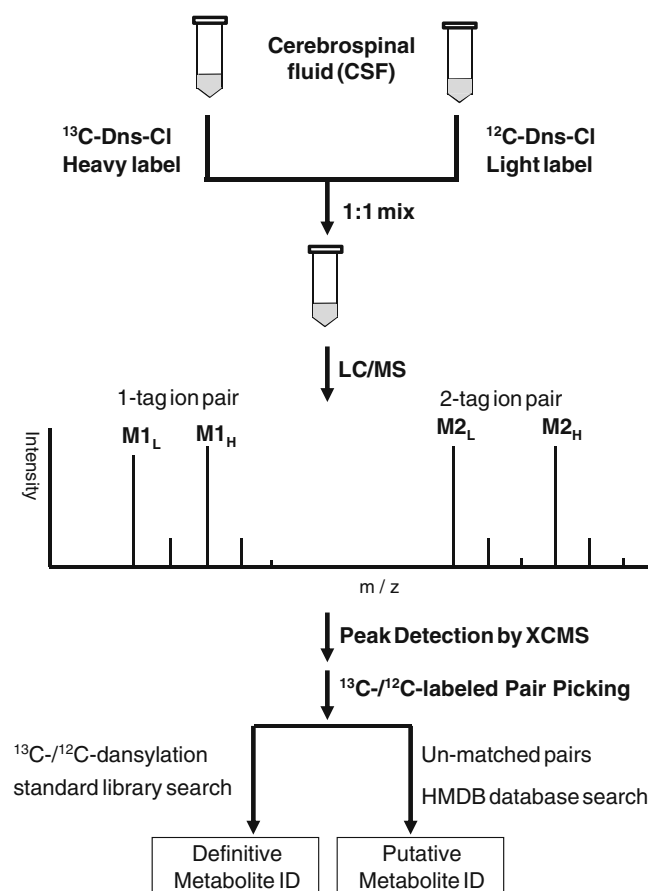


Figure 1. Workflow for CSF sample analysis and LC-MS data processing using the ^{12}C -/ ^{13}C -dansylation strategy

There are several important features of using the above strategy for metabolite identification. First of all, dansylation derivatization improves metabolite detection by enhancing ESI efficiency and improving chromatography retention and separation. This is illustrated in Figure 2 where it shows the base peak ion chromatograms of two reversed-phase (RP) LC-MS runs from the injections of the same amount of ^{13}C -

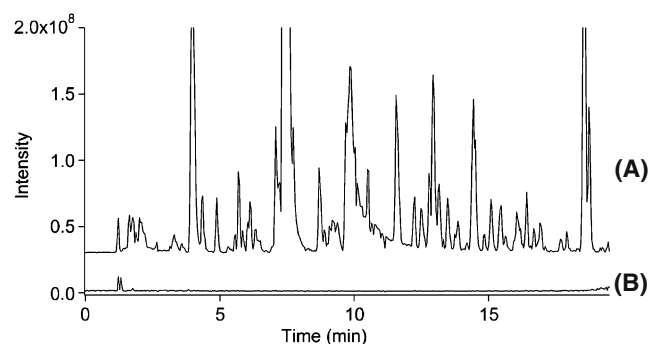


Figure 2. Base peak ion chromatograms of (A) 1:1 ^{12}C -/ ^{13}C -dansylated CSF sample #1 and (B) non-derivatized CSF sample #1 obtained by using LC-FTICR MS. In both cases, the sample amount injected was equivalent to about $0.5 \mu\text{L}$ of the original CSF sample

^{12}C -dansylated CSF (A) and un-labeled CSF (B). The injection amount was equivalent to 0.5 μL of the original CSF sample. As Figure 2b shows, without any sample pre-concentration, injection of 0.5 μL of un-labeled CSF on a 2.1-mm RP column hardly generates any MS signals. This chromatogram was found to be the same as that of blank injection; the two peaks shown at the beginning of the chromatogram are from the background, not the metabolites in the sample. In contrast, many peaks with signal-to-noise ratios between 3 and 4500 are observed in Figure 2a and they are distributed along the chromatographic elution profile. This is consistent with a previous study where signal enhancement of one to three orders of magnitude was found in ESI MS analysis of dansylated metabolites over their un-labeled counterparts and the labeled metabolites were better retained and separated in the RP column [20].

The signal enhancement can be attributed to several factors. One is related to the increased propensity of being charged for the labeled amines and phenols due to the presence of the dimethylamine moiety attached to the aromatic ring of the tag where a tertiary amine can be readily formed. The labeled compound has higher hydrophobicity than its unlabeled counterpart, making it easier to stay on the surface of the droplets during ESI. An elution solvent with higher organic solvent content, where a labeled compound is eluted out, as opposed to the unlabeled one eluted at void or high water content solvent, also enhances the ionization efficiency. On the chromatographic separation, the addition of the dansyl group containing a hydrophobic aromatic ring to a polar and hydrophilic amine or phenol increases hydrophobicity, thereby enhancing its retention on a RP column.

The sensitivity enhancement by dansylation derivatization is particularly important in analyzing CSF samples because of the limited amount of samples collectable from a patient compared to other body fluids such as blood and urine. The volume or amount of CSF samples available would be even smaller for biological model systems such as rats. Because no metabolite peaks were detected from the injection of the unlabeled CSF, all the metabolite peaks shown in Figure 2a that were generated from the injection of an equivalent volume of the labeled CSF should be from the amine and phenol derivatives. The second important feature is that $^{13}\text{C}/^{12}\text{C}$ -dansylated metabolites are perfectly co-eluted in RPLC and always detected as pairs in the same mass spectra. Thus, we can readily automate the process of peak picking and peak pairing (and relative quantification based on differential labeling of two comparative samples, although this is not the focus of the present work). Using isotope ion pairs has been employed to distinguish true metabolite signals from many other peaks detected in LC-MS [20, 31–33]. In this work, the accurate mass difference between heavy-labeled and light-labeled compounds indicates the presence of primary and secondary amines or phenol compounds as well as the number of reactive functional groups. Error in mass difference is the mass error between the theoretical mass difference and the measured

mass difference for $^{13}\text{C}/^{12}\text{C}$ -labeled ion pairs. The theoretical mass difference for one tag and singly charged ion pair is 2.00671 and two-tag singly charged ion pair is 4.01342. Error in mass difference was used as a key criterion to assign a $^{13}\text{C}/^{12}\text{C}$ -labeled ion pair. As an example, Figure 3a shows a singly charged spectrum of 1:1 molar ratio of $^{13}\text{C}/^{12}\text{C}$ -dansylated isoforms of 5-hydroxyindolacetic acid (5-HIAA), a physiologically important metabolite of serotonin [34]. The $^{13}\text{C}/^{12}\text{C}$ -labeled ion pair with error in mass difference of 0.47 ppm and matched retention time [difference in retention time = 2.4 s; variation of retention time in general is $<\pm 15$ s (see Supplemental Tables S1–S4)] to standards ensures the identity of 5-HIAA (see below). The spectral pattern clearly indicates only one reactive functional group exists. Figure 3b shows a doubly-charged spectrum of histidine and Figure 3c shows the corresponding singly charged spectrum. The errors in mass difference are 0.32 ppm for the doubly charged ion pair of histidine and 0.43 ppm for the singly charged ion pair. The difference in retention time to histidine standard is 1.8 s. The accurate mass differences and unique spectra patterns shown in Figure 3b and c unambiguously reveal that histidine has two reactive functional groups. It is interesting to note that the doubly charged ion pairs from the dansylated derivatives with two tags often exhibit very high signal intensity in the mass spectra. This may be due to the readiness of forming doubly charged ions by protonation of both amine groups on the two tags. In these cases, both doubly charged and singly charged ion pairs can be used to validate the presence of the same metabolite.

Another feature is related to the high stability of the dansylated metabolites. Dansylated compounds hardly fragment in-source or during the transfer to the mass analyzer in LC-MS. This enhances the molecular ion intensity and makes it less ambiguous in identifying the molecular ion peaks. Finally, dansylation increases the molecular ion mass by 234 Da for one tag and 467 Da for two tags, which effectively shifts their mass-to-charge ratios out of the low-mass region of a mass spectrum that typically contains more background noise from common contaminants and solvent clusters.

Using the workflow shown in Figure 1, a typical LC-MS run of a 1:1 $^{13}\text{C}/^{12}\text{C}$ -dansylated CSF sample generated about 14,000 peak features detected by the automated XCMS peak picking software [35]. Obviously, not all the features belong to the real signals from the metabolites. Isotopic ions, adducts, and fragment ions as well as multiply charged ions were treated as separated peak features in XCMS [17]. To pick the $^{13}\text{C}/^{12}\text{C}$ -dansylated ion pairs, the peak features in the XCMS output table were exported into Excel for processing. Because the error in mass difference between the $^{13}\text{C}/^{12}\text{C}$ -dansylated ion pair was found to be generally less than 2 ppm, we used 2 ppm error in mass difference as the first criterion to pick the ion pairs. The second criterion was based on the fact that the $^{13}\text{C}/^{12}\text{C}$ -dansylated pairs were perfectly co-eluted in RPLC; therefore, $^{13}\text{C}/^{12}\text{C}$ -labeled ion pairs must be shown in a same spectrum. At the last step in pair picking, only the

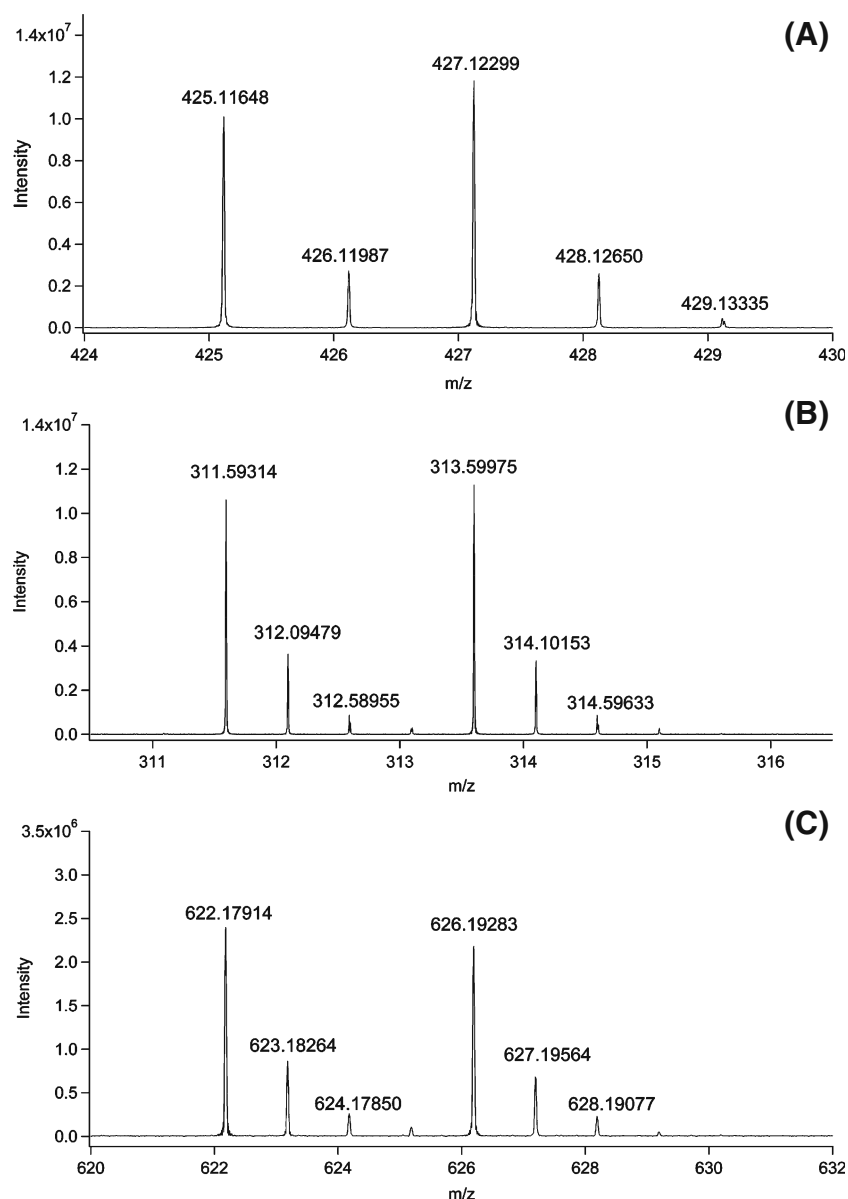


Figure 3. Expanded molecular ion regions of the mass spectra of (A) singly charged, one-dansylation-tag ion pair from 5-HIAA, (B) doubly charged, two-dansylation-tag ion pair from histidine, and (C) singly charged, two-dansylation-tag ion pair from histidine

protonated ion pair peaks were retained, while the ion pairs corresponding to isotopic peaks, common adduct ions, multimers, and multiply charged ions were eliminated. Nonreactive metabolites, interference ions, background ions, instrument noises, and electronic noises will not form $^{13}\text{C}/^{12}\text{C}$ -ion pairs with characteristic mass differences. For a given CSF sample, approximately 500 ion pairs can be confirmed to be from the $^{13}\text{C}/^{12}\text{C}$ -dansylated metabolites (see below).

The reproducibility for detecting the amine- and phenol-containing metabolites using dansylation in CSF samples has been examined. In total, four different CSF samples were analyzed, with each sample analyzed twice following the workflow shown in Figure 1, i.e., two experimental replicates (not repeat injections of the same mixture) were done on each sample. The results are summarized in Supplemental

Tables S1–S4 with each table listing the compound name (if known), retention time, masses of ion pair, and ion pair signal intensity. Figure 4a shows the comparison of the number of ion pairs detected from replicate runs of individual samples. The average number of ion pairs detected from the four samples ranges from 473 to 572 with an average of 519. These ion pairs were detected in the LC gradient time window between 1.6 and 24.5 min. There are 405 (71%), 353 (66%), 346 (69%), and 324 (68%) common ion pairs detected for samples 1 to 4, respectively. In addition, the signal intensities of common ion pairs from run to run are reproducible. The CV of the two dataset for samples 1 to 4 was found to be 3.2%, 5.7%, 6.8%, and 4.4%, respectively. Comparing the common ion pairs detected from the four samples, we found that 261 ion pairs were consistently detected (see Figure 4b), representing 64% of

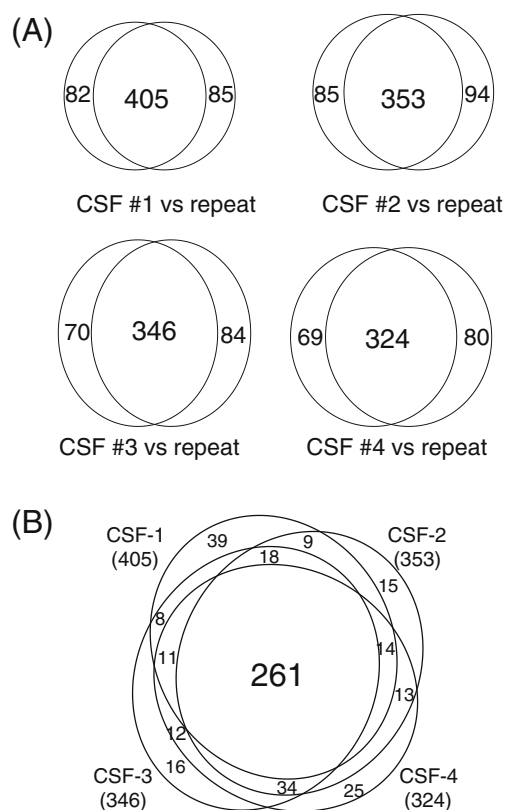


Figure 4. Comparison results of the common ion pairs detected among the four samples

the 405 ion pairs found in CSF-1, 74% of the 353 pairs in CSF-2, 75% of the 346 pairs in CSF-3, and 81% of the 324 pairs in CSF-4. There were 15 to 39 unique ion pairs found in individual samples. If we combined all the results obtained from the four samples, there were 1132 unique ion pairs found. The unique ion pairs found in the individual samples are likely the results of abundance differences; only the metabolites with the concentrations above the detection limit (low nanomolar) would be detected by dansylation LC-FTICR MS. These results also show the diversity of metabolites potentially present in human CSF. It should be noted that the low overlap of metabolites found in replicate runs or between samples is likely the result of under-sampling in one-dimensional-LC MS, much like in the shotgun proteome analysis where an overlap of 60% to 70% is commonly observed in replicate runs of peptides generated from trypsin digestion of whole cell lysates [36]. It remains to be seen whether multidimensional LC separation of metabolites prior to MS analysis can increase the number of common metabolites detected in replicates.

While the use of ion pairing allows the differentiation of the mass spectral peaks originated from the true metabolites *versus* those from other sources, accurate mass measurement of the molecular ions by FTICR MS offers the possibility of putative metabolite identification based on mass matches with a database of known metabolites. While several databases of

chemical compounds are available, we focused on our efforts on using the HMDB for putative identification of metabolites in CSF, as this database is composed of about 8000 human metabolites reported in the literature, as opposed to other database where many types of chemical compounds, in addition to the endogenous human metabolites, are enclosed. The results of the HMDB database search using the molecular masses of ion pairs (minus the dansyl group) are shown in Supplemental Tables S5–S8. For the 1132 unique ion pairs found in the four samples combined, 785 pairs (69%) do not match with any metabolites in the database, while 347 pairs (31%) match with one or more putative metabolites in HMDB database within ± 3 mDa in mass tolerance. Among the 347 matches, 90 pairs (26%) match with one metabolite. Because dansylation reaction targets primary and secondary amines as well as phenol groups in metabolites, we can use this information to eliminate some of the matches, i.e., matches with metabolites without an amine or phenol group are deemed to be false positive. Overall, there are 334 pairs (29%) matched with at least one metabolite in HMDB from the 1132 unique ion pairs.

To provide definitive identification of the ion pairs or metabolites detected in CSF, we used an in-house library of dansylation standards and compared the molecular ion masses and retention times of the standards with those of unknown metabolites. Our current library consists of 220 amine- and phenol-containing metabolites, which is almost double the number of standards we previously reported [20]. To build this library, for each standard, a dansylation reaction was carried out and the product was examined by LC-MS. Most compounds gave high conversion yields ($>90\%$). A list of these compounds along with their RPLC retention time and measured masses by LC-FTICR MS are shown in Supplemental Table S9. To minimize ion suppression, we divided the 220 standards into six groups (30–40 compounds per group), largely based on their retention properties and avoidance of closely eluted isomers on RPLC. The standards within a group were mixed and then ^{13}C - or ^{12}C -dansylated, followed by 1:1 mixing of the ^{13}C -/ ^{12}C -labeled standard mixtures, and injecting them into LC-FTICR MS for analysis. The accurate masses and retention time of each ion pair were determined from these experiments and the data shown in Supplemental Table S9 were used to compare with those found in the CSF samples run under the same LC and column conditions for positive metabolite identification.

In total, 85 metabolites are positively identified from the four samples combined, and they are shown in Table 1. The results of metabolites identified in individual CSF samples are listed in Supplemental Tables S10–S13. As these tables show, 76, 65, 70, 68 metabolites were reproducibly detected in the repeated differential labeling experiments from CSF samples 1 to 4, respectively. As an example, Supplemental Table S10 shows a list of identified metabolites from the replicate experiments of 1:1 ^{13}C -/ ^{12}C -labeled CSF sample #1. There are 76 common metabolites positively identified in both experiments. Only four metabolites in the first labeled experiment and three

Table 1. Combined list of identified metabolites from the replicate experiments of 1:1 ^{13}C -/ ^{12}C -labeled CSF samples #1–4 (the experimental data are shown in Supplemental Tables S10–S13)

Compound name	Compound name	Compound name
Phosphoethanolamine	r-Aminobutyric acid	L-ornithine
3-Methylhistidine	Hypoxanthine	Acetaminophen/or 4-acetamidophenol
Glucosamine	5-Hydroxymethyluracil	Acetaminophen/or 4-acetamidophenol
Taurine	3-Aminoisobutyric acid	Homovanillic acid
1-Methylhistidine	5-Aminopentanoic acid	3-/4-Hydroxyphenylacetic acid/ or 3-cresotinic acid
Arginine	2-Aminobutyric acid	Homocarnosine
Homoarginine	Cysteine-glutathione disulfide	Gentisic acid
Asparagine	Sarcosine	Lysine
Glutamine	Methylcysteine	4-Hydroxybenzoic acid
L-citrulline	Proline	Histidine
3-sn-Phosphatidylethanolamine	Methylamine	2-Aminooctanoic acid
Methylguanidine	Valine	1,3-Diaminopropane
Homoserine^a	Methionine	L-Tyrosinamide
Methionine sulfoxide	3-Hydroxypicolinic acid	1,4-Diaminobutane
Serine	Tryptophan	Cadaverine
Homocitrulline	Pipecolic acid	Tyrosine
Glutamic Acid	Phenylalanine	Cysteamine
Aspartic Acid	3-Hydroxymandelic acid	Metoprolol
4-Hydroxyl-proline	Isoleucine	Phenol
Amino adipic acid	Leucine	4-Nitrophenol
Iminodiacetic acid	L-cystathionine	Octopamine
Folic acid	L-norleucine	Tyramine
Threonine	Cystine	Serotonin
Diethanolamine	Hydroxyphenyllactic acid	Pyrocatechol
Ethanolamine	Homocystine	Spermidine
Glycine	5-HIAA	Thymol
Glycylproline	Dimethylamine	Deoxyepinephrine
Tyrosine methyl ester	Phenylpropanolamine	
Alanine	2,4-Diaminobutyric acid	

^aHighlighted in bold are the metabolites that have not been previously reported to be present in human CSF

metabolites in the second experiment are not commonly detected. Note that the signal intensities of these non-common metabolites are relatively low. Most of the high abundant metabolite ion pairs can be observed in both labeling experiments. There are 54 metabolites that are commonly detected in all 4 CSF samples. There are only 8, 2, 1, and 3 metabolites solely detected in CSF samples 1 to 4, respectively.

It is interesting to compare the 85 metabolites identified to the 329 CSF metabolites reported in the literature (www.csfmetabolome.ca). Our work identified 21 metabolites that have not been reported to be present in human CSF (see Table 1 with compound names highlighted in bold). Interestingly, three of them, homoserine, 4-hydroxy-proline, and cadaverine, were on the list of the standard compounds in the study of Myint et al. [18]. However, these compounds were not identified in their nano-LC-MS analysis of CSF samples. This can be most likely attributed to the detection sensitivity difference of the techniques used. Detection sensitivity improvement afforded by dansylation labeling allowed us to identify these compounds in CSF. The detection limit in their work was in the low micromolar range, which is typical for LC-MS without ion selection monitoring, while detection of dansylated metabolites such as amino acids in the low nanomolar range can be achieved [20]. Note that some of the 21 metabolites listed in Table 1 (highlighted compounds) are biologically relevant to the neuronal system. For example, tyramine is a suspected neurotransmitter/neuromodulator or co-transmitter with octopamine in central nerve system.

Identification of many of the remaining unknown ion pairs detected by the dansylation LC-MS method is, however, a major analytical challenge. One strategy of averting this problem is to carry out relative quantification of the metabolomes of a number of comparative samples (e.g., diseased *versus* controlled) first to discover one or a few putative biomarkers. Relative quantification is done by a process including the following steps: (1) aliquoting an individual sample into two halves and mixing the aliquots of the individual samples to form a pooled sample, (2) labeling the pooled sample with the heavy chain reagent and the individual sample aliquots with the light chain reagent, (3) mixing the light-mass-tagged individual sample with an aliquot of the heavy-mass-tagged pooled sample, and (4) injecting the mixture into LC-MS to determine peak ratios of ion pairs for relative quantification of metabolites. After the discovery of the putative biomarkers, major efforts are then devoted to the identification of these metabolites using techniques such as tandem MS, NMR and synthesis of standard compounds. A sensitive LC-MS method can then be developed for targeted analysis of these putative biomarkers so that they may be validated by high-throughput analysis of a large number of samples. We envisage that the dansylation LC-MS method described in this work will be useful in the initial biomarker discovery stage, as it can be used to profile the CSF metabolome in a more comprehensive manner than other techniques.

Conclusions

We report the development and application of an isotope labeling LC/MS technique for the detection and identification of metabolites in human CSF samples. It is shown that differential isotope labeling using dansylation chemistry is effective in analyzing many low abundance metabolites that are present in CSF. Without labeling, very few metabolite peaks were detected in LC-MS. With labeling, an average of 519 ion pairs were observed in a 25-min LC-MS run with the injection of an equivalent of 0.5 μ L of the original CSF sample. About 261 ion pairs (50% of the total number of pairs found in each run) were commonly detected in four different CSF samples with each sample analyzed twice. Unique ion pairs were found in individual samples and, in total, 1132 unique ion pairs were detected from the combined results. Because dansylated metabolites rarely fragment in the skimmer region of the interface and during the transport into the detection cell of FTICR-MS, these ion pairs are most likely from the true metabolites. By searching the Human Metabolome Database, 347 unique ion pairs (31%) matched with at least one metabolite in the database. Even if we only consider these matched pairs, this number is already greater than the 329 CSF metabolites reported in the literature (www.csfinetabolome.ca). To provide positive identification of the metabolites in CSF, we have constructed a dansylation library of 220 standard compounds. Using this library, 85 metabolites were identified and, among them, 21 metabolites have not been described in the literature to be present in human CSF.

Future expansion of the standard library will undoubtedly increase the metabolite coverage and our understanding of the CSF metabolome. The ion pair detection technique described in this work is focused on the use of dansylation chemistry to label amine- and phenol-containing metabolites. Other labeling chemistries targeted at different functional groups are being developed in this laboratory. We envisage that the high performance isotope labeling LC-MS technique with much improved detection sensitivity and ion pair detection specificity should enable comprehensive detection of metabolites in biofluids with unprecedented metabolome coverage.

Acknowledgments

The authors acknowledge support for this work by the Genome Canada and Genome Alberta through the Human Metabolomics Toolbox (HMDB) project and the Canada Research Chairs program. The authors thank Professor David Wishart and his group members for providing many of the metabolite standards used in this research, Jeff Xu, Jun Peng, and Ben Zhou for their contributions in the development of the software package used in metabolite data processing and identification, and Fiona Bamforth and Connie Cheung of Clinical Laboratory Medicine, University of Alberta for providing the CSF samples.

References

1. Blennow, K., Hampel, H., Weiner, M., Zetterberg, H.: Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nat Rev Neurol* **6**, 131–144 (2010)
2. Henley, S.M.D., Bates, G.P., Tabrizi, S.J.: Biomarkers for neurodegenerative diseases. *Curr Opin Neurol* **18**, 698–705 (2005)
3. Maurer, M.H.: Proteomics of brain extracellular fluid (ECF) and cerebrospinal fluid (CSF). *Mass Spectrom Rev* **29**, 17–28 (2010)
4. Alvarez-Sanchez, B., Priego-Capote, F., de Castro, M.D.L.: Metabolomics analysis I. selection of biological samples and practical aspects preceding sample preparation. *Trac-Trends Anal Chem* **29**, 111–119 (2010)
5. Wishart, D.S., Lewis, M.J., Morrissey, J.A., Flegel, M.D., Jeroncic, K., Xiong, Y.P., Cheng, D., Eisner, R., Gautam, B., Tzur, D., Sawhney, S., Bamforth, F., Greiner, R., Li, L.: The human cerebrospinal fluid metabolome. *J Chromatogr B Anal Technol Biomed Life Sci* **871**, 164–173 (2008)
6. Huang, J.T.J., Leweke, F.M., Tsang, T.M., Koethe, D., Kranaster, L., Gerth, C.W., Gross, S., Schreiber, D., Ruhmann, S., Schultze-Lutter, F., Klosterkötter, J., Holmes, E., Bahn, S. CSF Metabolic and Proteomic Profiles in Patients Prodromal for Psychosis. *PLoS ONE* **2**, 756 (2007)
7. Himmelreich, U., Malik, R., Kuhn, T., Daniel, H.M., Somorjai, R.L., Dolenko, B., Sorrell, T.C.: Rapid etiological classification of meningitis by NMR spectroscopy based on metabolite profiles and host response. *PLoS ONE* **4**, 8 (2009)
8. Jukarainen, N.M., Korhonen, S.P., Laakso, M.P., Korolainen, M.A., Niemitz, M., Soininen, P.P., Tuppurainen, K., Vepsäläinen, J., Pirttilä, T., Laatikainen, R.: Quantification of H-1 NMR spectra of human cerebrospinal fluid: a protocol based on constrained total-line-shape analysis. *Metabolomics* **4**, 150–160 (2008)
9. Dunne, V.G., Bhattachayya, S., Besser, M., Rae, C., Griffin, J.L.: Metabolites from cerebrospinal fluid in aneurysmal subarachnoid haemorrhage correlate with vasospasm and clinical outcome: a pattern-recognition H-1 NMR study. *NMR Biomed* **18**, 24–33 (2005)
10. Stoop, M.P., Coulier, L., Rosenling, T., Shi, S., Smolinska, A.M., Buydens, L., Ampt, K., Stingl, C., Dane, A., Muilwijk, B., Luitwieler, R.L., Smitt, P., Hintzen, R.Q., Bischoff, R., Wijmenga, S.S., Hankemeier, T., van Gool, A.J., Luiders, T.M.: Quantitative proteomics and metabolomics analysis of normal human cerebrospinal fluid samples. *Mol Cell Proteom* **9**, 2063–2075 (2010)
11. Koek, M.M., Bakels, F., Engel, W., van den Maagdenberg, A., Ferrari, M.D., Coulier, L., Hankemeier, T.: Metabolic profiling of ultrasmall sample volumes with GC/MS: from microliter to nanoliter samples. *Anal Chem* **82**, 156–162 (2010)
12. Rosenling, T., Slim, C.L., Christin, C., Coulier, L., Shi, S., Stoop, M.P., Bosman, J., Suits, F., Horvatovich, P.L., Stockhofe-Zurwieden, N., Vreeken, R., Hankemeier, T., van Gool, A.J., Luiders, T.M., Bischoff, R.: The effect of preanalytical factors on stability of the proteome and selected metabolites in cerebrospinal fluid (CSF). *J Proteome Res* **8**, 5511–5522 (2009)
13. Crews, B., Wikoff, W.R., Patti, G.J., Woo, H.K., Kalisiak, E., Heideker, J., Siuzdak, G.: Variability analysis of human plasma and cerebral spinal fluid reveals statistical significance of changes in mass spectrometry-based metabolomics data. *Anal Chem* **81**, 8538–8544 (2009)
14. Kaddurah-Daouk, R., Kristal, B.S., Weinshilboum, R.M.: Metabolomics: a global biochemical approach to drug response and disease. *Annu Rev Pharmacol Toxicol* **48**, 653–683 (2008)
15. Bourcier, S., Benoist, J.F., Clerc, F., Rigal, O., Taghi, M., Hoppilliard, Y.: Detection of 28 neurotransmitters and related compounds in biological fluids by liquid chromatography/tandem mass spectrometry. *Rapid Commun Mass Spectrom* **20**, 1405–1421 (2006)
16. Myint, K.T., Uehara, T., Aoshima, K., Oda, Y.: Polar anionic metabolome analysis by Nano-LC/MS with a metal chelating agent. *Anal Chem* **81**, 7766–7772 (2009)
17. Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., Siuzdak, G.: XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* **78**, 779–787 (2006)
18. Myint, K.T., Aoshima, K., Tanaka, S., Nakamura, T., Oda, Y.: Quantitative profiling of polar cationic metabolites in human cerebrospinal fluid by reversed-phase nanoliquid chromatography/mass spectrometry. *Anal Chem* **81**, 1121–1129 (2009)

19. Wikoff, W.R., Pendyala, G., Siuzdak, G., Fox, H.S.: Metabolomic analysis of the cerebrospinal fluid reveals changes in phospholipase expression in the CNS of SIV-infected macaques. *J Clin Invest* **118**, 2661–2669 (2008)
20. Guo, K., Li, L.: Differential C-12/C-13-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. *Anal Chem* **81**, 3919–3932 (2009)
21. Ji, C.J., Li, W.L., Ren, X.D., El-Kattan, A.F., Kozak, R., Fountain, S., Lepsy, C.: Diethylation labeling combined with UPLC/MS/MS for simultaneous determination of a panel of monoamine neurotransmitters in rat prefrontal cortex microdialysates. *Anal Chem* **80**, 9195–9203 (2008)
22. O'Maille, G., Go, E.P., Hoang, L., Want, E.J., Smith, C., O'Maille, P., Nordstrom, A., Morita, H., Qin, C., Uritboonthai, W., Apon, J., Moore, R., Garrett, J., Siuzdak, G.: Metabolomics relative quantitation with mass spectrometry using chemical derivatization and isotope labeling. *Spectr-Int J* **22**, 327–343 (2008)
23. Shortreed, M.R., Lamos, S.M., Frey, B.L., Phillips, M.F., Patel, M., Belshaw, P.J., Smith, L.M.: Ionizable isotopic labeling reagent for relative quantification of amine metabolites by mass spectrometry. *Anal Chem* **78**, 6398–6403 (2006)
24. Abello, N., Geurink, P.P., van der Toorn, M., van Oosterhout, A.J.M., Lugtenburg, J., van der Marel, G.A., Kerstjens, H.A.M., Postma, D.S., Overkleeft, H.S., Bischoff, R.: Poly(ethylene glycol)-based stable isotope labeling reagents for the quantitative analysis of low molecular weight metabolites by LC-MS. *Anal Chem* **80**, 917–9180 (2008)
25. Eggink, M., Wijtmans, M., Ekkebus, R., Lingeman, H., de Esch, I.J.P., Kool, J., Niessen, W.M.A., Irth, H.: Development of a selective ESI-MS derivatization reagent: synthesis and optimization for the analysis of aldehydes in biological mixtures. *Anal Chem* **80**, 9042–9051 (2008)
26. Carlson, E.E., Cravatt, B.F.: Enrichment tags for enhanced-resolution profiling of the polar metabolome. *J Am Chem Soc* **129**, 15780–15782 (2007)
27. Johnson, D.W.: Contemporary clinical usage of LC/MS: analysis of biologically important carboxylic acids. *Clin Biochem* **38**, 3–361 (2005)
28. Lamos, S.M., Shortreed, M.R., Frey, B.L., Belshaw, P.J., Smith, L.M.: Relative quantification of carboxylic acid metabolites by liquid chromatography-mass spectrometry using isotopic variants of choline. *Anal Chem* **79**, 5143–5149 (2007)
29. Han, J., Danell, R.M., Patel, J.R., Gumerov, D.R., Scarlett, C.O., Speir, J. P., Parker, C.E., Rusyn, I., Zeisel, S., Borchers, C.H.: Towards high-throughput metabolomics using ultrahigh-field fourier transform ion cyclotron resonance mass spectrometry. *Metabolomics* **4**, 128–140 (2008)
30. Wishart, D.S., Knox, C., Guo, A.C., Eisner, R., Young, N., Gautam, B., Hau, D.D., Psychogios, N., Dong, E., Bouatra, S., Mandal, R., Sinelnikov, I., Xia, J.G., Jia, L., Cruz, J.A., Lim, E., Sobsey, C.A., Shrivastava, S., Huang, P., Liu, P., Fang, L., Peng, J., Fradette, R., Cheng, D., Tzur, D., Clements, M., Lewis, A., De Souza, A., Zuniga, A., Dawe, M., Xiong, Y.P., Clive, D., Greiner, R., Nazyrova, A., Shaykhtudinov, R., Li, L., Vogel, H.J., Forsythe, I.: HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res* **37**, D603–D610 (2009)
31. Guo, K., Ji, C.J., Li, L.: Stable-isotope dimethylation labeling combined with LC-ESI MS for quantification of amine-containing metabolites in biological samples. *Anal Chem* **79**, 8631–8638 (2007)
32. Giavalisco, P., Kohl, K., Hummel, J., Seiwert, B., Willmitzer, L.: ¹³C isotope-labeled metabolomes allowing for improved compound annotation and relative quantification in liquid chromatography-mass spectrometry-based metabolomic research. *Anal Chem* **81**, 6546–6551 (2009)
33. Lin, L.-C., Wu, H.-Y., Tseng, V.S.-M., Chen, L.-C., Chang, Y.-C., Liao, P.-C.: A statistical procedure to selectively detect metabolite signals in LC-MS data based on using variable isotope ratios. *J Am Soc Mass Spectrom* **21**, 232–241 (2010)
34. Coccaro, E.F., Lee, R.: Cerebrospinal fluid 5-hydroxyindolacetic acid and homovanillic acid: reciprocal relationships with impulsive aggression in human subjects. *J Neural Transm* **117**, 241–248 (2010)
35. Benton, H.P., Wong, D.M., Trauger, S.A., Siuzdak, G.: XCMS2: processing tandem mass spectrometry data for metabolite identification and structural characterization. *Anal Chem* **80**, 6382–6389 (2008)
36. Wang, N., Li, L.: Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Anal Chem* **80**, 4696–4710 (2008)