FD-LC-MS/MS Method for Determining Protein Expression and Elucidating Biochemical Events in Tissues and Cells

Tomoko Ichibangase and Kazuhiro Imai*
Research Institute of Pharmaceutical Sciences, Musashino University;
Received January 26, 2012

To study biochemical events in tissues and cells, we have developed a novel proteomics approach, FD-LC-MS/MS, which consists of fluorogenic derivatization (FD), LC separation and detection/quantification of proteins in a biological sample, followed by the isolation and tryptic digestion of target proteins, and then their identification using nano-HPLC-MS/MS. Fluorogenic reagents such as 7-chloro-N-[2-(dimethylamino)ethyl]-2,1,3-benzoxadiazole-4-sulphonamide (DAABD-Cl) were designed to have high sensitivity for HPLC-fluorescence and -MS/MS detection and reactivity for cysteine residues in proteins. This comprehensive differential proteomics approach was applied to several tissues, such as mouse liver, mouse brain, horse muscle, breast cancer cell lines, and mouse heart, in order to study fluctuations in protein levels in tissues and cells.

Key words fluorogenic derivatization; FD-LC-MS/MS; proteomics; protein network

1. INTRODUCTION

Fluorometry is an extremely sensitive method for detecting small amounts of biomolecules such as proteins. Therefore, it should be suitable for the detection, quantification, and identification of proteins expressed in proteomics studies. Since many proteins are weakly fluorescent or non-fluorescent, derivatization with fluorogenic or fluorescent reagents is required. Fluorogenic reagents which do not fluoresce themselves but generate fluorescence after derivatization are strongly recommended due to their low background fluorescence, which affords highly sensitive detection of the derivatives. Ammonium-7-fluoro-2,1,3-benzoxadiazole-4-sulphonate (SBD-F) and 7-chloro-N-[2-(dimethylamino)ethyl]-2,1,3-benzoxadiazole-4-sulphonamide (DAABD-Cl) are hydrophilic and thiol-specific fluorogenic reagents, and both derivatized proteins retain their hydrophilicity and thus do not precipitate after the complete derivatization of proteins (Fig. 1). Moreover, since their relatively small molecules enable them to react completely with cysteine residues in proteins under mild conditions, and their derivatives are highly fluorescent with longer emission wavelengths than the natural fluorescence of biomolecules (less than 400 nm), these reagents enable the sensitive detection of the protein derivatives.

HPLC is a powerful tool for reproducible separation of compounds in a complex matrix of bio-samples. However, it has not been normally used for protein separation in bio-samples, with the exception of the FD-LC-MS/MS method we have developed. The experimental procedure is as follows: a protein mixture is derivatized with a fluorogenic reagent (FD), such as SBD-F or DAABD-Cl and then their derivatives are quantified by HPLC-fluorescence detection. Next, a derivatized target or subject protein can be isolated from the eluate of the HPLC without losing any amino acid sequence information, such as the type of protein isoforms and post-translational modifications, and then digested by trypsin. Finally, the peptide mixtures obtained are subjected to nano-HPLC-MS/MS, to identify the isolated protein using a probability-based protein identification algorithm.

The quantification of the derivatized proteins has high reproducibility because of the highly reproducible performance of the HPLC system. Moreover, no pre-treatment steps (e.g.,

---

*To whom correspondence should be addressed. e-mail: k-imai@musashino-u.ac.jp

© 2012 The Pharmaceutical Society of Japan

---

Fig. 1. Structures of the Fluorogenic Reagents Developed in This Study
precipitation, clean-up with the extraction column, or enzymatic protein digestion before HPLC) except for FD are required in the quantification, resulting in high sensitivity and high reproducibility because any loss of the existing proteins in the sample occurs during the analytical procedures. In contrast, other proteomics analytical methods, such as two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and LC-MS/MS, isotope-coded affinity tags (ICAT) and isobaric tags for relative and absolute quantification (iTRAQ), require certain pre-treatment steps (precipitation, clean-up with column, and enzymatic protein digestion).

In this paper, we describe the history of the development of the FD-LC-MS/MS method and its application as a comprehensive differential proteomics approach to extracts of mouse liver, mouse brain, thoroughbred horse muscle, and breast cancer cell lines as well as fluctuations in protein levels in tissues and cells to elucidate intracellular biochemical events.

2. DEVELOPMENT OF A NOVEL PROTEOMICS APPROACH: FD-LC-MS/MS

Two-dimensional PAGE (2D-PAGE) has been widely used to identify proteins from protein matrices. After clean-up of the protein sample with a mini-column, the extracted proteins are separated by 2D-PAGE, and then the proteins isolated are digested with enzymes to the peptides. They are usually identified using MS/MS.

Although 2D-PAGE enables us to separate intact proteins comprehensively, it requires a skillful handling technique to obtain reproducible results. To achieve high reproducibility throughout the analytical procedure and high sensitivity, different fluorescent cyanine dyes have been developed. However, because of the fluorescent properties of the dyes (fluorescent reagents), the amounts of the cyanine dyes are deliberately kept lower than in the reactive site of the protein, which results in the formation of multiple labeling products for each protein and hence several spots on the gel.

Multidimensional LC, one of the standard methods for proteomics analysis, has been used to separate enzymatically digested peptides followed by identification by MS/MS. In most LC-based methods, a mixture of the digested isotope-labeled peptides is separated by reversed-phase chromatography, and then the eluate is analyzed directly by mass spectrometry. Two typical stable-isotope labeling reagents, ICAT and iTRAQ, were developed in recent years and utilized for quantitative proteomics analyses. However, the molecular masses of these reagents including tandem mass tags are relatively high, which may cause the less soluble derivatized peptides to precipitate out.

Therefore, we have developed a hydrophilic, small-mass derivatization reagent for FD-LC-MS/MS. In the first stage, we used SBD-F as the fluorogenic reagent. The derivatized proteins were highly fluorescent, with long excitation (380 nm) and emission wavelengths (520 nm), and were suitable for femtomole detection owing to the fluorogenic properties due to the SBD moiety. For ESI-MS, the proteins with positive charges were proven to be highly detectable because of their easy ionization and simple fragmentation. Therefore, to improve the detectability of the modified proteins in MS, we next developed the hydrophilic and positively charged fluorogenic reagent DAABD-CI for thiolates as a replacement for SBD-F. The DAABD derivatives are positively charged in solution and resulted in a significant enhancement of detectability as compared with that with SBD-F (up to 100–3000-fold). Subsequently, to expand the feasibility of the method with the use of DAABD-CI, we investigated the HPLC separation conditions for the DAABD derivatives, and the improved method was applied to quantification and identification of the differentially expressed proteins in the extracts mentioned above.

7-Fluoro-N-[2-(diethylamino)ethyl]-2,1,3-benzoxadiazole-4-sulfonamide (DAEAEBSF) was also developed for simultaneous detection of the DEAEABS and DAABD-derivatives by HPLC-fluorescence detection. We chose the structures of the former derivatives because they yielded similar retention times in HPLC and different fluorescent characteristics to the latter. Actually, after each of the two samples was derivatized with DAEAEBSF or DAABD-CI, respectively, two reaction solutions were combined and subjected to HPLC analysis with in-line two fluorescent detectors whose excitation and emission wavelengths settle in the optimized wavelength of the former- and latter-derivatives, respectively. The resultant two chromatograms had patterns that were quite similar to each other, and the pair of fluorogenic reagents (DAEAEBSF and DAABD-CI) would be applicable to differential proteomics analysis using FD-LC-MS/MS.

In recent years, the new compound 7-fluoro-N-[2-(dimethylamino)ethyl]-2,1,3-benzoxadiazole-4-sulfonamide (DAABD-F) has been synthesized. Owing to the higher reactivity of the fluorine atom in DAABD-F as compared with the chloride atom in DAABD-CI, the reagent reacts with not only cysteine but also a certain lysine residue in proteins. Using DAABD-F, a wide range of applications of FD-LC-MS/MS would become possible.

FD-LC-MS/MS involves a simple set up consisting of a pump, a column, and a fluorescence detector. Currently, a comprehensive profiling analysis in HPLC-fluorescence detection requires a 10-h operation. However, when the elution time of a subject protein is determined, it would be possible to reduce the time required for a selected analysis of the subject protein by re-optimizing the separation conditions. It would be also possible to reduce the overall analysis time by adopting a higher-performance column, such as a recently available nonporous reversed phase column.

3. APPLICATION OF THE METHOD

In this section, application of the method to the extracts for differential analysis of the tissues and cells is described. It was also applied to mouse plasma and Caenorhabditis elegans.

3.1. Disease-Related Proteins in Livers of Hepatitis-Infected Mouse

In the post-genome era, a sensitive quantitative method for differential profiling analyses of clinical proteomics is required to understand the disease progress. Therefore, as a preliminary step, we adopted an FD-LC-MS/
MS method using DAABD-Cl to clarify disease-related liver proteins in hepatocarcinogenesis in transgenic (Tg) and non-transgenic (NTg) mice at three different developmental stages. The accuracy of the method calculated based on the reproducibility of the peak heights using representatives from high, medium and low peaks obtained from each individual mouse liver sample was as follows: The relative standard deviation (R.S.D., %) for each between-day peak was less than 16% (a relatively high peak), 17% (a medium peak), and 23% (a relatively low peak) \(n=3\). The reproducibility of the retention time was also calculated using the relatively low peak, resulting in a between-day R.S.D. of 0.41% \(n=3\). The disease-related proteins identified are listed in Table 1. At the age of 6 months, the expression of apoptosis-related proteins was suppressed. At the age of 12 months, proteins related to respiration, the electron-transfer system, and anti-oxidation were significantly up-regulated. At the age of 16 months, proteins related to defense, \(\beta\)-oxidation, and apoptosis were significantly suppressed. These fluctuating levels of proteins could explain the progression of hepatocarcinogenesis. Accordingly, this study has demonstrated that the method would be useful for clinical proteomics analysis because of its high resolution, high sensitivity, and high reproducibility.

In this analysis, a wide-pore (30 nm) reversed-phase column (particle size 3.0 \(\mu\)m) was utilized for protein separation. However, a non-porous small-particle reversed-phase column (particle size: 2.0 \(\mu\)m) was recently developed. The latter column was found to separate DAABD-derivatives with 3-fold greater peak capacity as compared to the former in gradient elution. This was because both the non-poromeric and smaller particles of the material reduced eddy diffusion and mass-transfer resistance on separation and resulted in high-resolution chromatography. Therefore, the separation of liver mitochondrial proteins that were only slightly separated with the wide-pore column (160 peaks in 260 min analytical time) was successfully achieved with the novel non-porous column (420 peaks)\(^{22}\).

### 3.2. Age-Related Proteins in Small Regions of Mouse Brain\(^{26}\)

The method was also applicable to the differential proteomics analysis of small regions of mouse brain. To identify age-related proteins in small regions of mouse brain, we improved the method in terms of sample preparation and HPLC separation, and applied it to the differential proteomics analysis of aging in cerebral cortex (Ctx), hippocampus (Hippo), and brainstem (BS). The method showed good accuracy with R.S.D.s<10% for between-day protein peak heights and much better sensitivity for the detection of proteins compared to other proteomics approaches. For example, only 40 \(\mu\)g of brain segment/injection was used in the study, while hundreds of micrograms of brain segment was required in general brain proteome analysis using 2-DE methods.\(^{26-30}\)

The existence of 28 regionally-specific age-related proteins in mouse brain was demonstrated in this study, seven of which have been reported for the first time. Figure 2 presents a comparison of the peak heights of each of the age-related altered proteins among three aging stages (4, 12 and 20 weeks) in Ctx. Considering many age-related proteins were observed, we were able to estimate the alterations in function with aging in each brain region, as described below. The time-series alteration of the most altered proteins in Ctx and Hippo was consistent with that of each ribosomal protein. The change in the level of protein synthesis ability could thus be a factor.

### Table 1. Significantly Altered Proteins between Tg and NTg Rat Livers

<table>
<thead>
<tr>
<th>Tg/NTg ratio</th>
<th>Identified protein name</th>
<th>Molecular mass (Da)</th>
<th>GI number</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 month</td>
<td>Marker</td>
<td>Major urinary protein (MUP)</td>
<td>17549</td>
</tr>
<tr>
<td></td>
<td>Respiration</td>
<td>(\alpha)-Globin</td>
<td>15076</td>
</tr>
<tr>
<td></td>
<td>Apoptosis</td>
<td>Eukaryotic translation elongation factor 1a1 (EF-1a1)</td>
<td>50140</td>
</tr>
<tr>
<td>12 month</td>
<td>Respiration</td>
<td>Hemoglobin (\beta)</td>
<td>15653</td>
</tr>
<tr>
<td></td>
<td>Electron transfer system</td>
<td>ATP synthase, H+ transporting, mitochondrial F1</td>
<td>5834</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ATP synthase, H+ transporting, mitochondrial F0</td>
<td>18752</td>
</tr>
<tr>
<td></td>
<td>Defence</td>
<td>Betaine-homocysteine methyltransferase (BHMT)</td>
<td>44992</td>
</tr>
<tr>
<td></td>
<td>Metabolism</td>
<td>Acetaldehyde dehydrogenase (ALDH)</td>
<td>54410</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>Diazepam binding inhibitor, splice form 1/(\beta)</td>
<td>15219</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sorbitol dehydrogenase precursor</td>
<td>40606</td>
</tr>
<tr>
<td>16 month</td>
<td>Defence</td>
<td>Glycine (N)-methyltransferase</td>
<td>32712</td>
</tr>
<tr>
<td></td>
<td>Fatty acid metabolism (containing (\beta)-oxidation)</td>
<td>Hydroxacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), (\beta)-subunit (HADH)</td>
<td>51353</td>
</tr>
<tr>
<td></td>
<td>Apoptosis</td>
<td>Ribosomal protein S29, isoform 1</td>
<td>6672</td>
</tr>
<tr>
<td></td>
<td>Metabolism</td>
<td>Cystatin B</td>
<td>11039</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PREDICTED: carbamoyl-phosphate synthetase 1 (CSP1)</td>
<td>165705</td>
</tr>
</tbody>
</table>

Asterisks indicate significant differences (two-tailed Student’s \(t\)-test, *\(p<0.05\), **\(p<0.01\)). GI number is simply a series of digits that are assigned consecutively to each sequence record processed by NCBI.
causing the changes in brain function with aging. There were significantly fewer altered proteins in BS (3 fluctuated proteins) than in other regions (15 and 10 fluctuated proteins in Ctx and Hippo, respectively). The results supported the morphological observation that the number of neuron functional units in the brain was rarely altered with aging in BS. These results verified that small brain regions could be targets for proteomics analysis due to the high sensitivity of the FD-LC-MS/MS method.

3.3. Energy Supply-Related Proteins in Fast-Running Thoroughbred Horse Muscle

To extend the applicability of the method, we applied it to thoroughbred horse muscle. To quantify the fluctuations in protein expression, the training-to-detraining (+/-) ratio for proteins was calculated, and the correlation of the ratio with the percentage of maximum oxygen consumption (VO_2max; an indicator of running speed) was investigated (Table 2). Sixteen proteins involved in energy supply, especially in anaerobic energy production,
increased with an increase in VO₂max, suggesting that this method was able to represent the biochemical events in fast horses and would be useful for evaluating the training effect in thoroughbred horses.

### 3.4. Breast Cancer-Related Proteins in Metastatic and Non-metastatic Breast Cancer Cells

Although several molecular markers for human breast cancer exist, their versatility has been limited. Therefore, we performed a differential proteomics analysis utilizing the method in 7 cancer cells and one normal cell. Since it is easy to compare a peak height on an LC-chromatogram of a subject sample with the corresponding peak height for that of other samples (normal versus cancer cells, metastatic versus non-metastatic cancer cells, etc.) using this method, we are not subject to any restrictions on the number of samples. Typical chromatograms are shown in Fig. 3. We were able to delineate the biochemical events involved in tumor progression and metastasis (Fig. 4). Additionally, we proposed sets of markers for metastatic breast cancer: the presence of cooperatively expressed annexin-2 and galectin-1 without tropomyosin-1 in a tissue would indicate its metastasis. Interestingly, in a metastatic cancer cell line, expression of the former two together with highly expressed cofilin-1 activated the Rho signal pathway (bold arrows in Fig. 4) to aggressively form disorganized actin filaments. Despite the excess expression of annexin-2 and galectin-1 in the normal cells, the highly expressed tropomyosin-1 counteracted the activity of cofilin-1 and stabilized the filaments, resulting in the restoration of the disorganization. This phenomenon suggests that enhancement of tropomyosin-1 could be used as a therapy for metastatic breast cancer.

<table>
<thead>
<tr>
<th>Identified protein name</th>
<th>Thoroughbred-1 VO₂max: 177 mL/(kg·min)</th>
<th>Thoroughbred-2 VO₂max: 181 mL/(kg·min)</th>
<th>Thoroughbred-3 VO₂max: 193 mL/(kg·min)</th>
<th>Thoroughbred-4 VO₂max: 202 mL/(kg·min)</th>
<th>Correlation with VO₂max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muscle creatine kinase (M-CK)</td>
<td>Similar to M-CK 0.27* 0.55* 0.77 2.14*</td>
<td>Similar to M-CK 0.43* 0.98 1.10 1.68*</td>
<td>Similar to M-CK 0.61* 1.05 1.16 1.83*</td>
<td>Similar to M-CK 0.43* 1.10 1.13 1.51*</td>
<td>0.917 0.933 0.939 0.867</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)</td>
<td>0.64* 1.02 1.08 1.10</td>
<td>0.43* 0.96 1.07 1.26</td>
<td>0.33* 1.02 1.27 1.59*</td>
<td>0.42* 0.78 1.85* 1.21</td>
<td>0.715 0.914 0.715 0.715</td>
</tr>
<tr>
<td>Similar to triosephosphate isomerase 1</td>
<td>0.43* 0.96 1.07 1.26</td>
<td>0.33* 1.02 1.27 1.59*</td>
<td>0.42* 0.78 1.85* 1.21</td>
<td>0.31* 0.87 2.14* 1.85*</td>
<td>0.886 0.914 0.715 0.886</td>
</tr>
<tr>
<td>Similar to pyruvate kinase 3</td>
<td>0.64* 1.02 1.08 1.10</td>
<td>0.33* 1.02 1.27 1.59*</td>
<td>0.42* 0.78 1.85* 1.21</td>
<td>0.31* 0.87 2.14* 1.85*</td>
<td>0.886 0.914 0.715 0.886</td>
</tr>
<tr>
<td>Similar to enolase 3, beta muscle isoform</td>
<td>0.43* 0.96 1.07 1.26</td>
<td>0.33* 1.02 1.27 1.59*</td>
<td>0.42* 0.78 1.85* 1.21</td>
<td>0.31* 0.87 2.14* 1.85*</td>
<td>0.886 0.914 0.715 0.886</td>
</tr>
<tr>
<td>Similar to glucose-6-phosphate isomerase</td>
<td>0.43* 0.96 1.07 1.26</td>
<td>0.33* 1.02 1.27 1.59*</td>
<td>0.42* 0.78 1.85* 1.21</td>
<td>0.31* 0.87 2.14* 1.85*</td>
<td>0.886 0.914 0.715 0.886</td>
</tr>
<tr>
<td>Other proteins</td>
<td>Not detected 0.67* 1.07 1.97* 2.05*</td>
<td>Hemoglobin, beta 0.37* 0.87 1.07 1.71*</td>
<td>Similar to malate dehydrogenase, cytoplasmic 0.63* 1.06 1.29 1.04</td>
<td>0.665 0.914 0.914 0.914</td>
<td></td>
</tr>
</tbody>
</table>

Asterisks denote significant differences (two-tailed Student’s t-test, p<0.05).
3.5. Cardioprotective-Related Proteins in Mouse Heart Pre-administered Docetaxel in a Model of Adriamycin-Induced Cardiotoxicity

Recently, to expand the method to include the study of toxicoproteomics, we applied it to mouse heart tissues isolated from control, intermittent-dosing (docetaxel (DOC)-adriamycin (ADR)), and simultaneous-dosing (DOC and ADR) groups. Clinical trials of anticancer therapy for metastatic breast cancer have reported that combination therapy of DOC and ADR is much more effective than previous combination therapies such as ADR-cyclophosphamide and fluorouracil-ADR-cyclophosphamide. Moreover, pre-administration of DOC in combination therapy has been shown to result in stronger antitumor effects and fewer ADR-induced cardiotoxic deaths in mouse models, however, no mechanism explaining this effect has yet been established. In this study, to identify the cellular processes affected by the pre-administration of DOC in the combination therapies, differential proteomics analysis using the above three groups was performed. As a result, nine proteins were observed to fluctuate differentially in heart tissue of the control and two treatment groups; seven of these proteins were involved in cellular energy production pathways, such as glycolysis, the tricarboxylic acid cycle, and the mitochondrial electron transport chain. Interestingly, a significantly higher level of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was observed in the DOC-ADR group, the group with the fewer cardiotoxic deaths, than in the DOC and ADR group, suggesting that GAPDH may counteract ADR-induced cardiotoxicity. GAPDH is a multifunctional protein involved in energy supply and antioxidative activity. Therefore, any of the functions of GAPDH may be relevant to the cardioprotective effects observed in the DOC-ADR treatment group. Although further analysis is needed, perhaps using a transgenic approach or another approach with the addition of...
of N6-naphthalenemethyl-20-methoxybenzamido-β-NAD+ for GAPDH inhibition, the present results suggest that GAPDH may have potential as a drug target for protective intervention and as a biomarker for the evaluation of cardioprotective effects in pre-clinical studies.

4. FUTURE TRENDS

We are currently investigating colorectal cancer-related and superoxide-related proteins in colorectal cancer cells and superoxide dismutase-deficient cells, respectively, and the preliminary results have revealed interesting intracellular phenomenon, such as a protein network for colorectal tumor progression and for increasing oxidative stress. In the future, the method should be extended to include the differential proteomics analysis of many different biological samples to identify protein networks, including the roles of certain proteins in animals.

We are also developing a more elaborate analytical system, including the use of a higher-performance column and two-dimensional HPLC system, to achieve the complete separation of proteins in tissues and cells. Also, the inexpensive automated system through the derivatization of proteins to the identification by HPLC-MS/MS should be considered to promote familiarizing FD-LC-MS/MS.

5. CONCLUSION

The FD-LC-MS/MS method described here has two unique features; it uses 1) hydrophilicity and a positively charged fluorogenic derivatization reagent, such as DAABD-Cl, to derivatize proteins, and 2) HPLC to separate the derivatized proteins. The reagent gives low background fluorescence affording the sensitive detection of derivatized proteins, while HPLC provides the reproducible quantitation of proteins. Actually, the data obtained has proved the method is highly sensitive and highly reproducible with respect to peak heights and retention times. Moreover, protein isoforms and post-translational modifications could be identified due to the isolation of only subject proteins and injection of the peptide mixtures obtained into the nano-HPLC-ESI-MS/MS after digestion.

In the present study using differential proteomics analysis, differences in protein expression between tissues and cells under different conditions, such as treated and nontreated tissue, were successfully demonstrated, and the biochemical events and possible mechanisms were described.

Acknowledgements We would like to express our gratitude to Drs. H. Saimaru, H. Asamoto, A. Koshiyama and K. Ohyama, and Mr. I. Yazawa for their valuable contributions to this study.

REFERENCES


