Chapter 21

Identification of Chemical-Adducted Proteins in Urine by Multi-dimensional Protein Identification Technology (LC/LC–MS/MS)

Matthew T. Labenski, Ashley A. Fisher, Terrence J. Monks,
and Serrine S. Lau

Abstract

Recent technological advancements in mass spectrometry facilitate the detection of chemical-induced posttranslational modifications (PTMs) that may alter cell signaling pathways or alter the structure and function of the modified proteins. To identify such protein adducts (Kleiner et al., *Chem Res Toxicol* 11:1283–1290, 1998), multi-dimensional protein identification technology (MuDPIT) has been utilized. MuDPIT was first described by Link et al. as a new technique useful for protein identification from a complex mixture of proteins (Link et al., *Nat Biotechnol* 17:676–682, 1999). MuDPIT utilizes two different HPLC columns to further enhance peptide separation, increasing the number of peptide hits and protein coverage. The technology is extremely useful for proteomes, such as the urine proteome, samples from immunoprecipitations, and 1D gel bands resolved from a tissue homogenate or lysate. In particular, MuDPIT has enhanced the field of adduct hunting for adducted peptides, since it is more capable of identifying lesser abundant peptides, such as those that are adducted, than the more standard LC–MS/MS. The site-specific identification of covalently adducted proteins is a prerequisite for understanding the biological significance of chemical-induced PTMs and the subsequent toxicological response they elicit.

Key words: MuDPIT, Covalent modification, Quinone, Quinol–thioether

1. Introduction

A number of xenobiotics are capable of giving rise to covalently adducted proteins within the kidney. The presence of such covalently modified proteins is frequently accompanied by the development of renal toxicity. Chemicals such as mycophenolic acid (1), S-(1,1,2,2-tetrafluoroethyl)-l-cysteine (2), and glutathione (GSH) conjugates of hydroquinone (HQ) all exert their toxic

effects on the kidney concomitant with the presence of covalently adducted proteins (3, 4). In addition to covalently adducting proteins within the S3 segment of the proximal tubule, GSH conjugates of HQ, including 2-(glutathion-*S*-yl)HQ, maintain the ability to redox cycle. The combination of these insults contributes to the ensuing cell and tissue necrosis. Moreover, because the targeted cells line the tubular lumen, dead and dying cells detach from the basement membrane and are sloughed into tubular fluid, resulting in the presence of both adducted and native proteins in tubular fluid, and ultimately in urine. Thus, urine is a perfect medium for the application of multi-dimensional protein identification technology (MuDPIT).

For any experiment generating samples that will be analyzed by mass spectroscopy, it is recommended that that sample be cleaned of any excess salt, small molecules, or other contaminants that may interfere with either protein digestion or the mass spectroscopic analysis. We have found that using a gravity-flow size exclusion column, followed by lyophilization, provides an appropriate level of sample clean-up. For experiments described herein, a solution digestion using trypsin was utilized. Following protein digestion and concentration of the peptide solution, the sample is then loaded onto the first HPLC column (strong cation exchange (SCX)) in preparation for MuDPIT analysis. MuDPIT utilizes two degrees of peptide separation, increasing the amount of peptide matches and protein coverage. In combination with data-dependent MS-MS scanning, more peptide ions are sampled and sequenced; relative to LC-MS/MS, this results in greater protein coverage per protein identified. Raw MS/MS data are analyzed using the matching correlation algorithm Sequest (5). The Sequest data output is then analyzed using Scaffold, a program that runs X!Tandem (6) on the Sequest output to validate identified peptides, and which also performs protein and peptide statistical analysis. Once modified peptides are identified, visual inspection and validation of the spectra are suggested, followed by further conformational analysis with a program such as P-Mod or FindMod. The larger the data set, the higher the likelihood of a false positive; it is, therefore, important to emphasize the use of multiple programs to validate the data.

2. Materials

2.1. Size Exclusion

1. PD-10 Desalting columns, code no. 17-0851-01 (GE Healthcare).

2.2. Solution Digest

1. Dithiothreitol (DTT), iodoacetamide, and ammonium bicarbonate (NH_4HCO_3) (Sigma).

2. 2 M DTT stock solution: 154.3 mg of DTT in 500 μ l of deionized distilled water (see Note 1).
3. DTT solution for reduction: 5 μ l of 2 M DTT in 1 ml of 0.1 M NH_4HCO_3 .
4. Iodoacetamide stock solution: 10 mg in 1 ml of 0.1 M NH_4HCO_3 .
5. Trypsin-modified sequencing grade (Promega): 20 μ g of lyophilized powder. Store at -20°C . Dilute to 0.1 $\mu\text{g}/\mu\text{l}$ and can be stored in solution at -20°C for several weeks.

2.3. MuDPIT Analysis

1. Silica tubing O.D. 363 μm , I.D. 250 μm (Polymicro, Phoenix, AZ) for strong cation exchange column.
2. Silica tubing O.D. 363 μm , I.D. 100 μm (Polymicro, Phoenix, AZ) for C18 column.
3. SCX resin, 5 μm Whatman partisphere.
4. C18 reverse phase (RP) resin, 5 μm Vydac.
5. HPLC grade methanol.
6. Compressed nitrogen tank with regulator, connected to a "bomb."

2.4. Mobile Phase

1. Mobile phase A: 10% methanol/0.5% formic acid, 0.01% TFA.
2. Mobile phase B: 98% methanol/0.5% formic acid, 0.01% TFA.
3. Mobile phase C: 10% methanol/0.5% formic acid, 0.01% TFA.
4. Mobile phase D: 500 mM ammonium acetate/10% methanol/0.5% formic acid, 0.01% TFA.

3. Methods

The main feature of MuDPIT is the use of two different types of HPLC separation columns; one made of strong cation exchange material and the second column composed of C18 material. Acidified peptides are loaded onto the cation exchange column, where they bind. A stepwise gradient of increasing salt steps will elute the peptides from the first column, when they will then pass into the second, C18 column. Prior to the next salt elution step, an organic gradient is run, which results in the release of peptides from the C18 column at various concentrations of the organic phase. The flow is in line with a nano-ESI source into the mass spectrometer. After the organic gradient is completed, another salt-based elution is performed, and the process is repeated as many times as necessary. Simple matrices, such as urine or a semi-enriched sample from either immunoaffinity or

Fig. 1. General schematic for site-specific identification of quinol-thioether adducted proteins from the semi-complex urine proteome from rats treated i.v. with MGHQ. 100 µg of urinary proteins were cleaned up by passing through a size-exclusion PD-10 column, followed by lyophilizing to dryness. Protein was then reconstituted in digestion buffer and digested with trypsin. Peptides were then loaded onto the mass spectrometer for MuDPIT analysis. Raw data from the mass spectrometry was first analyzed for both adducted and native peptides with Sequest. The Sequest output was then compared and validated using Scaffold (combines X!Tandem and Peptide Profit algorithms with SEQUEST). All adducted peptides were then validated both manually using lonGen and/or protein prospector, and with other software programs (P-Mod, FindMod).

tag-affinity, are best suited for MuDPIT analyses, although there has also been much success with more complex matrices. A general work flow for this set of experiments is shown in Fig. 1.

3.1. Size Exclusion

1. Equilibrate PD-10 desalting column with 25 ml of water and discard the flow through (see Note 1).
2. Dilute protein sample to 2.5 ml and add to the column. Disregard flow through.
3. Elute protein from column by adding 3.5 ml of water to the top. Collect flow through and freeze at -20°C immediately.
4. Lyophilize sample till dry.

3.2. Solution Digestion

1. Reconstitute samples in 100 µl of 0.1 M NH_4HCO_3 (see Note 2).
2. Most successful protein digests are conducted using trypsin as the proteolytic enzyme, with 50 mM Tris-HCl or 50–100 mM NH_4HCO_3 as the buffering system. These buffers

are usually in the pH range of 7.4–8.0. This is ideal for trypsin digestion to occur (see Note 3).

3. For proteins with disulfide bonds, a reduction/alkylation process will be necessary prior to trypsin digestion. These reduction/alkylation steps need to take place in the dark (see Note 4).
4. Prepare a 2 M stock solution of DTT and dilute 5 μ l of this with 1 ml of 0.1 M NH_4HCO_3 . Add 40 μ l of the diluted DTT solution to both control and treated samples, and incubate at 56°C for 45 min.
5. Remove samples from heat and allow to cool to room temperature.
6. Prepare a 55 mM stock solution of iodoacetamide and add 40 μ l of this stock solution to both control and treated samples. Incubate at room temperature for 30 min.
7. The ratio of trypsin to sample should be 1:50–1:20 (1 mg of trypsin to every 50–20 mg of protein). Add the appropriate amount of trypsin to control and treated samples. Gently agitate the tubes to mix.
8. Place the samples in a water bath at 37°C for 2 h and then add an additional amount (equal to the first amount added) of trypsin to each sample.
9. Continue to incubate at 37°C for 17–18 h to ensure complete digestion (see Note 5).
10. Trypsin digestion reactions can be terminated by freezing samples at –20°C, or by adding 4 μ l of glacial acetic acid.
11. Reduce the volume of the samples to approximately 10 μ l by vacuum concentration.

3.3. MuDPIT

3.3.1. Column Preparation

1. Cut a 35-cm long piece of silica capillary tubing (O.D. 363 μ m and I.D. 250 μ m) (7).
2. Using an alcohol burner, burn off insulation on outside of silica tubing in the middle 1.5 cm (see Note 6).
3. Resuspend the C18 RP resin in methanol and remove supernatant. Repeat this process three more times to remove fine particles from the mixture.
4. Pull column using a Sutter Instrument Co. Model p-2000 micropipette puller, and use a ceramic cutter (LC packing international) (see Note 7).
5. The column is then packed under high pressure (700 psi) with compressed nitrogen to 7 cm in a pressurized cell (bomb) connected to a regulator (7).
6. The SCX column is prepared in the same manner as the RP column, but without a tapered tip. A frit in a zero volume union is used as a barrier for the SCX to pack against. A 100- μ m

I.D. capillary packed with 7 cm of 5 μ m Vydac C18 reversed phase resin, and a separate 250- μ m I.D. capillary packed with 7 cm of 5 μ m Partisphere strong cation exchanger resin were used for the studies described herein.

3.3.2. MuDPIT Mass Spectrometric Analysis

1. For the nano LC/LC-MS/MS, a microbore HPLC system (Paradigm MS4, Michrom, Auburn, CA) was used with two separate SCX and RP columns.
2. The trypsin-digested peptides were pooled and acidified using trifluoroacetic acid and manually injected onto the SCX column, the effluent from the column being next fed through the RP column.
3. The 12-step MuDPIT analyses were as follows. Step 1 consists of a 5-min equilibration step with 100% buffer A, followed by another equilibration step for 5 min with 25% buffer B (75% buffer A), followed by a 40-min gradient from 25% buffer B to 65% buffer B, followed by a 10-min 65% buffer B and 10 min of 100% buffer A. Chromatography steps 2–12 follow the same pattern: 15 min of the appropriate % of buffers C and D, followed by a 2-min wash with 100% buffer C, a 5-min wash with 100% A, equilibration with 25% buffer B for 5 min, followed by a gradient from 25% buffer B to 65% buffer B for 40 min, followed by a 10-min 65% buffer B and 10 min of 100% buffer A. The stepwise gradient used was as follows: 0, 25, 50, 75, 100, 150, 200, 250, 300, 400, and 500 mM ammonium acetate, respectively, for the 11 salt steps. Eluting peptides at 350 nl/min were electrosprayed into a ThermoFinnigan LCQ-Deca XP Plus ion trap mass spectrometer (ThermoFinnigan, San Jose, CA) (8) with a spray voltage of 1.6–2.3 kV. Spectra were scanned over the range 380–2,000 amu. Automated peak recognition, dynamic exclusion, and daughter ion scanning of the most intense ion were performed using the Xcalibur v1.3 software, as described previously (9, 10).

3.4. Data Analysis

3.4.1. Database Searching

1. Analyze tandem MS spectra of peptides with TurboSEQUENT™, a program that allows the correlation of experimental tandem MS data with theoretical spectra generated from known protein sequences (5). The peak list (data files) for the search is generated by Bioworks 3.1. Parent peptide mass error tolerance was set at 1.5 amu and fragment ion mass tolerance was set at 0.5 amu during the search.
2. The criteria for a preliminary positive peptide identification: peptide precursor ions with a +1 charge having a Xcorr >1.0, +2 Xcorr >1.5, and +3 Xcorr >2.0. A dCn score >0.08 and a fragment ion ratio of experimental/theoretical >50% were also used as filtering criteria for reliable matched peptide identification (11). All matched peptides were confirmed by visual examination of the spectra.

3. X!Tandem (<http://www.thegpm.org>; version 2007.01.01.1) was used to query the ipi.RAT.v3.31 database. Sequest was selected to query the ipi.RAT.v3.31.fasta.hdr database (41,251 entries), assuming the digestion enzyme trypsin. Sequest and X!Tandem were searched with a fragment ion mass tolerance of 0.5 Da and a parent ion tolerance of 1.5 Da.
4. Modifications searched consisted of oxidation of methionine, iodoacetamide derivative of cysteine, +105 Da (HQ), +412 Da (MGHQ), +268 Da (NACHQ), +226 Da (CSHQ), and +283 Da (CysGlyHQ); at Lys, Arg, and Glu were specified in Sequest and X!Tandem as variable modifications.

3.4.2. Criteria for Protein Identification

1. Scaffold (version Scaffold-02.05.00, Proteome Software Inc., Portland, OR) was used to validate initial MS/MS-based peptide and protein identifications, which were subsequently accepted if they could be established at greater than 90.0% probability, as specified by the Peptide Prophet algorithm (12) and contained at least 1 identified peptide. Protein probabilities were assigned by the Protein Prophet algorithm (13). Proteins that contained similar peptides and that could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

3.4.3. Adduct Validation

All peptides found by Scaffold to be modified by chemical adduction were confirmed with P-Mod, and then manually validated with IonGen (14). P-Mod is an algorithm that screens data files for MSMS spectra corresponding to peptide sequences in a search list (15). Modification of the primary peptide sequence shifts the peptide mass, which may be experimentally observed as a difference between the measured mass of the modified peptide precursor ion (adjusted for charge state) and the predicted mass of the unmodified peptide. The mass shift will also be observed in the m/z values of fragment ions containing the modified amino acid. Scores with p -values greater than 0.01 are discarded as false positives. Upon collision-induced dissociation (CID) of the peptides, b- and y-ion fragments are generated: the b-ion series represents cleavage of the peptide bond with charge retention on the N-terminal fragment, and y-ions result from cleavage of the amide bonds with charge retention on the C-terminal side (16). Manual validation of MS/MS spectra is then used to confirm peptide sequence and adduct mass location. IonGen generates theoretical b- and y-ions for user-specified peptides containing adduct(s) (14).

1. Generate search file by inputting sequence of adducted peptide and any other types of modifications (oxidized methionines, alkylated cysteines, etc.).
2. Input.RAW datafile into search window and run program.

3. Look for peptide containing mass addition of interest (+105 amu for BQ modification) and make sure the amino acid adducted is a chemically appropriate target amino acid.
4. Look at p -value and make sure it is below the pre-determined cutoff ($p < 0.05$).

4. Notes

1. Water was chosen in this example, but other buffers can also be used, such as ammonium bicarbonate or ammonium acetate at 10 mM. When sample is lyophilized, the buffer will be removed, leaving only pure protein in the tube.
2. It is important to make sure that all samples are free of detergents before proceeding with trypsin digestion, and this can be achieved by protein precipitation methods.
3. Additional enzymes can be used for protein digestion, including chymotrypsin, endoprotease Asp-N, and glutamic acid-C, all of which have specific optimal pH ranges. Trypsin is the most common choice for proteomic procedures, so methodology using this enzyme is described here.
4. Reduction steps improve digestion efficiency by preventing disulfide formation, which assists in maintaining the protein in an unfolded state, exposing more tryptic sites for digestion.
5. As stated previously, the quinol-thioether conjugates are unstable in basic conditions. Consequently, when using compounds of this nature, the digestion procedure needs to be limited to approximately 3 h, to minimize quinol-thioether exposure to the basic buffering system necessary for trypsin digestion. Efficient digestion can be accomplished in less than 3 h.
6. If insulation remains present in the area of silica tubing being pulled, the laser will burn off the insulation onto the mirror, which will decrease the effectiveness of the laser.
7. This produces a column with a tapered tip that sprays directly into the mass spectrometer. The tapered tip also retains the C18 in the column.

Acknowledgments

This work was supported by GM070890 (SSL) and ES07091 (AAF). The authors acknowledge the support of the P30 ES06694 Southwest Environmental Health Sciences Center, in particular the Arizona Proteomics Consortium (APC). Our special thanks go to Dr. George Tsapralis, Director of the APC.

References

1. Asif, A. R., Armstrong, V. W., Volland, A., Wieland, E., Oellerich, M., and Shipkova, M. (2007) Proteins identified as targets of the acyl glucuronide metabolite of mycophenolic acid in kidney tissue from mycophenolate mofetil treated rats, *Biochimie* 89, 393–402.
2. Bruschi, S. A., Lindsay, J. G., and Crabb, J. W. (1998) Mitochondrial stress protein recognition of inactivated dehydrogenases during mammalian cell death, *Proc Natl Acad Sci USA* 95, 13413–13418.
3. Kleiner, H. E., Rivera, M. I., Pumford, N. R., Monks, T. J., and Lau, S. S. (1998) Immunochemical detection of quinol-thioether-derived protein adducts, *Chem Res Toxicol* 11, 1283–1290.
4. Kleiner, H. E., Jones, T. W., Monks, T. J., and Lau, S. S. (1998) Immunochemical analysis of quinol-thioether-derived covalent protein adducts in rodent species sensitive and resistant to quinol-thioether-mediated nephrotoxicity, *Chem Res Toxicol* 11, 1291–1300.
5. Yates, J. R., III, Eng, J. K., McCormack, A. L., and Schieltz, D. (1995) Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database, *Anal Chem* 67, 1426–1436.
6. Craig, R., Cortens, J. P., and Beavis, R. C. (2004) Open source system for analyzing, validating, and storing protein identification data, *J Proteome Res* 3, 1234–1242.
7. Jorde, L. B. (2005) *Encyclopedia of genetics, genomics, proteomics, and bioinformatics*, Wiley, Chichester.
8. Washburn, M. P., Wolters, D., and Yates, J. R., III. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology, *Nat Biotechnol* 19, 242–247.
9. Andon, N. L., Hollingworth, S., Koller, A., Greenland, A. J., Yates, J. R., III, and Haynes, P. A. (2002) Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry, *Proteomics* 2, 1156–1168.
10. Lantz, R. C., Lynch, B. J., Boitano, S., Poplin, G. S., Littau, S., Tsapralis, G., and Burgess, J. L. (2007) Pulmonary biomarkers based on alterations in protein expression after exposure to arsenic, *Environ Health Perspect* 115, 586–591.
11. Cooper, B., Eckert, D., Andon, N. L., Yates, J. R., and Haynes, P. A. (2003) Investigative proteomics: identification of an unknown plant virus from infected plants using mass spectrometry, *J Am Soc Mass Spectrom* 14, 736–741.
12. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search, *Anal Chem* 74, 5383–5392.
13. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry, *Anal Chem* 75, 4646–4658.
14. Fisher, A. A., Labenski, M. T., Malladi, S., Gokhale, V., Bowen, M. E., Milleron, R. S., Bratton, S. B., Monks, T. J., and Lau, S. S. (2007) Quinone electrophiles selectively adduct “electrophile binding motifs” within cytochrome c, *Biochemistry* 46, 11090–11100.
15. Hansen, B. T., Davey, S. W., Ham, A. J., and Liebler, D. C. (2005) P-Mod: an algorithm and software to map modifications to peptide sequences using tandem MS data, *J Proteome Res* 4, 358–368.
16. Standing, K. G. (2003) Peptide and protein de novo sequencing by mass spectrometry, *Curr Opin Struct Biol* 13, 595–601.
17. Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R., III. (1999) Direct analysis of protein complexes using mass spectrometry, *Nat Biotechnol* 17, 676–682.