

PROTEOMIC SOLUTIONS FOR ANALYTICAL CHALLENGES ASSOCIATED WITH ALCOHOL RESEARCH

Michael J. MacCoss, Ph.D., and Christine C. Wu, Ph.D.

KEY WORDS: Alcohol dependence; genetic theory of alcohol and other drug use; genetic trait; brain; animal models; proteins; protein analysis; proteomics; mass spectrometry; peptides.

Alcohol addiction is a complex disease with both hereditary and environmental influences. Because molecular determinants contributing to this phenotype are difficult to study in humans, numerous rodent models and conditioning paradigms¹ have provided powerful tools to study the molecular complexities underlying these behavioral phenotypes (Crabbe 2002). In particular, specifically bred rodents (i.e., selected lines and inbred strains) that differ in voluntary alcohol drinking represent valuable tools to dissect the genetic components of alcoholism. However, because each model has distinct advantages, a combined comparison across datasets of different models for common changes in gene expression would provide more statistical power to detect reliable changes as opposed to the analysis of any one model. Indeed, meta-analyses of diverse gene expression datasets were recently performed to uncover genes related to the predisposition for a high alcohol intake. This large endeavor resulted in the identification of 3,800 unique genes that significantly and consistently changed between all included mouse lines and strains (Mulligan et al. 2006).

Similar experiments also are crucial at the protein level. However, these analyses are not yet possible. Proteins do not conform to any one uniform sample preparation method and/or biochemical analysis. They display a broad range of physical and chemical properties (e.g., molecular weight or hydrophobicity) and are expressed over a very large dynamic range (up to 8 orders of magnitude) (Anderson 2005; Ghaemmaghami et al. 2003). Further complicating global proteomic comparisons are the added considerations that proteins often undergo extensive covalent modifications and that protein functions often are regulated by complex protein–protein interactions and the specific location of the proteins in the cell (i.e., their subcellular localization) (Grant and Wu 2007). Furthermore, because the number of biological replicates involved in behavioral analyses typically is high, robust high-throughput proteomic platforms will be required to handle the multitude of protein samples that can potentially result from the various brain regions for the numerous animal models and paradigms. Finally, these effects often are monitored over time courses, again inflating the total number of samples that need to be analyzed and compared. This article summarizes some general strategies for large-scale, high-throughput protein analyses and describes two new

proteomic strategies that appear promising for future studies in this field.

GENERAL STRATEGIES OF LARGE-SCALE PROTEIN ANALYSES

To overcome the complexities of proteins, proteomic methods routinely digest proteins into smaller pieces (i.e., peptides) prior to analysis. (Strategies using this approach are collectively termed shotgun proteomics.) The peptides then are separated by microcapillary chromatography and electrosprayed directly into a mass spectrometer placed at the outlet of the chromatography column for mass analysis and fragmentation of the peptides² using a technique known as data-dependent acquisition. In data-dependent acquisition, the mass spectrometer software makes real-time decisions about how the mass spectrometer is scanned, depending on data acquired in prior scans. Thus, the mass spectrometer acquires a “survey” mass spectrum, followed by tandem mass spectra (MS/MS spectra) of precursor ions identified in the earlier survey mass spectrum. This data-dependent acquisition, although powerful, is normally focused on the most intense signals and will ultimately be limited by the complexity of the mixture and the scan speed of the mass spectrometer.

The mass spectrometer’s ability to handle extremely complex mixtures of peptides can be enhanced either by biochemically fractionating the protein sample prior to digestion (thereby reducing the complexity of each resulting fractionated protein sample) or by adding an extra dimension of liquid chromatography (thereby increasing the separation on the peptide level) prior to mass spectrometry (Peng et al. 2003; Washburn et al 2001; Wolters et al. 2001). By improving peptide separation, the complexity of the set of peptides entering the mass spectrometer at each point in time is reduced and the mass spectrometer has more time to sample peptides eluting from the column with reduced interference from higher abundance peptide species. These multidimensional separation methods

¹ For these strategies, animals are trained (i.e., conditioned) to seek or avoid alcohol under different experimental parameters. Researchers can then use consequences to modify the occurrence of the alcohol-seeking behavior.

² For this fragmentation, peptides are collided with noble gases in the mass spectrometer to break them up into smaller partial peptides. If breaks occur randomly along the backbone of a peptide, the amino acid sequence of the peptide can be determined by finding the mass differences between the charged mass “ladders.”

MICHAEL J. MACCOSS, PH.D., is an assistant professor in the Department of Genome Sciences at the University of Washington, Seattle, Washington.

CHRISTINE C. WU, PH.D., is an assistant professor in the Department of Pharmacology at the University of Colorado School of Medicine, Aurora, Colorado.

interfaced with fast-scanning mass spectrometers³ (Blackler et al. 2006; Schwartz et al. 2002) are very powerful and have improved researchers' ability to identify large numbers of proteins from complex mixtures. However, this chromatographic separation and biochemical fractionation is slow, often resulting in an analysis time of greater than 24 hours per sample (MacCoss et al. 2002; Washburn et al. 2001). These long analysis times make analyses of replicate samples, multiple conditions, and many time points prohibitive. As a result, multidimensional chromatography essentially is incompatible with the throughput required for proteomic analysis of large numbers of samples (such as those commonly acquired in a behavior study). Furthermore, the comparison of peptides between samples using this approach is complicated, as proteins expressed in low abundance rarely are sampled because of the semirandom sampling of spectra by data-dependent acquisition (Liu et al. 2004).

An alternative to multidimensional chromatography is to use different mass spectrometers in distinct ways to improve the handling of complex mixtures (Conrads et al. 2000; Hu et al. 2005; Syka et al. 2004). Recent advances in mass spectrometry have facilitated the routine acquisition of mass spectra at high resolution with sufficient resolution and peak capacity to handle the complexity of the mixture with only a single dimension of chromatographic separation. Furthermore, proteomics researchers are beginning to adopt mass spectrometry-based approaches originally developed for small molecule quantitation workflows for the targeted analysis of hypothesized proteins (Anderson and Hunter 2006; Barnidge et al. 2004; Gerber et al. 2003). These advances in proteomic technologies provide a glimpse of future possibilities and applications toward the dissection of the molecular mechanisms of behavior. The following sections focus on two emerging technologies that look extremely promising for large-scale shotgun proteomic endeavors: label-free differential mass spectrometry for discovery-based comparative proteomics and targeted mass spectrometry for hypothesis-driven quantitative proteomics.

LABEL-FREE DIFFERENTIAL MASS SPECTROMETRY FOR COMPARATIVE PROTEOMICS (DISCOVERY PLATFORM)

Most proteomics methods have been developed to acquire MS/MS spectra on as many molecular species as possible, regardless of whether those peptides are of interest. This semirandom sampling of peptides by data-dependent acquisi-

tion focuses on the most abundant peptides first and then samples lower intensity peptides in subsequent scan events. Although improvements in mass spectrometer scan speeds have increased the depth to which peptides can be sampled (Blackler et al. 2006; Mayya et al. 2005), this approach routinely wastes the majority of instrument time on sampling the most abundant molecular species in the sample.

Because of this limitation, some groups have begun to focus their identification efforts only on peptides that differ in abundance between samples (Finney et al. 2008; Pasa-Tolic et al. 2002; Prakash et al. 2006; Wiener et al. 2004). This experimental workflow is similar to that used in a two-dimensional (2D) gel analysis,⁴ except that these analyses are performed on the peptide level using chromatographic separation and not on the protein level using gel electrophoresis. Specifically, in a 2D gel experiment, gels are run first to identify spots that differ between samples and then efforts are taken to determine the identities of the proteins in those spots. Similarly, in a "differential" mass spectrometry experiment, chromatographic separa-

³ Some mass spectrometers (e.g., linear ion trap mass spectrometers) can scan on the millisecond scale.

⁴ 2D gel analysis (or electrophoresis) is used to separate proteins in an extract based on their electric charge (i.e., isoelectric point) and mass (i.e., molecular weight). Protein extracts typically are first applied to thin gel strips with a pH gradient and exposed to an electric current. Under the influence of this current, the proteins migrate through the gel strip, with the distance and direction traveled depending on the electric charge of the proteins. The gel strips then are loaded on a similar type of gel and exposed to a second electric current flowing in a direction perpendicular to the first one. Under these conditions, the proteins migrate from the initial gel strip into the second gel, with the distance traveled depending on the mass of the proteins. The proteins can then be made visible as "spots" on the gel through various staining procedures.

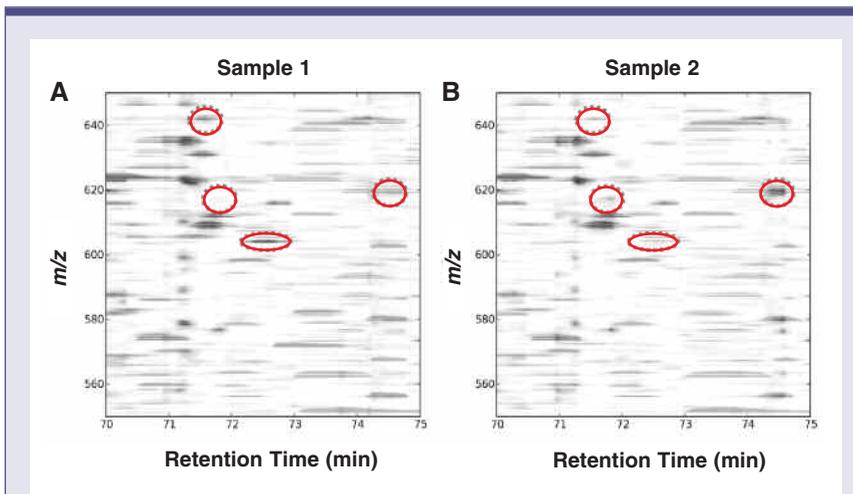


Figure 9 Finding differences between samples using differential mass spectrometry (dMS). Peptide maps are plotted as two-dimensional images following chromatogram alignment and intensity normalization. Statistical analysis software is used to find regions of mass-to-charge ratio (m/z) and retention time that differ in abundance between sample groups. Obvious visual differences (in the context of this figure) are illustrated in red.

tions are first carried out to identify peaks that differ in abundance between samples (see figure 9) and then efforts are taken to determine the identities of the peptides in those peaks. Peak abundances that change in a statistically

meaningful way (see figure 9, red circles) then are analyzed again to assign peptide identities.

The key experimental components to this differential mass spectrometry approach are (1) reproducible sample preparation, (2) an instrument that has sufficient peak capacity and dynamic range to handle complex mixtures with short analysis times, and (3) software for detecting regions of mass-to-charge ratio (m/z) and time that differ in abundance between samples. The mass chromatograms between experiments are first aligned, and subsequent comparisons between samples will evaluate the significance of intensity differences at any m/z in the elution profile (Finney et al. 2008; Meng et al. 2007). Advantages to this approach are the ability to measure differences without the use of stable isotope internal standards, a greater dynamic range for detection of differences, and the ability to measure differences between an unlimited number of aligned chromatograms. Finally, because chromatograms are aligned to one another, only a single MS/MS spectrum from any of the different samples (e.g., biological replicates or different time points) is needed to annotate the peptide identity of the difference region.

TARGETED MASS SPECTROMETRY (OR MS WESTERNS) FOR HYPOTHESIS-DRIVEN PROTEOMICS (QUANTITATIVE PLATFORM)

To address the limitations of sensitivity, selectivity, and throughput, there has been a recent shift toward the development and application of technologies for the targeted analysis of proteins within complex mixtures. Numerous derivations of targeted mass spectrometry using the specific acquisition of MS/MS spectra of peptides predicted *in silico* have been reported (Arnott et al. 2002; Chang et al. 2004), and more recently these methods have been based on the use of selected reaction monitoring (SRM) on triple quadrupole mass spectrometers (Anderson and Hunter 2006; Barnidge et al. 2004; Gerber et al. 2003). The concept of monitoring specific peptides from proteins of interest is

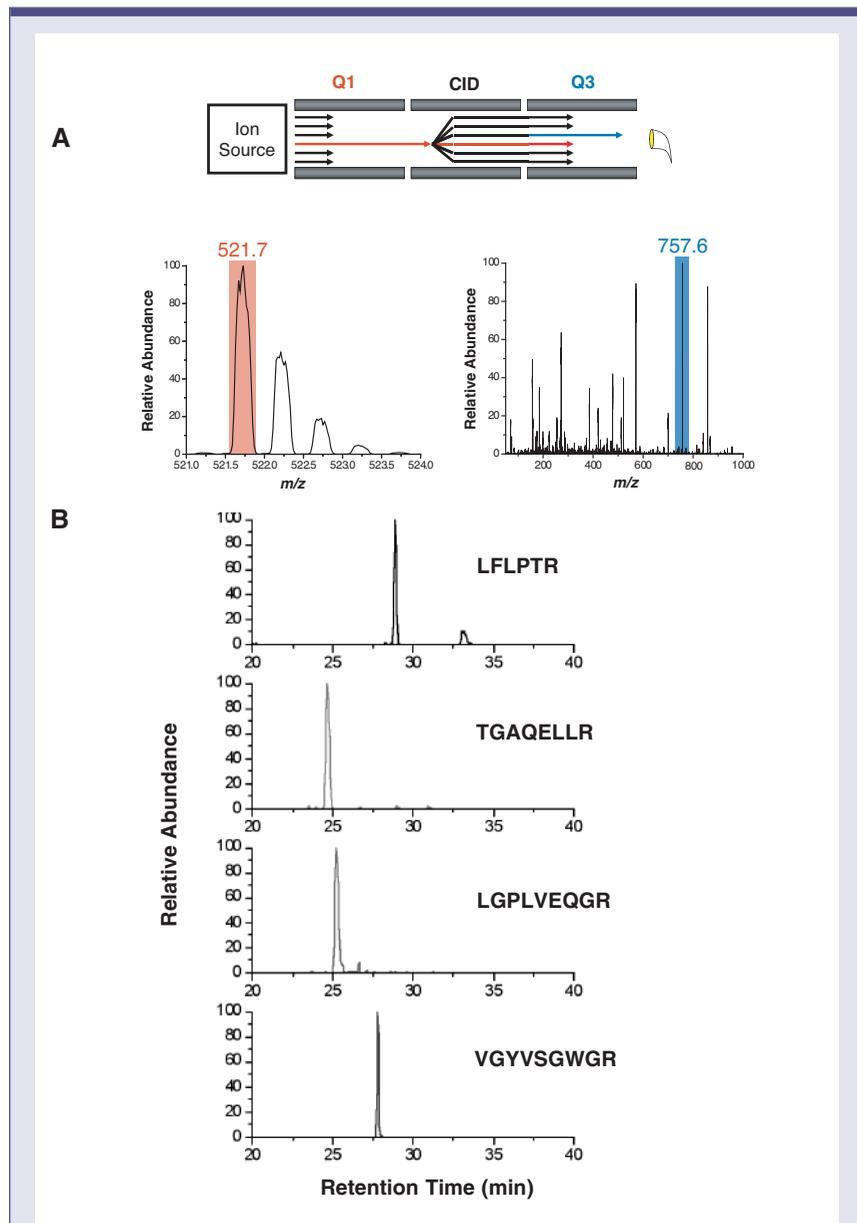


Figure 10 Illustration of selected reaction monitoring (SRM) on a triple quadrupole mass spectrometer. **A**) A predetermined precursor ion (mass-to-charge ratio, m/z 521.7) is selected in the first mass analyzer (Q1), fragmented by collision-induced dissociation (CID), and one of the resultant product ions (m/z 757.6) is selectively monitored in the second mass analyzer (Q3). **B**) Example of four different peptides generated by protein digestion with the enzyme trypsin (i.e., tryptic peptides) measured by SRM in human plasma.

well established. These methods have high specificity within a complex mixture and can be performed in a fraction of the instrument time compared with discovery-based methods. Ultimately, targeted studies are intended to complement discovery-based analyses and facilitate hypothesis-driven proteomic experiments.

Using the selectivity of multiple stages of mass selection of a tandem mass spectrometer (see figure 10), these targeted SRM assays are the mass spectrometry equivalent of a Western blot.⁵ In fact, Arnott and colleagues (2002) originally coined the term “MS Westerns” for the use of tandem mass spectrometry to target individual hypothesized peptides. Just as a Western blot only produces a signal for proteins within a complex mixture that are recognized by an antibody, a targeted mass spectrometry assay will only produce a signal for peptides that have a specific combination of precursor and product ion m/z . This combination of precursor and product ion m/z is extremely selective and referred to as a SRM transition. The advantage of using a targeted mass spectrometry-based assay is that it does not require creating any immunoaffinity reagents.

Targeted SRM assays can be developed for high-throughput assays that can measure multiple analytes in a single analysis (i.e., multiplexed quantitative assays) (Anderson and Hunter 2006). In a complex mixture, the chemical background of analytes that elute from a chromatography column at the same time (i.e., co-eluting analytes) can prohibit detection of a precursor ion in a data-dependent analysis. However, if the m/z of the precursor ion is known, a triple quadrupole mass spectrometer can be used to minimize the chemical interference using two separate stages of mass analysis to selectively monitor a unique peptide. The combined specificity of chromatographic retention time, precursor ion mass, and product ion mass allow the selective detection of a unique peptide and its respective protein within a complex mixture. Importantly, by combining SRM with peptide standards of known quantities, absolute peptide amounts can be determined (Desiderio and Kai 1983; Desiderio et al. 1983; Gerber et al. 2003; MacCoss and Matthews 2005). Furthermore, a mass spectrometry-based assay facilitates the use of stable isotope-labeled analogs of proteins/peptides as internal standards that can account for errors in sample preparation or the mass spectrometric measurement. Targeted SRM assays can be multiplexed to quantify multiple transitions for multiple peptides in any given experiment and can be automated for high-throughput analyses.

⁵ A Western blot (or immunoblot) is an analytical technique used to detect specific proteins in a given protein sample that involves three steps: (1) gel electrophoresis to resolve the proteins in the sample, (2) transfer of the proteins onto a membrane, and (3) detection of a selected protein using an immune molecule (i.e., antibody) that will bind specifically to that protein.

CONCLUSIONS AND FUTURE DIRECTIONS

The proteomic analysis of brain tissue poses a complex analytical problem; this problem is further exacerbated in behavioral studies. However, given the throughput capability of differential mass spectrometry and the software available for the analysis of these data, combined with the sensitivity and multiplexing capability of targeted mass spectrometry, the promise of large-scale proteomics analyses in neuroscience is a certainty. As always, however, the “devil is in the details,” and although these technologies currently are possible, it will require continued application-driven developments in mass spectrometry hardware, sample preparation, automated sample handling, and computational analysis before these approaches become routine. ■

FINANCIAL DISCLOSURE

The authors declare that they have no competing financial interests.

REFERENCES

- ANDERSON, L. Candidate-based proteomics in the search for biomarkers of cardiovascular disease. *Journal of Physiology* 563:23–60, 2005. PMID: 15611012
- ANDERSON, L., AND HUNTER, C.L. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Molecular and Cellular Proteomics* 5:573–588, 2006. PMID: 16332733
- ARNOTT, D.; KISHIYAMA, A.; LUIS, E.A.; ET AL. Selective detection of membrane proteins without antibodies: A mass spectrometric version of the Western blot. *Molecular and Cellular Proteomics* 1:148–156, 2002. PMID: 12096133
- BARNIDGE, D.R.; GOODMANSON, M.K.; KLEE, G.G.; AND MUDDIMAN, D.C. Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *Journal of Proteome Research* 3:644–652, 2004. PMID: 15253448
- BLACKLER, A.R.; KLAMMER, A.A.; MACCOSS, M.J.; AND WU, C.C. Quantitative comparison of proteomic data quality between a 2D and 3D quadrupole ion trap. *Analytical Chemistry* 78:1337–1344, 2006. PMID: 16478131
- CHANG, E.J.; ARCHAMBAULT, V.; MCLACHLIN, D.T.; ET AL. Analysis of protein phosphorylation by hypothesis-driven multiple-stage mass spectrometry. *Analytical Chemistry* 76:4472–4483, 2004. PMID: 15283590
- CONRADS, T.P.; ANDERSON, G.A.; VEENSTRA, T.D.; ET AL. Utility of accurate mass tags for proteome-wide protein identification. *Analytical Chemistry* 72:3349–3354, 2000. PMID: 10939410
- CRABBE, J.C. Alcohol and genetics: New models. *American Journal of Medical Genetics* 114:969–974, 2002. PMID: 12457395
- DESIDERIO, D.M., AND KAI, M. Preparation of stable isotope-incorporated peptide internal standards for filed desorption mass spectrometry quantification of peptides in biologic tissue. *Biomedical Mass Spectrometry* 10:471–479, 1983. PMID: 6616023
- DESIDERIO, D.M.; KATAKUSE, I.; AND KAI, M. Measurement of leucine enkephalin in caudate nucleus tissue with fast atom bombardment-collision activated dissociation-linked field scanning mass spectrometry. *Biomedical Mass Spectrometry* 10:426–429, 1983. PMID: 6616019

- FINNEY, G.L.; BLACKLER, A.R.; HOOPMANN, M.R.; ET AL. Label-free comparative analysis of proteomics mixtures using chromatographic alignment of high-resolution μ LC-MS data. *Analytical Chemistry* 80:961–971, 2008. PMID: 18189369
- GERBER, S.A.; RUSH, J.; STEMMAN, O.; ET AL. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences of the United States of America* 100:6940–6945, 2003. PMID: 12771378
- GHAEMMAGHAMI, S.; HUH, W.K.; BOWER, K.; ET AL. Global analysis of protein expression in yeast. *Nature* 425:737–741, 2003. PMID: 14562106
- GRANT, K.J., AND WU, C.C. Advances in neuromembrane proteomics: Efforts towards a comprehensive analysis of membrane proteins in the brain. *Briefings in Functional Genomics & Proteomics* 6:59–69, 2007. PMID: 17353200
- HU, Q.; NOLL, R.J.; LI, H.; ET AL. The Orbitrap: A new mass spectrometer. *Journal of Mass Spectrometry* 40, 430–443, 2005. PMID: 15838939
- LIU, H.; SADYGOV, R.G.; AND YATES, J.R. III. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Analytical Chemistry* 76:4193–4201, 2004. PMID: 15253663
- MACCOSS, M.J., AND MATTHEWS, D.E. Quantitative MS for proteomics: Teaching a new dog old tricks. *Analytical Chemistry* 77:294A–302A, 2005. PMID: 16134292
- MACCOSS, M.J.; McDONALD, W.H.; SARAF, A.; ET AL. Shotgun identification of protein modifications from protein complexes and lens tissue. *Proceedings of the National Academy of Sciences of the United States of America* 99:7900–7905, 2002. PMID: 12060738
- MAYYA, V.; REZAU, K.; CONG, Y.S.; AND HAN, D. Systematic comparison of a two-dimensional ion trap and a three-dimensional ion trap mass spectrometer in proteomics. *Molecular and Cellular Proteomics* 4:214–223, 2005. :15608339
- MENG, F.; WIENER, M.C.; SACHS, J.R.; ET AL. Quantitative analysis of complex peptide mixtures using FTMS and differential mass spectrometry. *Journal of the American Society for Mass Spectrometry* 18:226–233, 2007. PMID:17070068
- MULLIGAN, M.K.; PONOMAREV, I.; HITZEMANN, R.J.; ET AL. Toward understanding the genetics of alcohol drinking through transcriptome meta-analysis. *Proceedings of the National Academy of Sciences of the United States of America* 103:6368–6373, 2006. PMID: 16618939
- PASA-TOLIĆ, L.; LIPTON, M.S.; MASSELON, C.D.; ET AL. Gene expression profiling using advanced mass spectrometric approaches. *Journal of Mass Spectrometry* 37:1185–1198, 2002. PMID: 12489076
- PENG, J.; ELIAS, J.E.; THOREEN, C.C.; ET AL. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: The yeast proteome. *Journal of Proteome Research* 2:43–50, 2003. PMID: 12643542
- PRAKASH, A.; MALLICK, P.; WHITEAKER, J.; ET AL. Signal maps for mass spectrometry-based comparative proteomics. *Molecular and Cellular Proteomics* 5:423–432, 2006. PMID: 16269421
- SCHWARTZ, J.C.; SENKO, M.W.; AND SYKA, J.E. A two-dimensional quadrupole ion trap mass spectrometer. *Journal of the American Society for Mass Spectrometry* 13:659–669, 2002. PMID: 12056566
- SYKA, J.E.; MARTO, J.A.; BAI, D.L.; ET AL. Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *Journal of Proteome Research* 3:621–626, 2004. PMID: 15253445
- WASHBURN, M.P.; WOLTERS, D.; AND YATES, J.R. III. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology* 19:242–247, 2001. PMID: 11231557
- WIENER, M.C.; SACHS, J.R.; DEYANOVA, E.G.; AND YATES, N.A. Differential mass spectrometry: A label-free LC-MS method for finding significant differences in complex peptide and protein mixtures. *Analytical Chemistry*