Measuring modifications

Researchers are using MS to study proteins and how they interact with their environment.

BY RANDALL C. WILLIS

Now that the genome is well on its way to being understood, the real interest in molecular biology is in the behavior of the proteome—the complete complement of proteins in an organism. Making sense of the proteome requires a host of new tools, or rather, new methods of applying time-honored instruments, with MS being the key.

At the heart of the matter is the fact that cell and organism proteomes are complex in both protein variety and abundance. Whereas the genome consists of a relatively static array of DNA sequences, the proteome is composed of a finite number of proteins that can be modified in a seemingly infinite number of ways. These modifications can come at the post-transcriptional (e.g., alternate splicing) or post-translational (e.g., phosphorylation, glycosylation) levels.

Glycoproteins

MS has proven useful in the study of protein glycosylation, but it suffers from the poor ionization of glycosylated peptides compared with their unmodified forms, and episodes of gas-phase deglycosylation can make finding the site of modification difficult.

James Stephenson and colleagues at Purdue University (*www.purdue.edu*) and Oak Ridge National Laboratory (*www.ornl.gov*) have had some success in recent years with a "top-down" approach to protein sequence analysis using MS. Rather than extensively purifying a protein and digesting it with proteases, as is typical for MS, Stephenson's group subjects whole proteins to tandem MS, finding that the proteins tend to fragment preferentially at sites such as the N-terminal of proline and the C-terminal of aspartic acid, lysine, arginine, and histidine. But no one had determined how the fragmentation site preferences were affected by post-translational modifications or whether the modifications would survive the fragmentation process.

To address these questions, Stephenson's group performed whole-protein MS analysis on ribonuclease A and ribonuclease B, two proteins that differ solely in the attachment of an N-linked sugar (*1*). Mass spectra of the unglycosylated ribonuclease A showed that the protein fragmented at the same preferred residues as the proteins studied earlier. This pattern also held true for the glycosylated ribonuclease B. The only place where the two spectra differed was in the migration of the sugar-bearing peptide, and the difference was the molecular mass of the sugar moiety (Figure 1).

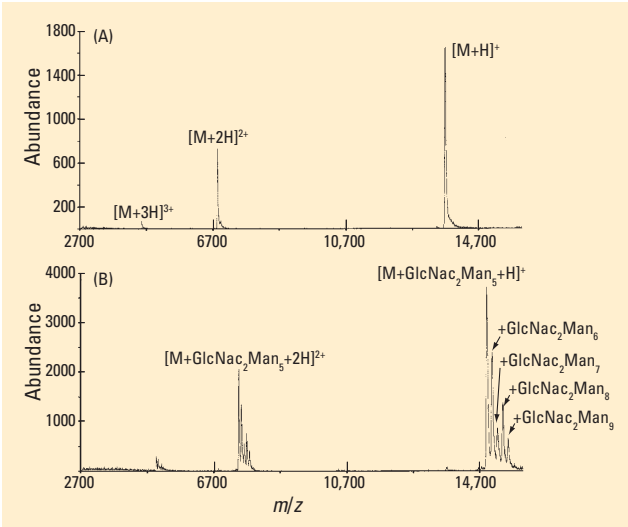

Figure 1. Sugar, sugar. The mass spectra of unglycosylated ribonuclease A (top) and glycosylated ribonuclease B (bottom) show the effects that sugar moieties can have on peptide ion formation. (Adapted with permission from Ref. 1.)

Because errant glycosylation is associated with some diseases, the researchers have developed a potentially useful tool.

Phosphoproteins

Phosphorylation plays a dominant role in cell-signaling pathways. Thus, a thorough understanding of where and when this event occurs is critical to determining how cells function. Various methods have been used to identify which amino acids in a protein are phosphorylated, but these methods are largely qualitative. To answer these questions quantitatively, Meredith Bond and colleagues at Case Western Reserve University School of Medicine (*http://mediswww.cwru.edu*) and the Cleveland Clinic Foundation (*www.clevelandclinic.org*) developed a capillary LC/electrospray ionization (ESI)-MS system that compares a phosphopeptide of interest to an internal reference peptide (*2*).

Normally, when a reference peptide is used in MS, it is added to the peptide mixture just before MS analysis. Thus, it does not go through the same reaction and purification steps as the native peptides, so it cannot be compared quantitatively. Bond's group developed the "native reference peptide" method, which relies on a peptide that is present in the protein digest being studied but is unmodified (Figure 2). This method also eliminates experimentally introduced variations. By taking the ratio of the MS peak areas of the reference to those of the peptide of interest, the researchers quantitatively assess the degree of phosphorylation. This ratio can then be converted to moles of phosphorylation per mole of protein by comparing the peak areas to a calibration curve of synthetic peptide standards. The researchers tested their method by following the phosphorylation kinetics of ser-

KEY TERMS: cell biology, clinical, high throughput, informatics, proteomics, screening, technique

ine-44 (Ser44) in cardiac troponin I (cTnI), a protein that might serve as a biomarker for heart disease.

The scientists treated cTnI with protein kinase C βII and adenosine 5′-triphosphate (ATP), and ran aliquots on denaturing polyacrylamide gels. They then digested the cTnI bands in-gel with trypsin and analyzed the resulting peptides by capillary LC/ESI-MS. They concentrated their efforts on the native reference peptide ^{121}NITEIADLTQK^{131} and the peptide ^{41}ISASR^{45}, where Ser44 is phosphorylated. As cTnI is phosphorylated, the phosphopeptide peak slowly replaces the nonphosphorylated peptide peak. This is indicated by changes (in the opposite direction) in the ratio of the chromatographic peak area of each of the two test peptides to the reference peptide.

Similarly, Brandon Ruotolo and colleagues at Texas A&M University (*www.tamu.edu*) and the U.S. National Institute on Drug Abuse (*www.nida.nih.gov*) developed a method that uses matrix-assisted laser desorption-ionization (MALDI)-ion mobility (IM)-time-of-flight (TOF) MS to distinguish between phosphorylated and nonphosphorylated peptides (*3*). In this case, the separation of the phosphorylated peptides relies on the fact that some peptide ions with well-defined secondary or tertiary structure will deviate from the traditionally linear mass-mobility relationship in the IM drift tube.

The researchers found that in the gas phase, phosphopeptides had a higher mobility than predicted solely on the basis of size and charge, corresponding to a smaller collision cross section (i.e., a more compact state). They confirmed this fact by performing molecular dynamics simulations to generate low-energy structures of several test peptides and phosphopeptides, and showed that distinct interactions occurred between amino acid side chains and the phosphate groups.

Lipoproteins

Lipids are also fundamental modifications that play an important role in protein structure and function, and there has been particular interest over the past several years regarding the effects on health of circulating high- and low-density lipoprotein (HDL and LDL) complexes. The presence of a lipid moiety complicates the process of eluting the proteins from a polyacrylamide gel because of strong hydrophobic interactions between the lipids and matrix. Thus, Ronald Macfarlane and colleagues at Texas A&M University tried to develop a method of isolating and characterizing lipoproteins from HDL—mainly apo A-I and II and apo C-I, II, and III—using immobilized pH gradient isoelectric focusing (IPG-IEF) and MALDI MS (*4*).

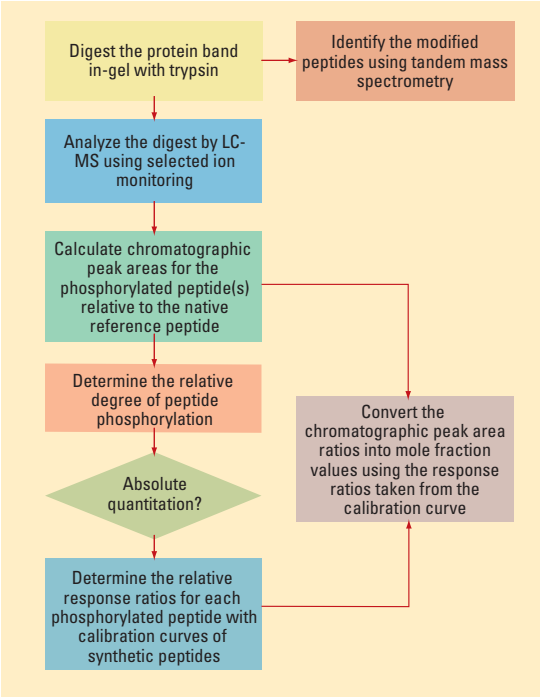

Figure 2. Following phosphorylation. The native reference peptide method of analyzing protein phosphorylation. (Adapted with permission from Ref. 2.)

To test their system, the scientists used IPG-IEF to create a pI profile of commercially available apo A-I and found that they could visualize 12 isoforms of the protein, 7 more than had been previously identified. They then tried to isolate the proteins from the gel bands using a variety of extraction methods, ultimately finding that a solution of formic acid, acetonitrile, 2-propanol, and water yielded the most protein (25%).

The researchers repeated their experiments with HDL fractions extracted from human blood and noted the appearance of several bands over the same pH range as the purified apo A-I. They then extracted protein from several gel slices along the pH range, and using MALDI MS, identified apo C-III isoforms (pH 4.0–4.5), apo A-II (pH 4.5–4.6), and apo A-I isoforms (pH 4.9–5.5). They also noted the appearance of other blood proteins, including human serum albumin and serum amyloid A.

The researchers are confident that their system can detect lipoprotein mutations and post-translational modifications associated with healthy and diseased states, making it a potentially useful prognostic and diagnostic tool.

Peptide–peptide interactions

MS has proven to be an essential tool for identifying isolated proteins, but in cells, proteins interact with other macromolecules. Thus, it was critical that MS technologies be developed for detecting and analyzing noncovalent complexes.

To determine the factors involved in the binding of one protein or peptide to another, ESI-MS has typically been the method of choice, but peptide–peptide interactions are often studied in the presence of lipids, detergents, or salts that suppress ESI. In such cases, MALDI appears to be the more suitable ionization choice. Amina Woods and her colleagues at the U.S. National Institute on Drug Abuse had previously shown that MALDI MS could be used to observe noncovalent complexes that involve a salt bridge between acidic and basic peptides, and that this association is strongly pH-dependent. Recently, however, Woods extended this analysis by incorporating IM into the process (*5*).

Desorption from MALDI into the high-pressure IM drift cell creates a different set of ionization conditions from standard high-vacuum MALDI, and recent evidence suggests that the buffer gas might stabilize ions and noncovalent complexes. Woods' group studied this possibility by looking for evidence of complex formation between the acidic hormone fragment Mini Gastrin I and the basic opioid peptide dynorphin 1-7.

The researchers found a near-linear relationship between m/z and mobility drift

time, such that they could distinguish signals from each of the peptides, the dynorphin–Mini Gastrin complex, and a Mini Gastrin homodimer, the last of which was not observed using ESI. The complex was, as expected, pH-dependent. Interestingly, the complex also showed a small degree of fragmentation, which was indicated by a peak that had the mobility of the complex but the *m/z* of dynorphin. Future research will focus on applying the technology to biological samples to look for biologically relevant complexes.

The applications of MS in proteomics and drug discovery are expanding rapidly, and FT-ion cyclotron resonance (ICR)-MS offers the added benefit of being able to identify noncovalent complexes using "soft" ionization techniques such as ESI and MALDI.

Peptide–disaccharide

Sugars play a large but undervalued role in cell functions such as biological recognition and signaling. Therefore, an understanding of how proteins and sugars interact is critical. To address this problem, Hans-Joachim Gabius and colleagues at various institutes recently undertook a study of the minimal structural requirements for sugar binding by polypeptides (*6*).

Animal lectins typically comprise more than 100 amino acids, complicating the analysis of their carbohydrate-binding behavior. Recent phage-display experiments with a library of pentadecapeptides, however, identified three peptides with high affinity for the Thomsen-Friedenreich (TF) antigen. The small size of the oligopeptides provides for potentially favorable pharmacokinetic properties and therefore might facilitate their use in biomedical applications.

Gabius and colleagues tested the peptides in both the absence and the presence of TF antigen and maltose (a disaccharide control) by NMR spectroscopy, ESI-MS, and molecular modeling to determine how the molecules interact. The NMR studies suggested that the peptides alone were predominantly unfolded and that the aromatic residues, long thought to form a hydrophobic core for folding, did not do so. This unfolded state remained largely unchanged when TF antigen was added, although there were specific interactions between the sugar and peptides that did not occur with maltose.

The bottom line

When moving from genome to proteome, it appears that all the current tricks of the trade (and more) will be required. And ultimately, MS techniques will remain at the heart of forward motion. But they are not your old professor's MS methods; rather, they are a complex of new adaptations of hyphenated and tandem instruments coupled to some good old-fashioned, and some fairly new, chemical techniques.

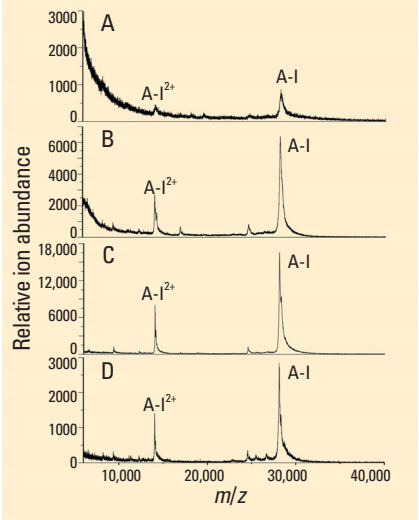

Figure 3. Shifting priorities. Differences in apo A-I lipid content are seen both in an unstained IPG gel (A–D) and in the MALDI-TOF mass spectra of recovered proteins from labeled excised bands. (Adapted with permission from Ref. 4.)

References

(1) Reid, G. E.; et al. *Anal. Chem.* **2002,** *74,* 577–583.
(2) Ruse, C. I.; et al. *Anal. Chem.* **2002,** *74,* 1658–1664.
(3) Ruotolo, B. T.; et al. *J. Proteome Res.* **2002,** *1,* 303–306.
(4) Farwig, Z. N.; Campbell, A. V.; Macfarlane, R. D. *Anal. Chem.* **2003,** *75,* 3823–3830.
(5) Woods, A. S.; et al. *J. Amer. Soc. Mass Spectrom.* **2002,** *13,* 166–169.
(6) Siebert, H.-C.; et al. *Biochemistry* **2002,** *41,* 9707–9717.

Randall C. Willis is senior associate editor of *Modern Drug Discovery*. Send your comments or questions about this article to mdd@acs.org or to the Editorial Office address on page 3. ■