Recent advances in absolute quantification of peptides and proteins using LC-MS

Abstract: Current bioanalytical methods for the quantification of peptides and proteins used in biomedical research and pharmaceutical R&D are predominantly based on immunoassay methodologies, such as enzyme-linked immunosorbent assays (ELISA). Although these antibody-based techniques are exceptionally sensitive, they suffer from issues, such as cross-reactivity, long development time, and high variability. Recently, liquid chromatography-mass spectrometry (LC-MS)-based quantification techniques for large molecules, including the bottom-up and top-down approaches, are gaining popularity due to their shorter development time, higher specificity, and ability to detect degradation products and post-translational modifications (PTMs). The multiple reaction monitoring (MRM)-based bottom-up LC-MS/MS approach, which targets enzymatic digestion-produced signature peptides as surrogates in a quantification assay, is a sophisticated method with good sensitivity, selectivity, and multiplex capacity, although it may miss critical information such as PTMs, sequence variants and isoforms, and truncation and degradation products. The LC-high resolution mass spectrometry (LC-HRMS)-based top-down approach quantifies intact target molecules, thus avoiding the shortcomings of the bottom-up approach. The current LC-HRMS approach is most suitable for peptides and proteins smaller than 10–15 kDa, though that may change with the development of more advanced mass spectrometers with higher resolving power and broader mass ranges. The drawbacks and bottlenecks of LC-MS-based large molecule quantification assays are low sensitivity, complicated sample preparation, and relatively low throughput. Technology advancement may address these challenges and expand their suitable applications. In this article, the recent advances and trends in this emerging field are reviewed, along with the basic workflows, analytical approaches, and other common issues.

Keywords: absolute quantification; bioanalysis; LC-MS/MS, LC-HRMS; protein quantification.

Introduction

In recent years, there is a rapidly increasing demand for quantification of peptides and proteins due to their proven promise as therapeutic agents. In addition, the shift of the drug discovery and development from “trial and error” to target-driven approaches has also created a demand for reliable quantification of peptide and protein biomarkers. Currently, the bioanalytical methods used in the quantification of peptides and proteins, in support of pharmacokinetic and biomarker studies, are mainly based on immunoassay methodologies. Although these antibody-based techniques are exceptionally sensitive, issues such as cross-reactivity, long development time, and high variability are stimulating interest in alternative assay platforms. At present, liquid chromatography-mass spectrometry (LC-MS)-based techniques, considered the “gold standard” for small molecules, are gaining increased interest in the field of large molecules due to their unique advantages (Ciccimaro and Blair 2010, Li et al. 2011a).

LC-MS-based assays have features that distinguish them from conventional ligand binding assays, such as enzyme-linked immunosorbent assays (ELISA), as summarized in Table 1. LC-MS assays have several important advantages. First, they require less development time. For example, two specific antibodies are required for each protein analyte in a sandwich ELISA, a typical immunoassay. The process of antibody production and selection, which is resource- and time-consuming, contributes to long assay development time. For example, two specific antibodies are required for each protein analyte in a sandwich ELISA, a typical immunoassay. The process of antibody production and selection, which is resource- and time-consuming, contributes to long assay development time. It may take up to a few months to develop and validate an immunoassay for toxicological or clinical studies. In contrast, a typical LC-MS-based assay that does not require an immuno-affinity enrichment process may take just days (or weeks) to develop, which is very attractive for discovery stage activities before substantial time and resources are committed to raise antibodies. The second advantage is that mass spectrometry can provide much higher specificity, which is a potential advantage when an immunoassay is subject to cross-reactivity or interference. While the presence of analogs to the target protein could interfere with an antibody-based assay, an LC-MS based assay would...
typically provide the capability to differentiate the interference from the targeted analyte. Finally, due to its high specificity, an LC-MS-based assay can be used to detect degradation product of proteins, as well as their post-translational modifications (PTMs), such as glycosylation and phosphorylation. Innovated mass spectrometers and software have also expanded researchers’ capability to conduct post-acquisition data-mining, which facilitates characterization and helps scientists understand the overall properties of molecular entities.

Currently, there are two major drawbacks of LC-MS assays that prevent them from being routinely used in bioanalytical work in an industrial setting. First, detection sensitivity is usually lower on LC-MS than in immunoassay-based platforms. The detection limit for a typical LC-MS assay is usually in the low μg/ml range when no immuno-enrichment of the target protein is applied, whereas the detection limit for a typical ELISA assay is often in the pg/ml range. Second, assay throughput for an LC-MS platform is lower than that of immunoassays. The current LC-MS approach often requires multiple steps of analyte enrichment, enzyme digestion, and clean-up, most of which are not adapted to the high-throughput microplate formats. Thus, it can take up to several days to process one batch of samples for LC-MS analysis, whereas only hours are required for an ELISA assay. The LC-MS data acquisition also involves many hours of injections and runs, yet it may only take a couple of minutes to read an ELISA assay plate. Nevertheless, it is believed that these roadblocks can be overcome with the evolving LC-MS technologies. It is expected that the platform will someday be routinely applied to bioanalytical support for large molecules in drug discovery and development.

In this article, we will review the recent advancements and trends in this emerging field, along with an introduction to the basic workflows, analytical approaches, and common issues. The primary focus of this review will be the targeted, absolute quantification of therapeutic biologics and biomarkers, excluding proteomics work, which is often conducted in a non-targeted, relative quantification fashion. We will particularly discuss liquid chromatography-high-resolution mass spectrometry (LC-HRMS) for intact peptide and protein analysis, an LC-MS strategy that is fast gaining popularity due to the development of improved mass spectrometry instruments.

**Workflows for quantification of peptides and proteins using LC-MS**

**Overview**

There are two common approaches for LC-MS-based quantification of peptides and proteins. (1) The bottom-up approach involves the enzymatic cleavage of a peptide or protein into small peptides, followed by the LC-MS/MS analysis of one or more of these proteolytic peptides, which are called signature peptides, as surrogate(s). (2) Top-down approach refers to analysis of an intact target molecule. This approach is often limited by the size...
of the analyte and is most suitable for peptides and proteins smaller than 10–15 kDa.

The bottom-up quantification approach has evolved from proteomics research. Many proteomics methods are based on enzymatic digestion of proteins into small peptides that are readily analyzed by mass spectrometers. Over the years, sensitive and selective bottom-up methods that are capable of identifying thousands of proteins within a single sample have been developed (Graumann et al. 2008). The bottom-up approach, however, has limitations when used for quantification, because it only looks at certain peptides of a protein, rather than the whole protein itself. Thus, it may miss critical information, such as PTMs, sequence variants and isoforms, as well as truncation and degradation products. The top-down quantification method, meanwhile, overcomes many of these shortcomings. Given that no proteolytic digestion is involved, the top-down approach analyzes intact peptides and proteins. Thus, it is capable of detecting PTMs (e.g., phosphorylation and acetylation), sequence variants (e.g., mutants, amino acid polymorphisms, and isoforms), and truncation or degradation products. A particular advantage of the top-down approach with full scan mass spectrometer acquisition mode is its data-mining capacity once LC-MS data are acquired. The full scan mode captures all information of a sample, and such data can be reprocessed post-acquisition depending on different requirements.

Due to the wide distribution of molecular weights and the highly diverse properties of peptides and proteins, there is no universal workflow suitable for all LC-MS analyses. Here, we propose a general workflow for LC-MS-based peptide and protein quantification, as shown in Figure 1. Depending on the size of a target analyte, sample composition, and sensitivity requirement, the sample preparation can range from a single step of protein precipitation to multiple steps that may involve the depletion of high-abundant proteins, fractionation, enzymatic digestion, and so on (Lu et al. 2009). A variety of LC modes can be selected to achieve optimal separation of the analyte from matrix components. For mass spectrometer detection, traditional LC-MS/MS acquisition modes as well as novel approaches, such as high resolution (HR) full scan MS, have been explored for either signature peptide or intact peptide and protein detection. In this section, recent advancements in major techniques of sample preparation, liquid chromatography, and mass spectrometer detection will be reviewed for their application in peptide and protein bioanalysis.

**Sample preparation**

**Protein precipitation**

For peptides and proteins smaller than 10–15 kDa, protein precipitation or solid phase extraction (SPE) is the preferred sample preparation method to separate them from a biological matrix. The relative small sizes of these peptides and proteins also make them good candidates for top-down analysis. Protein precipitation using acetonitrile or methanol can remove most of the larger proteins in a biological matrix while retaining peptides smaller than 1.5 kDa in the solvent. Peptides and proteins larger than 1.5 kDa tend to be insoluble in high organic contents and can be processed using SPE. PEGylated proteins, however, represent exceptions for this method. The attachment of polyethylene glycol (PEG) to peptides or proteins has been employed as a strategy to extend their in vivo circulatory half-life as well as to improve their chemical and physical stability, solubility, and potentially to reduce immunogenicity (Veronese and Mero 2008). Due to the PEG moiety, a PEGylated protein can remain in the extract supernatant during protein precipitation. Wu et al. (2011) reported a protein precipitation method, in which PEGylated proteins (protein drugs of 11–12 kDa coupled to 40 kDa PEG) were extracted using acidic isopropanol. Meanwhile, Xu et al. (2010) used protein precipitation to extract a peptide of 38 amino acids attached to a 40 kDa branched PEG from human plasma.

**Solid phase extraction (SPE)**

When protein precipitation is not applicable, SPE may provide a more selective way of sample cleanup. SPE also takes advantage of the size-exclusion principle by excluding large proteins that cannot enter the pores on the SPE sorbent. For example, a pore size of 40–80 Å can roughly eliminate
proteins larger than 20 kDa. Very often, ion exchange SPE can provide more selective cleanup as well as an orthogonal mode of separation when coupled with reversed-phase liquid chromatography (RPLC). In our laboratory, we used reversed phase SPE to extract different glycosylation forms of apolipoprotein C3 (ApoC3) proteins, which have molecular weights of roughly 8.8–9.7 kDa, from human plasma (Jian et al. 2013).

**Enzymatic digestion**

For most proteins larger than 10–15 kDa, there is no straightforward method to separate them from biological matrixes (except for immuno-capture, which is discussed later in this review). Size exclusion techniques have been shown to be somewhat successful in isolation of high molecular weight components and exclusion of unwanted matrix contaminants (Crowther and Venturini, unpublished data). These proteins are too large to be directly quantified using standard LC-MS assays and, therefore, need to be enzymatically digested into smaller peptides. Global enzymatic digestion has been used to digest all proteins, including target molecule, in a whole biological sample (Anderson and Hunter 2006, Kirsch et al. 2007, Kuzyk et al. 2009). Alternatively, enzymatic digestion is often performed after cleanup of samples or enrichment of target analytes (Li et al. 2011a). To improve the efficiency of digestion, different acceleration approaches can be utilized, such as high temperature, microwave, and organic solvent (Berna and Ackermann 2009, Li et al. 2009). Recently, Ouyang et al. (2012) developed a simple and efficient “pellet digestion” method based on trypsin digestion of protein analytes precipitated from plasma samples. Commonly used proteases for mass spectrometer analysis are trypsin, LysC, GluC, ApsN, and chymotrypsin, among others. Trypsin is a serine protease that specifically cleaves a protein at the carboxyl side of arginine and lysine residues (except when blocked by an adjacent carboxylic side proline residue). The average tryptic peptide is ideal for LC-MS analysis due to its relatively small size and single C-terminal-located basic residue. In many cases, the combination of two proteases is necessary if good signature peptides cannot be produced by single protease digestion.

**Depletion of high-abundant proteins**

In many cases, sample preparation of protein targets larger than 10–15 kDa requires more than enzymatic digestion due to sample complexity. An example is human plasma, where the total concentration of endogenous proteins is as high as 60–80 mg/ml. The complexity of the plasma proteome, in which the dynamic range of protein abundance covers 11 orders of magnitude, presents a great challenge to the quantification of selected proteins in plasma (Anderson and Anderson 2002). A simple global digestion of a plasma sample will not overcome the sensitivity issue in detecting low abundant proteins. Therefore, enrichment procedures prior to or after enzymatic digestion are often needed to reduce interference and improve sensitivity for LC-MS/MS analysis of a target protein, especially those circulating at low concentrations. One such enrichment procedure is the depletion of high-abundant serum proteins. It is estimated that 99% of the serum’s total protein mass is due to the top 20 most abundant protein species. Human serum albumin, the most abundant protein, is present at 35–50 mg/ml. Depletion kits using chemical affinity or immuno-affinity for the removal of serum albumin, immunoglobulin, and other high-abundant proteins have been shown to reduce protein content by up to 85% (Echan et al. 2005). Polasova et al. (2010) studied six commercial products for high abundance protein removal and showed that while performance varied, these products generally improved protein detection on 2D gel. Anderson and Hunter (2006) demonstrated that the depletion of six high-abundant proteins improved the reproducibility of LC-MS/MS-based protein quantification assay in human plasma. Current commercially-available technologies allow the removal of up to 20 high-abundant proteins, greatly improving the sensitivity of LC-MS detection.

**Immuno-capture enrichment**

Specific or nonspecific enrichment strategies can be applied to enrich a target analyte. Of these, immuno-capture enrichment is the most commonly used method. Examples of immuno-capture enrichment include capture of a target protein drug or signature peptides by specific antibodies, affinity enrichment of antibody drugs containing the Fc region of IgG by protein A and/or protein G, and capture of PEGylated proteins by anti-PEG antibody (Xu et al. 2010, Liu et al. 2011). Anderson et al. (2004) introduced a workflow called Stable Isotope Standards with Capture by Anti-Peptide Antibodies (SISCAPA) for the quantification of proteins using signature peptides. One key step of this workflow is the immuno-capture of signature peptides by polyclonal anti-peptide antibodies before LC-MS/MS analysis (Anderson et al. 2004). The advantage of this workflow is that, in theory, any low-abundant
protein may be potentially detected due to the affinity enrichment using an anti-peptide antibody. However, the drawback is the requirement of an anti-peptide antibody to be generated and used in this workflow, although polyclonal antibodies can be generated fairly quickly and at a reasonable cost in many instances using current technologies. The immuno-capture enrichment can also be performed online using a column loaded with antibody (Dufield and Radabaugh 2012).

**Bottom-up workflow for the quantification of peptides and proteins**

A typical bottom-up workflow for the quantification of large proteins (≥10–15 kDa) by targeted LC-MS/MS involves the enzymatic cleavage of a protein followed by the LC-MS/MS analysis of one or more so-called “signature peptides” in multiple reaction monitoring (MRM) mode. The data acquisition is normally performed on a triple quadrupole mass spectrometer with an electrospray ion source, although there has been a push to use high resolution mass spectrometers (HRMS), such as Orbitrap, in such applications. The most commonly used enzyme for protein digestion is trypsin. Positive-mode electrospray ionization (ESI) of tryptic peptides normally generates M+2H+ and M+3H+ molecular ions with mass/charge (m/z) values between 400 and 1100. Collision-induced dissociation (CID) fragments these molecule ions into predominately C-terminal y- and N-terminal b-ion series (Coon 2009). Signature peptides that are predicted to uniquely and stoichiometrically represent the targeted protein are monitored by MRM for quantification.

A signature peptide should be chosen based on a specific region of interest in a target protein, or simply a unique sequence without homology to other non-target proteins. It may have a length of about seven to 25 amino acids and should not contain reactive amino acids (Met, His, Trp, Cys, etc.) or areas with known PTMs unless the quantification targets one or more of these PTMs. The sensitivity, stability, specificity, and LC property should be determined experimentally and, perhaps, with in-silico computational support (Anderson and Hunter 2006, Kamiie et al. 2008). It is considered good practice to select three to five signature peptides as different peptides from one protein can vary widely in their MS response. Recovery from sample processing and strong matrix effects may also affect the signal intensities of these peptides in biological matrices such as plasma and serum (Keshishian et al. 2007). Furthermore, monitoring multiple signature peptides provides a better chance to cover any potential truncation or modification of a target protein (Cao et al. 2010). There are exceptions to these rules in certain circumstances. For example, in the current pursuit of biologic drugs in the pharmaceutical industry, antibody drug candidates are being evaluated in every major pharmaceutical company’s pipelines. Rather than designing a unique signature peptide and developing a new assay for each antibody drug, Furlong et al. (2012) proposed the use of a single universal surrogate peptide to quantify a variety of monoclonal antibody (mAb) and Fc-fusion protein candidates in plasma samples obtained from animal species commonly used in pre-clinical drug development. This approach could speed up the assay development for quantification of different antibody drugs.

Peptide separation has been mostly achieved by RPLC, but in some cases, hydrophilic interaction liquid chromatography (HILIC) conditions are used to improve the retention and ionization efficiency of hydrophilic peptides. Two-dimensional LC using orthogonal separation mechanism, such as ion-exchange chromatography (IEC)-RPLC or RPLC-HILIC, has been utilized to fractionate and clean up samples, thus improving the sensitivity of the detection (Gilar et al. 2005, Keshishian et al. 2007, Liu et al. 2009).

Particle sizes and dimensions of columns that are suitable for peptide separation are generally the same as those used for small molecule applications. However, pore sizes of conventional LC packing materials, normally between 60–150 Å, are often too small for large molecules because these molecules cannot get access to the surface chemistries within these small pores. Thus, in recent years, packing materials with larger pore sizes (~300 Å) have been developed for separation of large molecules. Experiences gained from proteomics have also pushed the application of capillary and nano columns with inner diameter (I.D.) as narrow as 75 μm in peptide separation. The low flow rate (a typical flow rate for a 75 μm I.D. nano column is 200–300 nl/min) of these columns greatly improves ionization efficiency of ESI. For example, Onisko et al. (2007) quantified attomole level of Prion protein from brains of terminally ill Syrian hamsters using nano HPLC technology. A challenge for capillary and nano HPLC is the need to use sophisticated instrumentation for accurate and reproducible delivery of flow at the low μl to nl/min level. This has been a struggle for HPLC manufacturers, but some are slowly achieving acceptable performance.

Various factors may influence peptide ionization efficiency during mass spectrometer data acquisition, thus affecting the sensitivity, accuracy, and precision of a quantification assay. Alternations in matrix interferences,
mobile phase composition, or mass spectrometer operating parameters may change the abundance of specific ion types (i.e., different charge states or adduct ions). Ion-pairing agents, such as trifluoroacetic acid (TFA), are commonly used as mobile phase additives for peptide chromatographic separations due to their ability to improve peak shape by suppressing undesirable interactions between peptides and the stationary phase. In addition, they promote retention of peptides on RP-HPLC by ion-pairing with these peptides, which increases their hydrophobicity. The trade-off for using TFA is that TFA is known to suppress ESI signal intensity due to its ability to form gas-phase ion pairs with positively charged ions. Post-column infusion of propionic acid and isopropanol has been shown to alleviate ion suppression caused by TFA (Apffel et al. 1995). Alternatively, acetic acid (0.5%) or propionic acid (1%), when directly added to the TFA-containing mobile phase, can also effectively reduce ion suppression (Shou and Weng 2005). In many cases, formic acid has been used to replace TFA altogether (Lu et al. 2009, Li et al. 2011b, Liu et al. 2013a).

Software packages used for LC-MS/MS data acquisition and analysis of signature peptides are the same as those used in small molecule applications. Unlike a small compound, which typically generates a singly charged molecular ion, a peptide may produce M+2H+ and M+3H+ ions. The optimization of instrument parameters is necessary to obtain the best signals from different charge states. Signals from multiple MRM transitions of one or more signature peptides can be summed for enhanced signal intensities.

**Top-down workflow for the quantification of peptides and proteins**

**LC-MS/MS approach**

The MRM-based LC-MS/MS technology has also been used in top-down approach for peptide and protein quantification. In this approach, instead of a signature peptide, the molecule ion of an intact peptide or protein is selected as the parent ion for CID fragmentation, after which a suitable fragment ion is chosen for quantification in MRM mode of analysis. Ji et al. (2003b) reported using the approach to quantify rK5, a protein drug candidate with a molecular weight of 10,464 Dalton in monkey plasma.

This approach, however, has not been widely applied and, so far, there have been very few related publications. One serious inherent challenge for this approach is that the MRM mode of detection requires fragmentation of a parent ion to produce unique product ions in order to achieve desired specificity, whereas it is extremely difficult to fragment intact proteins of sizes over 10 kDa by CID. In fact, in another publication by Ji et al. on rK5 quantification, a protein similar to rK5 in structure but with a slightly reduced molecular weight was used as an internal standard and could not be fragmented. “Pseudo-MRM”, i.e., monitoring the parent ion to parent ion transition with very low collision energy, had to be conducted, which significantly sacrificed the specificity of the detection (Ji et al. 2003a). In our laboratory we could not fragment human ApoC3 proteins (9–10 kDa) when we used a similar approach (Jian et al. 2013). Newly developed fragmentation methods, such as electron-transfer dissociation (ETD) and electron-capture dissociation (ECD, mainly for Fourier transform mass spectrometers, or FTMS), which are more suitable for breaking down longer peptides or even entire proteins, may promote the application of this top-down quantification approach in the future (Horn et al. 2000, Coon et al. 2005).

For quantification of PEGylated proteins, Li et al. (2011b) developed a unique in-source fragmentation approach, in which a 20 kDa PEGylated protein was fragmented in ion source by CID to generate a signature peptide and was then monitored by MRM as a surrogate of the whole protein.

**Full scan LC-HRMS approach**

When a quantification target is a peptide or protein smaller than 10–15 kDa, it is a good candidate for top-down analysis by full scan LC-HRMS workflow. The sample preparation methods include protein precipitation, SPE, immuno-capture, etc., as discussed in the sample preparation section earlier. However, in contrast to the LC-MS/MS bottom-up workflow for larger proteins, enzyme digestion is not needed and no signature peptide is involved in the LC-HRMS workflow. Instead, an intact peptide or protein target is subject to analysis by a high-resolution mass spectrometer (HRMS) in full scan mode.

The LC component of the workflow is very similar to that of LC-MS/MS. HPLC columns with larger pore sizes (~300 Å) are recommended for easy access of analytes to the surface chemistries within the pores of the packing materials. For reverse-phase separation, packing materials with shorter hydrocarbon chains (e.g., C3, C4, and C8) may be chosen over commonly used C18 for increased hydrophobicity of large peptides and proteins. Micro or nano flow rates may be used to improve sensitivity of an LC-HRMS assay. However, our experiences suggest that samples processed by protein precipitation or SPE may be too crude for nano flow and may result in frequent clogging of columns or instruments.
For analyte detection, mass spectrometers with high resolving power and accuracy are essential to the workflow. In contrast to MRM analysis of signature peptides, full scans of intact peptides or proteins are performed using a mass spectrometer. A typical spectrum contains a charged envelope made of multiple peaks of different charge states. Each charge state has its own multiple isotopic peaks if the target peptide or protein is small enough and the mass spectrometer has sufficient resolution. Such spectra are considerably more complex than those of singly charged small compounds and doubly or triply charged small peptides. Attention should be given to the optimization of mobile phase components and ionization conditions in order to obtain the optimal peak quality. A data analysis strategy is needed to determine which peaks should be included in quantification. One or two charge states or isotopic peaks with high quality may be chosen if the peak intensities are high; it is also ideal to include as many peaks as possible to improve the final signal intensity and reproducibility of the quantification if the peak signals are weak.

Early experiments of LC-HRMS for peptide and protein quantification were performed on FTMS instruments. In 1997, Padley et al. demonstrated quantification of lysozyme, a 14 kDa protein using an electro-spray source followed by a linear ion trap coupled with an FTMS, although no internal standard was included and the linearity only covered 1.5 orders of magnitude (Padley et al. 1997). Gordon and Muddiman quantified cyclosporin A (CsA), a small cyclic peptide immunosuppressant using a similar type of FTMS instrument. Even though the peptide was small (~1.2 kDa) and all samples were pure standards in neat solutions, this study had the key elements of the full scan LC-HRMS approach, including a calibration curve with an analog protein (cyclosporin G, or CsG) as internal standard (Gordon and Muddiman 1999a). Using the same strategy, Gordon’s group also measured the concentrations of equine heart cytochrome c with bovine heart cytochrome c as an internal standard (Gordon and Muddiman 1999a). Using the same strategy, Gordon’s group also measured the concentrations of equine heart cytochrome c with bovine heart cytochrome c as an internal standard. Charge states of 7–14 were observed with each charge state isotopically resolved, and the four most abundant isotopic peaks from the dominating charge states of 7 and 8 were chosen for quantification (Gordon et al. 1999).

The rapid development of advanced and novel HRMS and the need to quantify biologics and biomarkers in the pharmaceutical industry have greatly accelerated LC-HRMS technologies. In the last several years, there has been great improvement in the resolution and mass accuracy of the quadrupole time-of-flight (Q-TOF) type of mass spectrometers. Such improvement and their fast scan speed and intrinsic high mass range in full scan mode have made Q-TOF a suitable choice for LC-HRMS analysis. Orbitrap, a recently developed novel high resolution mass analyzer, is another option. The latest generation of Orbitrap has expanded mass range and faster scan speed, which makes this instrument an important tool for top-down analysis (Zubarev and Makarov 2013).

Comparison studies have demonstrated that HRMS quantification of peptides and proteins can achieve a similar level of sensitivity to that of MRM while affording better specificity and higher efficiency (Dillen et al. 2012, Plumb et al. 2012). Ruan et al. performed a well-designed quantification assay for lysozyme in human plasma using an LTQ Orbitrap. In their work, the calibration curve was generated by spiking human lysozyme and chicken lysozyme (used as internal standard) in monkey plasma, and SPE was used for sample preparation. They reported that assay sensitivity was sufficient to measure the endogenous lysozyme in human plasma samples (Ruan et al. 2011). Meanwhile, Liu et al. reported the quantification of a recombinant monoclonal antibody in monkey serum using a Q-TOF instrument. After performing Protein A affinity purification, limited Lys-C digestion was performed to release the 47 kDa human Fab fragment, which was subject to LC-MS full scan analysis. The isotopically labeled antibody was used as an internal standard throughout the sample preparation steps to ensure accurate quantification. The actual quantification was based on the intensities of Fab peaks in the deconvoluted mass spectra. Concentrations of the antibody in multiple monkey serum samples determined by LC-MS were in good agreement with the values obtained from ELISA (Liu et al. 2011). Recently, Gucinski and Boyne developed and validated quantification assays for two insulin variants and human growth hormone using LTQ Orbitrap (Gucinski and Boyne 2012).

In a recent work in our laboratory, we developed an LC-HRMS method for the quantification of different glycoisomers of intact apolipoprotein C3 (ApoC3) in human plasma. ApoC3 exists mainly in three glycoisofoms, ApoC3-0, ApoC3-1, and ApoC3-2. The O-linked sugar moiety is bound to threonine in position 74 and consists of one residue of galactose, one residue of N-acetyl-galactosamine, as well as one and two residues of N-acetyleneuraminic acid (NeuNAc, known as sialic acid) for ApoC3-1 and ApoC3-2, respectively, while ApoC3-0 lacks the entire sugar chain (Figure 2). Changes in ApoC3 glycoisofom ratios have been observed in cases of a number of disorders, including obesity, kidney diseases, liver diseases and sepsis, and may thus provide important information for the diagnosis, prognosis, and evaluation of therapeutic responses of these conditions.
While available antibodies could not differentiate among these ApoC3 glycoisoforms, the sizes of these molecules (8.8–9.7 kDa) made them good candidates for LC-HRMS analysis. Human plasma samples were processed with SPE and then subjected to LC-HRMS analysis in full scan mode using a Q-TOF mass spectrometer. Isotopic peaks for each targeted glycoisoform at two charge states were extracted using a window of 50 mDa and then integrated into a chromatographic peak (Figure 3). The peak area ratios of ApoC3-1/ApoC3-0 and ApoC3-2/ApoC3-0 were calculated and evaluated for assay performance. Our results indicated that these ratios could be determined with excellent reproducibility in multiple subjects. This method was applied in a preliminary biomarker study of diabetes by analyzing plasma samples collected from normal, pre-diabetic, and diabetic subjects (Jian et al. 2013).

Internal standard strategy

A well-designed internal standard strategy is critical for the successful development of an LC-MS-based protein quantification assay. Variations may be introduced in any step in a workflow during sample processing, liquid chromatography, and ionization, but a good internal standard should track and compensate for such variations as much as possible. The pros and cons of different internal standard strategies have been extensively reviewed (Pritchard et al. 2011, Bronsema et al. 2012, Pailleux and Beaudry 2012) and will be briefly discussed in this section.

Stable isotope-labeled signature peptide

Similar to the case of small compound application, stable isotope-labeled signature peptides have been used as internal standards for LC-MS/MS-based bottom-up quantification of peptides and proteins. They are typically added at an early stage of sample processing, such as during enzymatic digestion. This approach is very straightforward and a stable isotope-labeled peptide is often easy to synthesize at low cost. Gerber et al. called these peptides “absolute quantification” (AQUA) peptides and used the strategy to quantify phosphorylated proteins in whole-cell lysates (Gerber et al. 2003). However, one of the major drawbacks of this type of internal standard is that it cannot correct for variations in enzymatic digestion or any other sample processing steps prior to the digestion; thus, it may cause serious bias in the quantification results (Brun et al. 2007, Li et al. 2012).

Stable isotope-labeled extended signature peptide

An alternative internal standard designed to cover the variations in enzymatic digestion step is the stable isotope-labeled extended signature peptide, which has extra amino acid sequences containing protease digestion sites flanking both sides of the signature peptide. These extended internal standards are added during the sample digestion step and go through the digestion process; therefore, variations in digestion efficiency are expected to be corrected to some extent. An example of such approach is the QCAT strategy developed by Beynon et al. QCAT is an artificial recombinant protein that has multiple signature peptides linked together. Stable isotope labeling may be incorporated into a QCAT protein during translation, and trypsin digestion will release these stable labeled signature peptides. Thus, a QCAT protein can be used as internal standards for quantification of multiple proteins (Beynon et al. 2005). However, the digestion sites in this type of internal standards are often more accessible than those in target proteins; therefore, they may not track the digestion kinetics of the target proteins. Comparison studies have shown
that a cleavable stable isotope-labeled peptide could not ideally correct for the digestion efficiency, thus offering little advantage over a regular stable isotope-labeled signature peptide internal standard (Barnidge et al. 2004, Brun et al. 2007).

**Stable isotope-labeled intact protein**

As discussed above, the primary drawback of stable labeled signature peptides and those with flanking sequences is that variation in sample processing is not...
tracked and corrected by such internal standards. The recovery of an intact target protein from enrichment steps may not be tracked because it is likely that the labeled signature peptide is not compatible with the enrichment methods, thus, has not been spiked in the sample at those stages. The variation in enzyme digestion efficiency may also not be tracked. Therefore, when a signature peptide is not detected in an MRM assay from a certain sample, it is often unclear whether it is due to truly low concentration, poor recovery from enrichment steps, or poor enzymatic digestion. For the same reason, when a signature peptide is detected and quantified, the results may not be reliable due to variation in recovery and digestion efficiency among study samples. Furthermore, the stable labeled signature peptide internal standard is not compatible with full scan LC-HRMS workflow in the top-down approach.

To address these issues, stable isotope-labeled intact proteins have been used as internal standards. These standards are usually added at the start of sample processing, or even during sample collection, thus providing better control over all the steps of the quantification workflow and improving the accuracy and precision of the bioanalytical assay. There are different ways to produce a stable labeled protein internal standard. For a small protein, the stable isotope-labeled version can be chemically synthesized (Jian et al. 2013). Brun et al. proposed a Protein Standard Absolute Quantification (PSAQ) approach, in which their intact protein internal standards were made by in vitro protein synthesis in the presence of stable labeled amino acids in a cell free system (Brun et al. 2007). Alternatively, such internal standards can be produced by stable isotope labeling with amino acids in cell culture (SILAC) (Ong et al. 2002). Heudhi et al. (2008) obtained their stable labeled intact antibody internal standard from antibody producing cells (SP2/0 Ag 14.0 cells) grown in medium supplied with stable labeled threonine. To facilitate the quantification of a large number of different antibody drug candidates in early discovery, Li et al. (2012) proposed a stable isotope-labeled intact antibody as a common internal standard for different antibody quantification assays.

**Protein analog**

Even though they are considered the best internal standards, stable isotope-labeled intact proteins are costly and take relatively long time to make. A protein analog may be used as an internal standard in the absence of an isotope-labeled protein. A protein analog internal standard should resemble the target protein as much as possible to better track the behavior of the target protein. For example, horse myoglobin has been used as internal standard for human myoglobin (Mayr et al. 2006). Interestingly, Yang et al. reported using bovine fetuin as an internal standard for the quantification of two totally different proteins, a recombinant human growth hormone somatropin and a monoclonal antibody. In their LC-MS/MS analysis, bovine fetuin peptides were selected based on the proximity of their retention time to that of the target signature peptides, similar recoveries, and absence of interference from target proteins or plasma constituents (Yang et al. 2007).

**Calibration strategy**

A calibration curve is necessary for any absolute quantification of peptides and proteins by LC-MS. The criteria and considerations for calibrator selection are very similar to those employed for the internal standard. An intact protein should be chosen over a signature peptide or an extended signature peptide if possible. Berna et al. reported that there was a ~15% bias when they quantified Myosin Light Chain 1 protein in rat serum using immunocapture followed by LC-MS/MS analysis with a calibration curve made from a signature peptide. They found that the negative bias was caused by recovery loss due to immunocapture recovery, digestion efficiency, and ionization suppression (Berna et al. 2007). Similarly, Cao et al. observed significant underestimation when two signature peptides were used independently for quantification of a same protein (Cao et al. 2010).

A calibration curve should be prepared in the blank matrix (i.e., the target biological sample devoid of the analytes) whenever possible. If a matrix has endogenous analyte, a surrogate matrix with similar property may be used. For example, a calibration curve may be prepared in monkey plasma if study samples are human plasma (Ruan et al. 2011). A similar approach was used by Cao et al. (2010) when they quantified carbonyl reductases CBR1 and CBR3 in human liver.

Calibration standards and quality control (QC) samples may exhibit different sample extraction recovery rates, especially during enzymatic digestion, from incurred samples due to different matrix contents, thus leading to biased quantification results. Liu et al. proposed a “standard-addition test” to verify the ruggedness of a protein LC-MS/MS assay for incurred samples. After the initial measurement of an incurred sample, the sample was mixed with a standard sample containing a
known amount of analyte in one-to-one ratio. The resulting sample was measured again and the result was compared with the theoretical concentration (original + added) to reveal the accuracy (Liu et al. 2013a).

For assay validation, similar experiments as those for the LC-MS/MS methods for small molecules, such as accuracy, precision, sensitivity, selectivity, stability, and so on, should be evaluated. To address the unique aspects of protein quantification, procedures such as high-abundant protein depletion, immuno-capture enrichment, and enzymatic digestion should be evaluated for their efficiency, depending on the specific workflow.

**Common issues in the quantification of peptides and proteins using LC-MS**

Compared with small compound applications, the quantification of peptides and proteins has similar challenges, such as stability, adsorption, and solubility, among others. It also has its unique problems, such as specific protein binding and PTMs. Given the nature of peptides and proteins, methods to address these challenges and issues are quite different from those of small compounds.

**Stability**

Stability of peptides and proteins is a major concern in large molecule quantification. Cao et al. (2010) showed significant instability of multiple candidate signature peptides under various storage conditions over time, and that the extent of instability varied markedly. Many different factors affect peptide and protein stability. Peptides and proteins, especially at low concentrations are susceptible to proteolysis caused by naturally occurring protease in all organisms. To address this issue, protease inhibitors may be added in one or more steps of sample preparation process (Olivieri et al. 2001, Rai et al. 2005). Protease inhibitors are commercially available in individual or cocktail format and are easy to use. All peptide and protein samples should be handled on ice or at 4°C because most proteases function optimally at room temperature to 37°C.

Some amino acids are prone to oxidation, including methionine and cysteine, and to a lesser extent, tryptophan and histidine. Hence, prolonged exposure to atmospheric oxygen should be minimized. Given that cysteine tends to form intra- and inter-molecule disulfide bonds, a common practice is to reduce disulfide bonds completely by dithiothreitol (DTT) or 2-mercaptoethanol treatment, followed by alkylation with iodoacetamide or iodoacetic acid. Such reduction and alkylation steps should eliminate the variability of disulfide formation (Herbert et al. 2001, Kirsch et al. 2007).

Another important factor for stability is pH. Fragmentation occurring at the C-terminus of an Asp residue (peptide bond Asp-Xaa, where Xaa is any residue) is one of the most frequent degradation pathways of mAbs under mildly acidic conditions (Vlasak and Ionescu 2011). Deamidation of asparagine and glutamine also happens at acidic conditions (Robinson and Robinson 2004). Therefore, peptide and protein samples should be processed and stored in a buffered solution at near neutral pH as much as possible.

Finally, protein samples should be handled gently. For example, repetitive freezing/thawing, vigorous vortexing, and foaming can cause damage to proteins, and should thus be avoided during sample preparation or storage.

**Adsorption**

The amphiphatic nature of peptides and proteins makes them readily adsorb to most surfaces, which can lead to inaccurate quantification results (Wilson et al. 2010, Goebel-Stengel et al. 2011). In general, positive charged peptides tend to adsorb to a glass surface (with negative charge), while neutral peptides tend to adsorb to polypropylene due to hydrophobic interaction. However, adsorption is not easy to predict and only manifests itself during experiments. There are many different ways to prevent peptide and protein adsorption, but each method has to be tested empirically. Adsorption is more severe when peptides and proteins are at very low concentrations in matrix-free aqueous solutions, so it is good practice to make high concentration stocks and spike them directly into plasma or serum samples. Adding displacement agents, such as a structural analog (Fouda et al. 1991) or a protein rich solution (e.g., bovine serum albumin) may limit adsorption (Schwartz et al. 1997, Goebel-Stengel et al. 2011). Using organic-aqueous solution (at least 20% organic) is another option (Hyenstrand et al. 2001, van Midwoud et al. 2007). It is convenient to use specially-treated tubes and plates that are commercially available for low peptide and protein adsorption, but attention should still be paid to each individual product. For example, contrary to the common belief, Goebel-Stengel et al. (2011) showed that siliconization did not improve but significantly decreased the recovery of most of their tested peptides.
Specific protein binding

A potential problem of the LC-MS/MS assay for peptide and protein quantification is the specific protein binding in test samples. Specific protein binding can be illustrated by variations in quantification of growth hormone (GH) between different immunoassays. Growth hormone-binding protein (GHBP) has a high affinity to growth hormone, and in human serum and plasma samples, up to 50% of growth hormone forms complex with GHBP. Therefore epitopes might not be accessible for certain GH antibodies due to steric hindrance in an immunoassay, which leads to underestimation of GH concentrations (Bidlingmaier 2008).

With the rapid development of biologic drugs, one situation that deserves attention is the formation of anti-drug antibodies (ADAs). Biologics have been shown to be immunogenic in animals in pre-clinical tests and in human patients (Warnke et al. 2012). Aside from the problems they may cause to biologics, such as loss of efficacy, ADAs post issues in quantification of these biologics; ADAs and a drug have to be separated in the sample preparation process in order to avoid any misrepresentation of the total drug concentration in a specimen. Ji et al. (2007) described a sample preparation method, in which monkey plasma samples were denatured with 8 M guanidine hydrochloride to disassociate a small protein drug from its ADAs before solid phase extraction. Meanwhile, ADAs themselves may need to be quantified in order to assess immunogenic potential of a biologic drug during the development phase. Many available technologies measuring ADAs, such as ELISA, cell-based assay and surface plasmon resonance, are subject to interference by high circulating concentrations of the protein therapeutics that bind to ADAs. Neubert et al. (2008) took advantage of the affinity between a PEGylated human growth hormone analog (hGHA) and its ADAs by capturing the hGHA-ADAs complex with protein G magnetic beads first, and then quantifying the hGHA, whose concentration was proportional to those of total ADAs.

PTMs

PTMs are covalent processing events that change the properties of a protein through proteolytic cleavage or through the addition of a modifying group to one or more amino acids. Even though PTMs have been observed in both prokaryotes and eukaryotes (Huq and Wei 2006), they are more prevalent in eukaryotes because of the presence of organelles such as endoplasmic reticulum (ER) and Golgi apparatus. Some common forms of PTMs are phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, lipidation, and proteolysis.

Due to the fact that they may potentially happen to any peptide or protein, PTMs have to be considered when a workflow is designed for quantification. In general, sequences containing potential PTMs should be excluded when signature peptides are selected. N-linked glycosylation can be removed by treatment of PNGase F enzyme when quantification targets are glycoproteins. Meanwhile, proteins with PTMs of particular interest may be targets for quantification. The quantification of proteins with particular PTMs presents additional challenges in the development of an LC-MS/MS method, because the assay must target specific sequences containing the PTMs, but these sequences may not be amenable to LC-MS/MS analysis. Some peptides may be difficult to detect due to low ionization efficiency, large peptides may fall outside of the detection range of a mass analyzer, and short hydrophilic peptides may not be retained on an HPLC column. Aside from trypsin, other proteases or a combination of proteases may be needed to generate suitable peptides. Liu et al. (2013b) described a general workflow for quantification of PTMs and protein isoforms as well as discussed assay optimization using phosphorylated and citrullinated peptides as examples.

Conclusions and future perspectives

As we have discussed, despite the significant challenges associated with LC-MS-based assays for peptide and protein quantification, these techniques have shown promise as emerging complementary platforms for large molecule bioanalysis. The MRM-based LC-MS/MS bottom-up assays are a natural extension of proteomics research. First, a peptide or protein of interest is identified by proteomic discovery, and then the MRM assay is used to quantify its concentration. Knowledge obtained from proteomics over the last two decades has greatly accelerated the development of such assays. The triple quadrupole mass spectrometers, which have been used in small compound quantification for many years and have become sophisticated over the years, are perfect for LC-MS/MS-based peptide and protein quantification. Furthermore, in the past several years SRMAtlas, a knowledge database has been developed (www.srmatlas.org) based on high-quality measurements of natural and synthetic peptides using triple quadrupole mass spectrometers. Users
are able to query MRM transitions from a variety of data sources, and this setup has greatly facilitated the assay development. Currently, triple quadrupole mass spectrometer data covers 97% of yeast and 99% of Mycobacterium tuberculosis proteome, and SRMAtlas will expand the coverage to 99% of human (~157 K peptides, ~20,200 proteins) and 80% of mouse (~60 K peptides, ~12,000 proteins) proteome, as well as ~7 K peptides/~3000 proteins of empirical human glycoproteome (Picotti et al. 2008, 2013, Hüttenhain et al. 2013, Schubert et al. 2013). With the multiplexing capacity of analyzing large number of targets in a single run, LC-MS/MS-based assays may eventually be chosen over ELISA assays, which require antibodies for each of the analytes to be identified or developed.

Meanwhile, the LC-HRMS-based top-down peptide and protein quantification assays are still in their early stages even though they have shown great potential. One of the issues in this workflow is that it is more suitable for peptides and proteins with small sizes. For quantification of proteins with high molecular weights, a mass spectrometer with high resolving power is required. The higher the resolving power of a mass spectrometer the better its selectivity; therefore the better sensitivity as a narrower extracted ion chromatogram (EIC) window can be utilized to remove the background interference. Figures 4 and 5 illustrate the effect of resolving power on the separation of isotopic peaks and the effect of EIC window on assay selectivity and sensitivity, respectively. Based on their model, Ruan et al. (2011) calculated that 600,000 practical resolving power was needed to quantify an intact antibody of about 150 kDa. Mass spectrometers that can handle higher mass ranges are also in demand because some proteins may not carry sufficient charges to fit their m/z values into the narrow detection ranges of current mass spectrometers. Aside from the requirement for high end mass spectrometers, more user-friendly quantification software is needed to handle multiple charge states and isotopic peaks resulting from intact peptide and protein analysis. Currently in our lab, we must manually sum the chosen charge states and isotopic peaks to obtain final quantification results. Software packages that can perform automatic summing, accurate deconvolution, and any other user-friendly features are expected to be developed by commercial suppliers.

One of the major bottlenecks for both LC-MS/MS- and LC-HRMS-based workflows is their relatively low...
throughput in sample preparation; major sample preparation steps must be automated to improve the throughput. Routine automation of some sample preparation procedures for small compounds, such as SPE and protein precipitation, can be easily transferred to peptide and small protein applications. The automation of immunocapture and enzymatic digestion procedures has been attempted and can be adapted for peptide and protein quantification workflow (Chelius et al. 2008, Berna and Ackermann 2009, Richardson et al. 2011). Considerable efforts have also been made towards the automation of immuno-capture techniques by commercial suppliers. Examples include the automation of Dynobeads procedure by Life Technologies (www.lifetechnologies.com) and MSIA technology by Thermo Fisher Scientific (www.thermoscientific.com). These individual technologies must be battle-tested extensively, standardized, and ultimately linked together to improve the overall efficiency of the LC-MS-based quantification workflows.

In summary, we believe that the absolute quantification of peptides and proteins by LC-MS has its unique advantages and is complementary to traditional quantification methods, such as immunoassays, especially when antibodies are not readily available or cross-reactivity exists in available antibodies. Technological advancements should be able to address many of the current technical issues and bottlenecks to expand its suitable applications. Growing needs from biomedical research as well as the pharmaceutical and diagnostics industries will eventually push the LC-MS-based peptide and protein quantification methodologies into the mainstream of bioanalysis.

Acknowledgments: We wish to thank Dr. Jonathan Crowther, Dr. Anna Luczak, and Dr. Steve Alexander as well as Nathan Connolly, for their critical reading and helpful discussions of this review.

Received September 4, 2013; accepted December 29, 2013; previously published online February 13, 2014

References


Horn, D. M.; Ge, Y.; McLafferty, F. W. Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins. Anal. Chem. 2000, 72, 4778–4784.


Ji, Q.; Rodila R.; El-Shourbagy, T. A. A sample preparation process for LC-MS/MS analysis of total protein drug concentrations in monkey plasma samples with antibody. J. Chromatogr. B. 2007, 847, 133–141.


Li, H.; Rose, M. J.; Holder, J. R.; Wright, M.; Miranda, L. P.; James, C. A. Direct quantitative analysis of a 20 kDa PEGylated human calcitonin gene peptide antagonist in cynomolgus monkey


Stanley (Weihua) Zhang received his PhD in Molecular Biology and Biochemistry from Rutgers, The State University of New Jersey, in 1999. His professional career covers several different areas, including proteomics-based biomarker discovery, peptide and protein characterization, as well as small and large molecule quantification. He has been with Johnson & Johnson for over 8 years, and currently focuses on assay development of companion diagnostics.

Wenying Jian received her PhD in Pharmacology from the University of Pennsylvania in 2005. Her current work at Janssen Research and Development involves biochemical and metabolic development and validation for drug candidates, metabolites, and biomarkers in support of drug development in preclinical and clinical stages. Her research expertise and interest center on the application of advanced LC-MS methodologies in the detection, identification, and quantification of endogenous small and large molecules, including biomarkers, drugs, and their metabolites in complicated biological matrices.