Proteomics: from hypothesis to quantitative assay on a single platform. Guidelines for developing MRM assays using ion trap mass spectrometers

Bomie Han and Richard E. Higgs

Advance Access publication date 25 June 2008

Abstract

High-throughput HPLC-mass spectrometry (HPLC-MS) is routinely used to profile biological samples for potential protein markers of disease, drug efficacy and toxicity. The discovery technology has advanced to the point where translating hypotheses from proteomic profiling studies into clinical use is the bottleneck to realizing the full potential of these approaches. The first step in this translation is the development and analytical validation of a higher throughput assay with improved sensitivity and selectivity relative to typical profiling assays. Multiple reaction monitoring (MRM) assays are an attractive approach for this stage of biomarker development given their improved sensitivity and specificity, the speed at which the assays can be developed and the quantitative nature of the assay. While the profiling assays are performed with ion trap mass spectrometers, MRM assays are traditionally developed in quadrupole-based mass spectrometers. Development of MRM assays from the same instrument used in the profiling analysis enables a seamless and rapid transition from hypothesis generation to validation. This report provides guidelines for rapidly developing an MRM assay using the same mass spectrometry platform used for profiling experiments (typically ion traps) and reviews methodological and analytical validation considerations. The analytical validation guidelines presented are drawn from existing practices on immunological assays and are applicable to any mass spectrometry platform technology.

Keywords: multiple reaction monitoring; assay validation; spiked recovery; HPLC-MS; quantification; proteomics; LTQ

INTRODUCTION

HPLC-MS proteome profiling of serum proteins has the potential to identify new biomarkers for many different applications including disease progression, efficacy and toxicity of therapeutic intervention [1]. However, the technological advancements in HPLC-MS instrumentation and data processing have largely not fully delivered on the promise of demonstrable clinical impact on treatment or drug development. This failure can largely be attributed to the masking effect of the abundant serum proteins on less abundant proteins and false positive findings due to underpowered study designs coupled with non-robust data analysis approaches [2]. Improvements in sample preparation/fractionation techniques to enrich for lower abundance proteins as well as improved study design and data analysis should increase the chances of identifying clinically useful markers directly from serum. Alternatively, performing profiling experiments on samples more proximal to the biology of interest (e.g. tissue, urine, cerebrospinal fluid, etc.) may increase the chances of identifying protein changes that would otherwise be very difficult to...
find in serum directly [2]. Irrespective of the source of the initial findings or hypotheses, the next step in the biomarker development process is often a high-throughput, sensitive and selective serum assay. The multiple reaction monitoring (MRM) assay is an attractive option at this step of the biomarker development process given its sensitivity, selectivity, the speed at which the assay can be developed and the quantitative nature of the assay. Many analytes in the sub-to-low microgram per millilitre range in plasma are simultaneously quantifiable using an MRM approach with minimal sample preparation [3]. Protein profiling studies are often carried out using ion trap mass spectrometers (e.g. LTQ from Thermo Scientific, Waltham, MA) and if the MRM assay is developed on this same platform, the speed of assay development can be extremely rapid by using the peptide fragmentation data generated in the profiling study to inform the MRM assay development. In cases where plasma protein levels are below the sensitivity of the MRM, single antibody enrichment approaches are possible with development times typically much faster than two-antibody ELISA assay development [4–6]. The purpose of developing the MRM assay is to (i) confirm the initial finding from the protein profiling study using a more selective assay, possibly in the same study samples used for profiling and (ii) begin the biological validation analyses by applying a more selective and higher throughput assay on new study samples. The rationale for repeating measurements of peptides on the initial study samples using an MRM method is that quantification of any peptide signal from a profiling study relies solely on the parent ion mass and is susceptible to interference by unrelated peptides while the MRM method uses product ion intensity and has higher selectivity due to double filtering of parent and product ion mass. Prior to using the MRM assay for biological validation on new study samples, it is important to first perform an analytical validation of the MRM assay in a manner consistent with how the assay will be used (‘fit for purpose validation’) [7]. Since many hypotheses will fail during the first few steps of confirmation, it is desirable to have a mechanism to evaluate multiple hypotheses quickly. In other words, it is advantageous to have methods for relative comparisons of protein concentration between control and test samples without substantial investment in generating standard reagents before attempting to set up a fully quantitative assay. In the first part of this article, we describe methods to quickly build MRM assays and to test initial hypotheses using an ion trap mass spectrometer. In a later part, we focus on building a quantitative and high-throughput MRM assay. We used the LTQ from Thermo Scientific coupled in line with an HPLC system for all experiments. Although some discussions are specific to ion trap mass spectrometers, many are applicable to a broad range of mass spectrometry technologies. We will try to delineate between instrument-specific discussions and guidelines on MRM assays in general. Analytical method validation guidelines for quantitative MRM assays are presented based on adapting the best practices for immunological assay validation literature and are applicable to any mass spectrometry platform.

GUIDELINES

In this section, we enumerate a set of guidelines we have found useful in rapidly developing MRM assays to confirm initial findings from a proteome profiling study, configure the assay to be quantitative and validate the assay using a spiked recovery experimental design to estimate the limits of quantification and general performance of the assay (Figure 1).

Figure 1: Overall flow scheme for MRM method development and analytical validation.
Peptide sequence selection and considerations for trypsin digestion

Since most profiling studies are performed using trypsin as a proteolytic enzyme, we will focus our discussion on tryptic peptides in this section. Not all tryptic peptides can be detected in a reproducible manner due to the presence of closely eluting peptides of similar mass and low confidence in peptide identification due to a poor MS/MS fragmentation pattern. Because targeted MRM methods often have 10 times or more sensitivity than proteomic profiling methods, peptides are more likely to be detected reproducibly in an MRM. However, not all peptides identified from a profiling study will be suitable for quantitative purposes. The following guidelines are applicable to all types of mass spectrometer and may be considered when selecting peptides for MRM assays:

- Generate a list of tryptic peptides from the amino acid sequence of the protein (e.g. using GPMAW software from Lighthouse data, Odense M, Denmark) [8]. When profiling data are available, identify which of the tryptic peptides have been detected from the profiling studies. Peptides containing approximately 7–30 residues are generally optimal for MRM analyses.
- Avoid peptides that are too hydrophilic or too hydrophobic. This criterion usually goes hand-in-hand with the length of the peptide, but many peptides can behave differently between the two criteria and, therefore, it is worth checking the hydrophobicity of the peptide independently of the peptide length [9]. Tryptic peptides are usually separated through a reverse-phased HPLC column before they are subjected to mass spectrometric analysis. After loading of the sample onto the column using low or no organic solvent, it is desirable to wash the column with the loading solvent to remove salts and other additives such as urea and divert them away from the mass spectrometer before peptides are eluted with higher organic solvent. Peptides that are too hydrophilic and do not bind to the column tightly will be lost during this initial washing step. GPMAW or other peptide analysis software provide information on the concentration of acetonitrile for elution of the peptide from a typical C18 column (e.g. HPLC index value) [8]. Peptides with HPLC indices <5 will fail to bind to the column tightly enough under most experimental conditions. Conversely, peptides with HPLC indices >30 often give unsatisfactory results due to interference with other hydrophobic polymeric ‘chemical noise’.
- Avoid peptides that contain potential modification sites. Some of the common modifications to consider when selecting peptides for an MRM are discussed subsequently: Met residues can be oxidized. Asn residues in the glycosylation consensus sequence (N-X-S/T) are often glycosylated, and Asn residues followed by a small residue can be deamidated in the following order: Gly > Ala > Ser or Thr [10]. The molecular weight of the deamidated form is 1 mass unit higher than the intact peptide. Once deamidated, it can adopt one of the two alternative configurations that often separate during HPLC. Therefore, peptides containing Asn followed by small residues often show three peaks on HPLC, two of which have identical mass that is one unit higher than the original peptide. Such peptides have been successfully used for quantification but careful monitoring of all forms is required [11]. N-terminal Gln or Glu residues can undergo cyclization to form pyroglutamic acid. Formation of the pyroglutamic ring is thermodynamically favoured and often goes to completion, especially at non-neutral pH. However, both intact and modified peptides need to be monitored before the peptide can be used for quantification.

- Peptides that contain internal Arg or Lys residues are incomplete digestion products and should be avoided. However, when Arg or Lys is followed by Pro, the site is consistently resistant to trypsin digestion and can be considered reproducible. Although not as resistant as the K/R-P site, when Arg or Lys is followed by multiple acidic residues, the site is often quite resistant to trypsin digestion. Peptides with another trypsin cleavage site within two or three residues of the N-terminal portion of the given tryptic peptide can often result in incomplete digestion at the desired cleavage site and are not desirable. If used, both cleavage forms of the peptide need to be quantified and added together to measure the total peptide concentration because the relative concentration of each form may not be highly reproducible.

- Peptides with a single Pro often result in MS/MS fragmentation patterns dominated by fragmentation at the C-terminus of the Pro residue. Since these peptides often do not produce other cleavage products with significant intensity, identification of the peptide depends heavily on a single MS/MS transition and, therefore, confidence for
identification of the peptide by MRM is low. Even when multiple minor fragment ions from other positions can provide enough sequence information to assign peptide identification with high confidence from the full MS/MS scan, fragments with signal intensity <10% of the most intense ion are not useful in most cases for quantification with an MRM. If such a peptide is chosen for development of a quantitative assay, a full MS/MS spectrum of the parent ion needs to be examined at each step during assay development and optimization to ensure that no unrelated product ions are present. In addition, it is highly recommended that the retention time is compared to a synthetic peptide at the early stages of evaluation to increase the confidence of the identification.

Peptides with His residues can take up additional charge for every His present. For example, a tryptic peptide with a single His is likely to have three positive charges under acidic conditions instead of the typical two positive charges for most tryptic peptides. However, when the His residue is close to either the N- or C-terminus, suppression of the charge on the His residue is often observed due to charge repulsion. In this case, the same peptide can take either two or three positive charges. Therefore, both charge states need to be examined to determine which state produces better MS/MS fragmentation patterns and higher overall signal intensity.

Cys residues need to be reduced and alkylated. Commonly used methods utilize reduction by DTT followed by alkylation with iodoacetamide. Alkylated N-terminal Cys residues can undergo cyclization in a manner similar to pyroglutamic acid formation from N-terminal glutamine. Reduction by triethylphosphine followed by alkylation with 2-iodoethanol can avoid this problem [12]. An additional advantage of the second method is that the reducing and alkylation agents can be removed by evaporation. In both cases, modification of the Cys residue should be included in the calculation of the peptide mass.

A common contaminant of trypsin preparation is chymotrypsin. Trypsin should be treated with tosyl phenylalanyl chloromethyl ketone (TPCK) to inactivate chymotrypsin. TPCK-treated trypsin can acquire chymotrypsin-like activity due to partial self-digestion, resulting in further digestion of the tryptic peptides at non-canonical sites (or over-digestion) and lowering the yield of the tryptic peptides. To minimize over-digestion, it is recommended that the amount of trypsin is titrated based on the total amount of protein in the sample. Ratio of trypsin to total protein of 1:100 (w/w) is recommended. Alternatively, Lys-modified trypsin (e.g. trypsin-gold from Promega, Madison, WI) can be used, which is resistant to self-digestion. A comparison of trypsin from different vendors indicates that Lys-modified trypsin does not produce over-digestion even when an excess of trypsin is used.

Protein databases should be searched to check for known polymorphisms of the peptide and to determine if the peptide sequence is unique to the protein of interest. Although peptides with more than seven residues are unlikely to be found in unrelated proteins by chance, proteins in the same family or isoforms may contain stretches of highly homologous sequences. Depending on the application it may be desirable to select peptides shared by different species so that the same assay can be used for different clinical and pre-clinical applications. Conversely, selecting peptides with sequences different from the species used as a dilution matrix for calibration standards or control samples can also be important.

The peptides resulting from these selection criteria are likely to have between 7 and 30 amino acid residues, HPLC index value between 5 and 30, two positive charges or a single dominant charge state under acidic conditions, devoid of potential modification sites, no internal trypsin cleavage sites and unique to the protein(s) of interest.

**Confirmation of MS/MS identification and building preliminary MRM method**

One advantage of having profiling proteomics data is that full MS/MS spectra are already available for prioritization based on their fragmentation spectrum. If purified protein is available, MS/MS spectra for individual tryptic peptides can be also obtained by trypsin digestion. Confirmation of the peptide identification can be obtained by comparison of the full MS/MS spectrum to a list of theoretical m/z values of all product ions (Figures 2 and 3E). If neither profiling proteomics data nor purified protein is available, preparation of synthetic peptides may be necessary to obtain full MS/MS spectra of the peptides to evaluate.

Peptides satisfying the criteria listed above with confirmed MS/MS identification can be prioritized based on the presence of three or four major MS/MS fragment ions since it may require quantification of
two or three high-quality MS/MS transitions to build a successful MRM assay. Up to five MS/MS transitions are often selected for evaluation during these initial stages of assay development. Based on this evaluation, three transitions that give high signal and clean background are selected for further development. Although the signal intensity of each transition is an intrinsic property of the peptide, background signal or interference is a property of the sample which is subject to change for biological effects from disease or treatment, sample preparation or HPLC conditions. Therefore, it may be necessary to revisit the selection of transitions once all assay conditions are optimized and additional biological samples have been analysed.

The use of the ion trap instrument has a major advantage for this initial confirmation stage of assay development because transition from the profiling study to the more quantitative MRM assay is seamless. Nearly identical instrument parameters can be used to set up the MRM that were used to collect full MS/MS spectra from the initial proteome profiling study. When a different instrument is used for MRM development, the fragmentation pattern of the parent ion does not always translate to a new instrument requiring an optimization of instrument parameters. While synthesis of the target peptide alone can take a couple of weeks before optimization of the method can be initiated on a different instrument, one can have a fully functional MRM method on the same instrument on the same day the profiling data become available.

While the LTQ instrument suffers from the requirement of having a wider parent ion selection
Figure 3: Example demonstrating how the full MS/MS and three transitions are used to identify the 2+ EGFR-5 peptide NLQEILIGAVR in an MRM method. Total ion chromatogram from summing the three transitions (A). Total ion chromatogram from the full MS/MS scan (B). Large, unrelated peak from summed MRM transitions (c) and corresponding MRM spectra (C). Desired peptide peak (d, e) with corresponding MRM transition spectra (D) and full MS/MS spectrum (E).
window compared to other instruments, we take advantage of this wide selection window by examining higher mass species from natural isotopes of the peptide. Instead of placing the centre of the mass selection window on the monoisotopic mass of the parent ion, we use the average mass as the centre of the selection window and, thus the centre of the mass selection window is biased to allow molecules with higher isotopes. When multiple isotopic masses are selected for the parent ion, the abundance of species with different isotopes in the product ions will reflect the natural abundance of these isotopic masses. Under this condition, a clean MRM shows a natural isotopic distribution within the product ion \( m/z \) window (Figure 3D). Theoretical isotopic distributions can be calculated from natural abundances of the common isotopes \((^{2}H\text{ and }^{13}C)\) and peptide composition. On the other hand, when a parent ion of a different species with a similar \( m/z \) value is captured and produces a product ion with a similar \( m/z \) value, the product ion is unlikely to produce the same isotopic distribution pattern (Figure 3C). Such chemical noise often shows no isotopic distribution at all because only a single isotopic parent mass was isolated. The isotopic distribution patterns of product ions in the MRM is a powerful tool for both correct identification of the peptide and for monitoring the degree of contamination of the MRM signal by unwanted peptides or small molecules. This information can guide development of sample preparation or HPLC methods for better separation of the target peptide from interferences. We often observe changes in the distribution pattern of the \( m/z \) values within the product ion window towards more ideal distribution pattern as the HPLC method is improved. While the isotopic distribution pattern is a powerful tool, it is also important to remember that the observed pattern may deviate significantly from the theoretical pattern when the overall signal intensity is low. The following guidelines may be considered when building preliminary MRM methods following initial profiling studies:

Select three to five major MS/MS fragments. Relative signal intensity of different fragments (often shown as percentage of the most intense fragment) is determined by the intrinsic probability of fragmentation at each position under the given instrument parameters. Because of this probabilistic nature of signal intensity, higher intensity signals are more stable. For the same reason, higher intensity signals also give better reproducibility, especially at lower concentrations of the peptide.

Higher \( m/z \) fragments are more informative than lower \( m/z \) fragments in terms of probability of false identification. It is less likely to obtain the same \( m/z \) value from an unrelated combination of amino acid residues when the \( m/z \) value is higher than when it is lower. Therefore, the \( y9+ \) ion is preferred over the \( y2+ \) ion in the example shown in Figure 2.

MS/MS ions that did not result from fragmentation of the peptide backbone should be avoided if possible. These ions include the unmodified parent ion, water loss from the parent ion and side chain fragmentations, of which water loss is most commonly observed. These ions do not provide any more sequence information than the parent ion itself and, therefore, lack the specificity feature we are seeking in the MRM assay.

The same product ion can often take multiple charge states. Choose only one charge state to monitor from each product ion.

Once the desired product ions are selected, build an MRM method by putting together parent mass, parent mass selection window and mass selection windows for three to five product ions. For the parent ion, use the \( m/z \) value for the average mass (as opposed to monoisotopic mass) for centre of the parent ion selection window. For the mass selection window of the parent ion, use 3 mass units for the width of the selection window when the parent ion \( m/z \) value is <700 and 4 mass units when it is >700. For product ions, mass selection windows are from −1 mass unit relative to the monoisotopic mass to +2 \((m/z < 700)\) or +3 \((m/z > 700)\) mass units. In this manner, the lowest mass ion will be exactly 1 mass unit away from the left edge of the window, allowing easy identification of the correct \( m/z \) value for the target ion from the spectral data.

**Preliminary LC/MS/MS assay method development**

Depending on the nature and intensity of the background, stronger MS/MS transitions do not necessarily result in low background signals. Signal intensity as well as background noise needs to be evaluated from actual experimental samples under HPLC conditions to be employed for the MRM assay. Guidelines for building the initial LC/MS/MS method are described in this section. This preliminary method will be further optimized once it is
determined that development of a high-throughput quantitative assay is worth pursuing.

HPLC conditions for a protein profiling assay are usually optimized for detection of as many peptides as possible. Typically, profiling studies employ a lengthy chromatographic separation of the sample, often exceeding 2 h per run. For initial MRM evaluation, a much shorter chromatographic separation (e.g. 19 min) can be used (Table 1). Including 1 min of sample injection time, this prototypical method allows testing three samples an hour. Column eluate is diverted away from the mass spectrometer during the first 2 min and the last 3 min.

This method is not designed to deliver optimal performance of the HPLC column, but rather, is designed to give reasonable resolution for as many different peptides as possible with a relatively short run time without optimization for individual peptides. The HPLC is run under steady state but not under equilibrium, and the retention time of most peptides does not match with the HPLC index value. This is why up to 45% acetonitrile is used during resolution of the peptides (first 16 min) while the HPLC index of most of the selected peptides is <30. A blank sample at the beginning of a series of samples should be used as well as whenever a batch is interrupted and restarted. Shortening the column equilibration step with the loading solvent (17.4–19 min) can result in destabilization of peak shape and retention time, particularly for less hydrophobic peptides. Although many commonly reported methods use 5% acetonitrile during sample loading, we use no acetonitrile during sample loading and for the first 1.5 min to make sure that hydrophilic peptides can be captured by the column.

Many investigators stop trypsin digestion by acidifying the solution to prevent over-digestion. However, many amphiphilic peptides show reduced solubility under acidic conditions, resulting in the loss of signal during storage at 4°C. A 100 mM ammonium bicarbonate solution for trypsin digestion, which produces a pH of about 7.5, may be used to load samples directly onto the column without adding acid when Lys-modified trypsin is used.

The same study samples that were used in the profiling study can be used to identify peptides of interest and to evaluate different transitions for refinement of the MRM method. We collect full MS/MS spectra and MRM data for the same parent ion in the same run. It is important to change the number of microscans to one in all targeted methods to reduce cycle time whereas more microscans are used during profiling studies. Due to the increased acquisition time for full MS/MS spectra, up to about five different peptides can be profiled in this manner with an LTQ. Examination of the major peaks in the total ion chromatogram (TIC) of the MRM data (peaks c and d in Figure 3A with corresponding mass spectra in Figure 3C and D) can quickly locate the target peptide. While peak c was mostly composed of a product ion whose m/z value was close enough to be captured within the selection window (Figure 3C m/z = 515.9 compared to 515.4 for the correct peptide), peak d contained all three major fragmentations with correct isotopic distribution. Since the full MS/MS data and the MRM data were obtained in the same run, the exact position of the target peptide in the TIC of the full MS/MS is known from the MRM data. The full MS/MS spectrum with the same retention time (peak e in Figure 3B and corresponding spectrum in Figure 3E) matched the MS/MS spectrum of the peptide from the profiling study (Figure 2) providing confirmation that the MRM is configured for the correct peptide. Examination of the MRM spectra of other major peaks (peak c in Figure 3A and corresponding mass spectrum in Figure 3C) and noise near the peak of interest can reveal the nature of potential interferences and guide the selection of transitions for MRM that have lower background interference.

It is common to find additional transitions in the full MS/MS scan that cannot be explained by the given parent ion. These are likely to arise from fragmentation of unrelated peptides with similar parent m/z values considering that the sample at this stage has not been enriched for the target protein.

### Table I: Prototypical MRM chromatographic method

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–1.5</td>
<td>100% @ 200 µl/min</td>
<td>0–12% @ 200 µl/min</td>
</tr>
<tr>
<td>1.5–2.0</td>
<td>0–12% @ 200 µl/min</td>
<td>12–45% @ 200 µl/min</td>
</tr>
<tr>
<td>2.0–16.0</td>
<td>12–45% @ 200 µl/min</td>
<td>45–80% @ 200 µl/min</td>
</tr>
<tr>
<td>16–16.2</td>
<td>45–80% @ 200 µl/min</td>
<td>80% @ 600 µl/min</td>
</tr>
<tr>
<td>16.3–173</td>
<td>80% @ 600 µl/min</td>
<td>80–0% @ 600 µl/min</td>
</tr>
<tr>
<td>17.3–174</td>
<td>80–0% @ 600 µl/min</td>
<td>100% @ 600 µl/min</td>
</tr>
<tr>
<td>17.4–18.9</td>
<td>100% @ 600 µl/min</td>
<td>100% @ 200 µl/min</td>
</tr>
<tr>
<td>18.9–19.0</td>
<td>100% @ 200 µl/min</td>
<td></td>
</tr>
</tbody>
</table>

A: 0.1% formic acid in H2O, B: 0.1% formic acid in acetonitrile, 2mm × 50 mm 2.5 µm CI18 column (XBridge from Waters, Milford, MA) at 5°C.
Furthermore, the full MS/MS spectrum from this short HPLC (Figure 3E) is likely to contain more unexplained transitions than the original full MS/MS spectrum from the profiling study (Figure 2A and B) due to poorer HPLC separation. Ultimately, the identity of the peptide needs to be confirmed using a synthetic peptide for both retention time and MS/MS fragmentation, but this process can be postponed until a decision is made to develop a more rigorous quantitative assay. It is recommended that more than one peptide from the same protein be monitored to reduce the possibility of being misled by incorrectly assigned peptide or confounded transitions.

Once all the target peptides are identified, the peptides may be grouped by retention time into multiple segments and an MRM-only method (i.e. no full scan MS/MS acquisition) can be set up. The maximum number of scans in a segment depends on the peak width and scan time, which is affected by HPLC conditions and mass spectrometer instrument parameters, respectively. In general, the total number of measurements across a peak should be more than 10 from shoulder to shoulder in order to ensure high-quality peak integration. Under the HPLC condition given earlier with an LTQ mass spectrometer, no more than four scans can fit in a single time segment satisfying this criterion. HPLC conditions can be optimized to more evenly distribute peptides into different segments to maximize the number of peptides measured within a single run.

Confirmation of profiling experiment results

With a preliminary MRM method in place, the initial findings from a profiling study can be confirmed with the more sensitive and selective MRM assay. Although it is both necessary and informative to use a wide mass selection window for product ions during signal collection for the reasons explained earlier, the signal-to-noise ratio can be substantially improved during peak area integration of the individual peptide by restricting signal integration to a more narrowly defined product ion mass range. For product ions with m/z values of <700, abundance of fragments with heavy isotopes are usually <20% of the monoisotopic fragment. Therefore, integration of only the monoisotopic mass by choosing a mass range from monoisotopic m/z –0.5 to monoisotopic m/z + 0.5 of the product ion can eliminate most of interferences without sacrificing the signal significantly. For product ions with m/z values >700, the abundance of the $^{13}$C or $^2$H heavy isotope becomes significant, often exceeding that of the monoisotopic fragment. In this case, the mass integration range can be extended to (monoisotopic m/z –0.5 to monoisotopic m/z + 1.5) for higher signal intensity as long as there is no interfering noise within the extended window.

Due to the low concentration of most analytes and the abundance of interferences in the sample at this stage, methods are needed to assess the reproducibility of the peak area measurement for each peptide or transition. In addition to the analysis of isotopic distribution, another important and powerful tool for evaluating if any of the measured transitions are confounded by unrelated interferences is the analysis of the relative intensities among different product ions. A significant difference between the study sample and the synthetic peptide in the relative intensities of different product ions is an indication that the indicated product ion in the study sample is contaminated by an unrelated interference. Such transitions can be eliminated during the construction of the MRM method. Even when a standard peptide is not available, analysis of the ratios of signal intensities among different transitions in the study samples can be used to detect those transitions that are contaminated by unrelated compounds because the relative intensity of individual transitions is independent of parent peptide concentration. A high value of standard deviation in these ratios among study samples is often an indication that measurement of certain product ions in some samples is substantially confounded by unrelated compounds. Such transitions can be eliminated from sample analysis and subsequent MRM assays. This analysis cannot be applied to those peptides with a single dominant fragment ion as is often observed from peptides with a single Pro for reasons discussed above during peptide selection.

Since multiple peptides are used for the same protein and no calibration standards are usually available at this stage, peak areas for individual peptides are compared directly between treatment groups. Peak areas for individual transitions from the same peptide may be summed. Average values of the peak area within a treatment group can be compared to that of the control group using an analysis of variance (ANOVA) model which can be readily implemented using commercial data analysis software like Excel or JMP [13, 14]. When multiple peptides are measured for the same protein, the ratio of the
average peak area of the treatment group to that of the control group will be reasonably reproducible (e.g. CV <20%) among all peptides from the same protein. Large variation in this ratio may indicate high level of noise or false identification for some of the peptides measured.

If the initial findings are confirmed with the preliminary MRM method then proceeding with the development of an analytically validated quantitative assay follows. Synthetic peptides need to be prepared for each chosen peptide as part of final confirmation of peptide identification while development of the quantitative assay is being pursued. Full MS/MS spectrum, relative intensities of different transitions and HPLC retention time should be confirmed. Mass spectrometer instrument parameters can be further optimized using the synthetic peptide for maximum sensitivity.

Quantitative and high-throughput MRM development
Certain applications identify potential biomarkers from a proteome profiling analysis of tissue with a goal of developing an accompanying plasma-based assay. Many times the sensitivity of the MRM with minimal depletion of the abundant plasma proteins will not be sufficient. In such cases, it may be possible to improve the overall sensitivity using quantitative sample preparation methods (e.g. precipitation for small peptides, Liposorb for apolipoprotein enrichment, etc.) while in other cases some form of immunoaffinity enrichment will be required. Preliminary HPLC conditions may be further optimized to minimize run times while maintaining adequate chromatographic separation. While the approaches taken for sample preparation and chromatographic separation are highly dependent on the specifics of the application, sample matrix and peptides being quantified, we provide some guidelines that may be broadly applicable. Sensitivity of the assay depends more on the sample preparation method than mass spectrometric methods. The following guidelines may be considered during development of a high-throughput assay method.

Take advantage of unique features of the target protein or peptide for enrichment method. For example, small peptides are much more soluble in the presence of organic solvents such as methanol, or acetonitrile than larger proteins. Careful choice of organic solvent concentration can eliminate most proteins by precipitation while leaving the target analyte in solution. Examine if the target protein has unusually low or high pI, certain types of glycosylation or affinity for certain small molecules. High or low molecular weight may be leveraged to selectively enrich by using molecular sieve membranes such as Centricon filter units. These non-immunological enrichment methods tend to have a large capacity and are not subject to limitations by competition when an internal standard is employed.

When an antibody is used for enrichment, consider immobilizing the antibody to a solid surface such as magnetic beads. If the antibody is used in solution without pre-conjugation to a solid surface, immunoglobulins in the sample can compete during the capture step. Such competition not only limits the capacity of the capturing method but also substantially increases noise during MS/MS measurement of the analyte.

Consider the throughput of the assay. Even though robotic handling of samples is not practical for most labs, use of 96-well plates or similar devices and multi-channel pipettes for all steps during sample preparation can increase the throughput of the assay without capital investment. Use of a 96-well filter plate for all steps of separation not only increases the throughput but may also improve recovery of the analyte and uniformity among samples. For example, protein–antibody–solid resin complexes can be captured and washed within a filter plate then eluted into a collection plate. Alternatively, the whole complex can be digested directly with trypsin within the filter plate without pre-elution after sealing both top and bottom of the plate. When using 96-well filter plates, we strongly recommend using centrifugation rather than a vacuum manifold for higher handling capacity, uniformity of filtration time between samples and lack of cross-contamination when collecting filtrates.

Consider that physico-chemical properties of the analyte peptide may be significantly different from those of the parent protein once the protein is digested. Addition of a chaotropic agent such as urea, or detergent may be necessary to keep both parent protein and digested peptides in solution. Although use of detergent is considered incompatible with mass spectrometry by many researchers, detergent can be separated from the analyte peptides by HPLC and diverted away from the mass spectrometer.

Once the sample preparation method is developed to enable handling of a large number of samples, HPLC run time determines the overall
throughput of the assay because the HPLC-MS method measures one sample at a time. Efforts to reduce HPLC run time and to separate analytes from interferences is one of the most rewarding investments for assay development.

Best separation of the analyte from other peptides is usually achieved under isocratic elution of the column. When multiple peptides are measured in one run, multiple step gradients rather than a continuous gradient, composed of a series of isocratic elution each optimized for different set of target peptides can be used.

Steady-state, rather than equilibrium operation of the column is a practical method to achieve minimum necessary separation with short run times. Isocratic elution conditions need to be empirically determined because the elution condition under steady-state can be significantly different from that under equilibrium. One caveat for steady-state operation is that the retention time can change noticeably when the column is changed, even when the new column came from the same manufacturer as the old column.

HPLC run times of <5 min are achievable without sacrificing resolution when the number of analytes multiplexed in the MRM is small. It may be necessary to run the column at a higher temperature (e.g. 55°C) or at a higher flow rate, or both to allow faster equilibration between mobile and stationary phase in order to decrease overall run time. Check to ensure that the number of measurements across the width of the peak is at least 10. Instrument parameters such as ion collection time or dwell time can be adjusted to decrease cycle time, although care should be exercised not to compromise instrument performance. If the peak is too sharp, it may be necessary to purposely introduce peak broadening. Using lower organic solvent concentration can achieve this effect with an added benefit of increased separation from interferences although this is likely to increase overall run time.

Test the extent of carry-over of the analyte peptide from one HPLC run to the next and make sure that the carry-over does not compromise the desired precision of measurement. For example, 1% carry-over from one sample with 100 times the signal than the next will produce 100% increase in the measured signal for the second sample. Although it is not common to observe such a wide range of concentrations in biological samples within one experiment, more than a 100-fold difference in concentration between adjacent samples can routinely occur if the sample injection order is randomized including calibration standards. Include a blank injection before the lowest calibration standard and after the highest standard.

When the HPLC flow rate is increased to shorten the run time, sensitivity of the mass spectrometer is usually decreased. Mass spectrometer instrument parameters should be adjusted to maintain instrument sensitivity. Higher gas flow, higher capillary temperature and higher lens voltage are some of the parameters to adjust.

**Internal standard development**

The use of isotopically labelled internal standards may be required to normalize for variability in sample handling, sample preparation, injection volume and mass spectrometer performance. The same concentration of isotopically labelled internal standard will be spiked into all samples including calibration standards. Ratio of peak area of the analyte (either individual transition or summed transitions for the same peptide) to that of the isotopically labelled internal standard can be used to construct a calibration curve and calculate concentration of the analyte. If the random variability from the HPLC and mass spectrometer exceeds the variability from sample handling and preparation, normalizing to an internal standard can result in an overall more variable measure of an MRM transition. Therefore, data-driven decisions should be taken on whether or not a transition should be normalized to an internal standard. The following guidelines may be applied to the development of internal standards:

The ideal internal standard will be a whole protein uniformly labelled with heavy isotopes. This will guarantee identical behaviour of the internal standard with the target analyte during all steps of sample handling and measurement. However, such reagents are often not practical to generate. Peptide internal standards are easier to prepare by simple chemical synthesis as long as the internal standard peptide behaves similarly through all steps of sample preparation and analyte enrichment. For example, when antibody enrichment is used, the target analyte peptide is often different from the antibody epitope. A hybrid peptide containing both analyte peptide with heavy isotope and the epitope peptide can be synthesized with appropriate trypsin cleavage sites. A similar peptide without heavy isotopes may also be used as an external calibration sample when
a protein calibration standard is not available (e.g. assay for procollagen type-I N-terminal pro-peptide [PINP]) [11].

For peptide internal standards with a heavy isotope label, a common method to prepare them is to use isotopically pure amino acid precursors during peptide synthesis at a specific site. Make sure that the difference in the m/z (not mass) values between the labelled and unlabelled peptide is bigger than the m/z selection window of the parent ion to avoid any cross-contamination between the standard and analyte. If this is not practical, for example, due to cost constraints, then ensure that the labelled amino acid residue or its unlabelled counterpart is included in all product ions measured for the MRM to reduce the cross-contamination.

While the labelled internal standard (e.g. 15N- or 13C-labelled peptide) has the same chemical property as the target analyte, the background interference can be different due to the m/z differences with the standard relative to the target.

The internal standard does not necessarily have to be chemically identical to the target analyte. When multiple analytes are measured, it is often impractical to have a chemically identical internal standard for all analytes in a multiplexed assay. In these cases, non-matched and non-labelled standard protein can be spiked into all samples and peptide(s) from the standard protein can be used to normalize peak areas for all analytes in the same sample to correct for variation in the sample recovery or instrument performance. It is important that the standard protein mimics the behaviour of the target proteins during sample preparation and that the analyte sample does not contain the same peptide.

The internal standard should be added at the lowest concentration that gives reasonable signal intensity without significantly affecting capacity of the enrichment method, but high enough concentration so that measurement of the internal standard does not become another source of assay variability.

Internal standards should be added at the earliest point of sample preparation as possible. The internal standard can compete with the target analyte if a fixed capacity enrichment method is used. Therefore, it is necessary to have external calibration samples to compare the ratio of the analyte to the internal standard—a simple ratio to linearly extrapolate concentration is not acceptable for calibration.

### Preparation of calibration samples

A quantitative assay requires an external concentration response curve suitable for regression of the transition peak area to the concentration of analyte in the calibration samples. The following guidelines apply to the design of a suitable external calibration sample set:

Prepare serial dilutions of the standard protein at a fixed ratio, rather than fixed interval, into a dilution matrix to prepare calibration standards. Serial dilutions of a fixed ratio will provide even spacing on the log scale for better calibration curve fitting (e.g. four-parameter logistic fit using log10 concentration). The calibration standard spiked into the dilution matrix does not have to be the entire protein of the target peptide(s), but it should mimic the behaviour of the target peptide during all courses of sample preparation and analysis. A short peptide can be used as long as it can reproduce the behaviour of the full-length protein (e.g. PINP assay) [11].

Choose a dilution matrix for the calibration standards that mimics the experimental sample in composition but does not contain the target analytes. Plasma from other species (provided target analyte peptide sequences are different from those used to develop the MRM) is often useful. Using phosphate buffered saline solution as a background matrix should be avoided.

Pooled plasma from the target animal species can be used as a source of protein with serial dilutions made into plasma from another species. Experimental samples should also be diluted into the same dilution matrix at a fixed ratio. This method provides ideal assay conditions because the calibration samples and the study samples have similar overall composition. Protein concentrations in the study samples can be reported as percent of the pooled plasma, which provides the concentration in terms of deviation from a reference concentration. Although not a popular method, concentration in this scale provides biological relevance of the measured concentration without having to know what the normal concentration of the protein is.

Survey a broad collection of experimental samples to estimate the concentration range of the target analyte in study samples before preparing a large volume of calibration standards. The same protein can show more than 10-fold difference in concentration at different ages or between different strains of the same age animals. For example, the apoB protein concentration in LDL receptor knockout mice is
more than 10-fold higher than the wild-type mice. Assay sensitivity, dynamic range and concentration range of the calibration samples need to be optimized with consideration of the known concentration range of the target analyte.

Prepare a large volume of calibration standards, aliquot and freeze. Unless stability during multiple freeze–thaw cycle has been tested, use each aliquot only once. Also prepare a large batch of positive control samples to be used for in-study validation by combining a large pool of samples with similar target analyte concentrations. Make aliquots and freeze. Include the positive control in all experiments. The positive control samples can be used to track historical performance of the assay over time and to bridge between different batches of calibration standards.

Data processing
Given a set of MRM transitions to monitor, there are multiple choices for how to integrate the chromatographic peaks, combine the transition peak areas, and fit a calibration curve to the calibration standards. Data-driven decisions using a precision profile generated from replicate analyses of the calibration standards are recommended as an objective means to compare different data-processing approaches [7, 15, 16]. The following guidelines list several topics to consider in the data-processing stage of assay development:

Mass spectrometer instrument data systems typically provide several options for peak integration. Ideally peak integration can be referenced to an offset relative to the elution of an internal standard rather than relying on fixed chromatographic time windows. For the LTQ, one transition can be chosen for each peptide as a reference by searching for the highest peak within a fixed retention time window, and then the rest of the transitions for the same peptide can be assigned by searching for a peak with the closest retention time. Baseline subtraction is usually not an issue with MRMs due to their specificity relative to other, high background, measurements (e.g. HPLC-UV/Vis spectroscopy). The specific analytes of interest, background matrix and sample preparation/chromatographic separation will determine which peak integration approach is optimal for a particular assay.

It is important to qualify each transition for any potential interference caused by biological modulation related to the disease or treatment under investigation. Fitting a fixed-effects linear ANOVA with a TREATMENT_GROUP x TRANSITION interaction term in the model is a simple method to test for significant group-related interferences. These ANOVA models are readily implemented in desktop data analysis programs like Excel or JMP [13, 14]. Transitions with background signal affected by the biological process identified by a statistically significant TREATMENT_GROUP x TRANSITION interaction term should be avoided. Testing for this interaction periodically as part of the in-study validation phase is also recommended.

If the peak area for an individual transition is low, consider summing it with another transition. Transitions with large differences in peak area are not good candidates for summation. The determination of whether or not summing transition areas is advantageous should be guided by the precision profile results.

A weighted four-parameter logistic fit of the transition or summed transition peak area to the log10 concentrations of the spiked calibration standards is recommended [17]. Weighting is typically proportional to a power of the mean variance function. Alternatively, inverse of the x-value (log10 concentration in this case) or inverse of x2 can be used especially when single measurements are used for calibration standards. Without proper weighting, samples with high concentration will dominate the regression and optimal fitting of low concentration samples would not be achieved. Other calibration regression models may be considered (e.g. an exponential model), especially in cases where system saturation is not observed. The most common source of system saturation comes from the sample enrichment method due to the limited capacity of the enrichment reagent. Mass spectrometer detector saturation is usually not limiting with biological samples since the instrument dynamic range is many orders of magnitude higher than sample concentration ranges and concentration of even the most abundant protein in most biological system is far less than instrument capacity. Performance obtained on the precision profile analysis or the spiked recovery total error profile (discussed later) should be used to inform the choice of calibration regression model.

Transition (or summed transition) peak areas are backfitted against their corresponding calibration curve to obtain concentration estimates for each
transition (or transition sum). The mean or the weighted mean (weighting done proportional to the backfitted standard error) of the transitions from a peptide are used to estimate the peptide concentration. Likewise, the mean or the weighted mean of the peptide concentrations from a protein are used to estimate the protein concentration.

**Spiked recovery analytical validation**

Assay dynamic range, or working range, extends from the lower limit of quantification (LLOQ) to the upper limit of quantification (ULOQ). Concentrations between the LLOQ and ULOQ have acceptable limits of accuracy, precision and total error. Estimates of the working range are often the most useful performance metrics for an assay. Samples measured above the ULOQ need to be diluted in order to report reliable analyte levels (provided that dilution linearity has been demonstrated). Samples with analyte levels below the LLOQ cannot be reliably quantified due to sensitivity limitations. The term LLOD (lower limit of detection) is often mistakenly used to describe the assay sensitivity or confused with LLOQ. LLOD is the lowest concentration that results in signal intensity significantly higher than background (usually defined as +2 or +3 SD of the background signal level). The spiked recovery experiment consists of testing a series of validation samples with known, spiked, concentrations of the analyte of interest analyzed over multiple replicates within a day (batch) and over several days (batches). The difference between the true and estimated analyte concentration of these validation samples obtained by backfitting to their associated standard curve is used to create profiles of bias, precision and total error over the concentration range of the validation samples. The LLOQ and ULOQ are estimated based on the lowest and highest validation sample concentrations that meet the criteria for bias, precision and total error. It is important to include multiple (≥3) days, or batches, as well as multiple (≥2) replicates within-day in order to account for both within- and between-day variability effects. For many assays the inter-day variability will be much larger than the intra-day variability. The equations for estimating the error terms as well as guidelines for error thresholds to define LLOQ and ULOQ are reported in DeSilva et al. [15].

**CONCLUSION**

The gap between clinically deployed biomarker assays and discoveries made using modern proteomics discovery technologies is now widely recognized and garnering attention from many investigators. This gap, to a large extent, is due to the lack of efficient approaches to determine which biomarker discoveries warrant the substantial investment required to see a discovery achieve clinical utility. The targeted MRM assay is a powerful tool that can be used as a link between biomarker discoveries and robust, validated and high-throughput clinical assays. We have shown that an ion trap mass spectrometer (e.g. LTQ) can be used to develop highly sensitive and quantitative MRM assay [11]. By strategically using the same mass spectrometry platform (e.g. ion traps) for the discovery and MRM validation phases of biomarker development, the efficiency of transitioning to the more selective and sensitive MRM approach is maximized. This efficiency is largely due to the ability to use the proteomic profiling MS/MS fragmentation spectra from the discovery phase as starting points for selecting good candidate transitions for MRM development without having to wait for standard peptide synthesis. By using the approach outlined here in which the isotope pattern of product ions are evaluated for interferences, candidate MRM transitions can be quickly evaluated. By selecting interference-free transitions from a template chromatographic method, substantial assay development time can be saved by avoiding time-consuming chromatography optimization. This is particularly important when multiple analytes are measured simultaneously, in which case case optimization of chromatographic separation for individual peptide becomes extremely difficult, if not impossible, as the number of analyte peptides grows. Ability to assess interference for each transition at all stages of assay development and implementation significantly reduces the risk of making substantial investments into assay development only to find later that the initial discoveries were misled by interferences. Multiple isotope peaks from product ions may be summed for quantification, thereby increasing the sensitivity achievable for interference-free transitions. By applying the guidelines included in this report, we have routinely translated biomarker discoveries into MRM methods within a day. Biomarker discoveries that confirm with the more selective and sensitive MRM methods can then be taken
forward for assay optimization and conversion to a fully quantitative and validated assay. The ability to efficiently develop a multiplexed panel of MRM assays for biological confirmation will help expedite the translation from biomarker discovery to demonstrable impact. Detailed examples are provided in the Supplementary Data for most of the steps described earlier in the article to aid speedy development of MRM assays starting from results of profiling studies.

SUPPLEMENTARY DATA
Supplementary Data are available online at http://bfgp.oxfordjournals.org/.

References