Multiple reaction monitoring (MRM) is a mass spectrometry technique based on selection of a peptide ion and one or more characteristic fragment ions. The MRM method has previously been shown to be specific, accurate, and reproducible between laboratories. Through the use of retention time windows, MRM-based quantitation can also be multiplexed to analyze and quantitate hundreds of proteins per run, thus increasing the throughput of this type of assay and making it rapid enough for clinical applications. Although MRM-based quantitation is typically used for the clinical validation of potential biomarkers, in this paper, we demonstrate the use of a panel of MRM assays for biomarker discovery.

We previously reported on a method for the quantitative analysis of the 45 most abundant plasma proteins, based on liquid chromatography MRM/mass spectrometry, with quantitation accuracy provided by the use of stable isotopically labeled standard peptides. In this paper, we describe the use of this MRM method for the discovery of biomarkers characteristic of the presence or absence of coronary artery disease. Quantitative analysis of peptides specific for 44 of these abundant proteins was performed on plasma from a group of 38 patients classified as coronary artery disease-positive or coronary artery disease-negative. Statistical analysis of the results resulted in the discovery of a panel of 5 proteins with discriminative concentrations between patients with and without coronary artery disease, with a \( P \) value of <0.05. An initial evaluation of the analytic pipeline by cross-validation resulted in an estimated sensitivity and specificity of 74% for the test samples.

This demonstrates that MRM-mass spectrometry-based quantitation offers a potential methodology for both biomarker discovery and biomarker validation. Similar to other discovery methodologies (eg, isobaric tags for relative and absolute quantitation), MRM-mass spectrometry provides the sensitivity and accuracy needed in the discovery phase, but, unlike the others, MRM-mass spectrometry offers the high throughput needed for clinical validation. Thus, this advantage eliminates the need for the development and validation of different discovery and clinical validation methods.

**Introduction**

**Mass Spectrometry**

Mass spectrometry (MS) is an analytical chemistry technique based on the determination of the mass (actually, the mass-to-charge ratio) of an analyte ion. This analyte can be fragmented inside the mass spectrometer, which is under vacuum, to give charged product ions. This process is known as collision-induced dissociation. The masses of the fragment ions can then be determined in a second stage of MS, giving rise to the terms MS/MS or tandem mass spectrometry (see Figure 1a). Because the fragments and their abundances are determined by the structure of the analyte, this fragmentation pattern and the masses of the product ions are characteristic of the analyte. Both the molecular ions and the fragments can be used for analyte identification and quantitation.

MS-based quantitation is usually based on peak heights or peak areas that can be either acquired during full-scan acquisitions over the entire mass range or by selecting masses that are characteristic of the compound being monitored (selected ion monitoring or selected ion recording). Because the presence of stable isotopes (\(^{13}\)C, \(^{15}\)N, \(^{18}\)O) causes predictable shifts in the masses of the ions that contain them, without changing the chemical properties of the analyte, stable isotope-labeled analogues of target analytes have long been used as internal standards for MS-based quantitation. Absolute quantitation using stable isotope-labeled standards is based on determining the relative peak heights or peak areas of the analyte in the sample as compared with a standard of known concentration, either in the MS or in the MS/MS mode. This means that the masses of the light and heavy precursor ions are monitored. The advantage of using stable labels is that both forms can be injected in the same run, thus compensating for any run-to-run variations in sensitivity, any sample-specific suppression of the ion signals, or any shifts in retention time.

With an instrument capable of tandem MS, additional specificity can be obtained by monitoring a precursor ion and 1 of its collision-induced dissociation-generated product ions (selected reaction monitoring or multiple reaction monitor-
If stable isotope-labeled standards are used, the precursor and product ions of the internal standard should co-elute with those of the endogenous component in the sample, thus improving the resulting specificity. This process will also reduce the possibility of a false-positive identification and inaccurate quantitation. Multiple reaction monitoring (MRM) using stable isotope-labeled standard peptides (SIS peptides) has been called the gold standard of MS-based quantitation methods and has also been found to be reproducible between laboratories. In addition, MRM analysis can be multiplexed (ie, the use of timed ion selection windows during the course of peptide elution allows the quantitation of hundreds or even thousands of target molecules in a single analysis).

All of the above techniques were first applied to small-molecule analysis: the analysis of compounds that could be easily volatilized.

**MS-Based Proteomics**

In the 1980s, advances in MS ionization techniques for nonvolatile or thermally labile analytes (electrospray and matrix-assisted laser desorption/ionization) led to the application of MS to the study of proteins in biological systems. Improvements in instrumentation and software led to rapid advances in proteomics, starting in the 1990s and continuing through the present day. The rapid advance of MS-based proteomics is also owed to the fact that collision-induced dissociation results in cleavage between the amino acids of a peptide, thus facilitating both the prediction of peptide MS/MS mass spectra and the interpretation of experimentally derived spectra.

Most proteomics experiments are still done by using a combination of enzymatic digestion of a protein or the proteins in a biological fluid such as plasma, followed by an on-line or off-line high-performance liquid chromatographic (LC) separation of the resulting mixture of peptides, followed by MS. This peptide-based approach is termed bottom-up proteomics because the identity (and other information on the intact protein) is derived from information on the peptides generated from it.

**Quantitative Proteomics**

Now, proteomics is moving from a qualitative science (“which proteins are present in the sample?”) to a quantitative science (“how much of a particular protein is in the sample?”). Quantitative proteomics actually first proceeded through a relative quantitation stage (ie, is more or less protein present as a function of treatment?) This relative quantitation stage has led to the discovery of many potential biomarkers through exhaustive comparative studies, with the relative protein expression levels determined by differential in-gel electrophoresis (DIGE) or other 2-dimensional gel methods that analyze intact proteins or by a variety of peptide-based stable label or label-free techniques (for a recent review, see).

Bottom-up proteomics, described above, relies on the use of peptides as surrogates for their parent proteins. The use of stable isotope-labeled internal standard peptides (SIS peptides) has enabled MS-based absolute quantitation of the parent protein, based on the relative abundances of a spiked-in SIS peptide analog of a peptide characteristic of and specific to that protein (a proteotypic peptide). The digestion process must be either complete (1 mole of protein = 1 mole of each peptide), or it must be reproducible, so correction factors can be calculated. This is a critical and often-ignored
part of a proteomics experiment, but progress is being made on characterizing various digestion and extraction protocols.7

Plasma Proteomics
Plasma is readily available and is the most widely tested biological fluid in human diagnostic assays, but it is also a particularly challenging biological matrix for proteomics studies because of the wide dynamic range of proteins present (1010).8 Depletion of abundant proteins through the use of antibodies is often required, especially for the detection or quantitation of very low abundance analytes. This depletion step adds significantly to the cost of the analysis, as well as being a potential source of variability. High-abundance plasma proteins, however, can also be potential biomarkers, and these proteins can be measured without depletion. In fact, 31 of the 45 most abundant proteins in plasma have been previously proposed as biomarkers of CVD in cardiovascular-related literature.9

MRM in Quantitative Proteomics
Although MRM has been used for decades to measure small molecules, it has only been recently applied to the quantitation of proteins.6,10,11 To date, most research efforts have mainly focused on optimizing specific MRM assays, improving sample processing methods, and developing new software tools to manage MRM experiments.12–21 Previous work has shown the utility of MRM assays to perform targeted discovery proteomics,22–24 as well as to validate candidate protein markers.25,26 In this study, we demonstrate that a highly multiplexed MRM analysis can also be used for biomarker discovery.

MRM Method Development
The method that our laboratory uses for MRM method development has been previously described in detail.27 Briefly, an in silico theoretical digest of the target proteins is performed. BLAST (Basic Local Assignment Search Tool) searching is used to determine a peptide sequence that is unique to the target protein and is between 5 to 25 amino acids in length, which is optimal for MS detection. Peptides containing oxidizable amino acids (eg, methionine, cysteine) are avoided, if possible, because multiple versions of the same peptide can be formed. Similarly, peptides likely to give only partial cleavage (ie, peptides that contain multiple adjacent cleavage sites, such as KK, KR, etc) are also avoided. A list of up to 8 candidate peptides for each protein is determined based on these theoretical considerations, but the peptides are further screened by solubility and LC retention time based on data from a real plasma digest and/or by synthesis of the corresponding SIS peptides. The SIS peptides are then used to tune the accelerating and fragmentation voltages for maximum sensitivity of 1 peptide fragment from each peptide. Ideally, 3 of these peptides will be suitable for use in the final multiplexed method.

The development of the final MRM assay involves multiple steps, which are shown schematically in Figure 2a. This involves selection of candidate peptides, synthesizing SIS versions of these peptides, using these SIS peptides to improve the sensitivity of the assay, and (for the quantitation accuracy) creating a mixture of SIS peptides whose concentrations are similar to those in the samples. Since the final assay is done by online LC/MRM-MS, peptide retention times must also be considered in the final multiplexed version of the assay (Figure 2b).

In our laboratory, we developed a multiplexed MRM-based quantitation method for the top 45 most abundant proteins in human plasma, with quantitation based on SIS-labeled analogues of their proteotypic peptides.28 These 45 proteins can be quantitated in a single LC/MS/MS analysis, with high sensitivity and high quantitation accuracy. Specificity is achieved through the use of 3 peptide-specific criteria: the retention time, the peptide molecular weight, and the masses of the peptide fragments. We have found that optimized tuning of the MS and MS/MS parameters, using the SIS peptide as the standard, results in an average 11.4-fold increase in sensitivity. This tuning involves selection of the optimum precursor ion charge state, tuning the declustering potential and the collision energies, followed by selection of the most sensitive and specific fragment ion (Figure 1b).27 In this way, attomole sensitivities can be achieved, with quantitation precision of <6% and reproducibilities of <20% CV (coefficient variation), with the consumption of only 14 nL of sample per run, making MRM-based quantitation suitable for clinical analyses.

In the study described here, our MRM-based quantitation method was used to determine if a panel composed of a subset of these 44 proteins could be used to distinguish between patients positive (n = 19) or negative (n = 19) with respect to coronary artery disease (CAD), based on MRM analysis of their proteotypic peptides. A panel of 5 protein markers was first identified that individually discriminated between the 2 groups of patients (with a P value of <0.05) and, together, can achieve a satisfactory classification performance. Cross-validation results showed that a panel of proteins measured by our MRM-based methodology and identified by the same analytical methods can correctly classify 74% of the test samples.

Materials and Methods

Study Design
Two cohorts of 19 patients each were studied: The control group consisted of 19 CAD-negative patients. A second group of 19 patients were diagnosed with severe CAD. These patients had 1 or more lesions, and >50% stenosis. Patients were all male but were not age-matched.

Sample Collection and Processing
Plasma samples from 38 patients were collected in ethylenediaminetetraacetic acid tubes and stored at ~80°C until processed (for further details, see29). Sample preparation involves denaturation, reduction, and alkylation.29 In this study, plasma-tryptic digests were diluted with ammonium bicarbonate and denatured with sodium deoxycholate. The denatured plasma samples were reduced with tris(2-carboxyethyl)phosphine, and alkylated with 10 mmol/L iodoacetamide. Ammonium bicarbonate was used to lower the sodium deoxycholate concentration before tryptic digestion.

Synthesis of SIS Peptides
Isotopically labeled standard peptides corresponding to each of the 44 target proteins were synthesized, as described in Kuzyk et al.27,28 and purified by reversed-phase high-performance LC. To ensure
accurate quantitation, the concentration of each synthetic peptide was determined by acid hydrolysis, followed by amino acid analysis (AAA). Capillary zone electrophoresis with ultraviolet detection (190 to 210 nm) was then used to correct the AAA-determined peptide concentrations to reflect the percent purity of the target peptide. For accuracy, the mixture of peptides added to each plasma digest was “balanced” so that the concentration in the standard mixture approximated the concentration in normal plasma. The mixture of SIS peptides was added after the digestion step to avoid degradation of the SIS peptides.

Online LC/MRM-MS Analysis

For quantitation of unknown samples, a mixture of stable isotopically labeled internal standards corresponding to each of the target peptides, synthesized and purified as described above, was added to each plasma digest. Samples were desalted and concentrated prior to MS analysis by solid phase extraction, and the eluted samples were frozen and lyophilized to dryness. Prior to LC-MRM/MS analysis, samples were reconstituted in 0.1% formic acid to a concentration of ~1 μg/μL, and the SIS mixture was added. Samples were extracted and analyzed in random order.

For the online LC/MRM-MS analysis, an Eksigent NanoLC-1D Plus high-performance LC was used for the injection of 1 μL of the desalted plasma digest samples. Peptides were separated by online reversed-phase capillary LC, using a 75-μm identification column, a flow rate of 300 nL/min, and a 41-minute gradient. An AB Sciex 4000 QTRAP with a nano-electrospray ionization source was used for all LC-MRM/MS analyses. Details on the peptides used, the

---

Figure 2. A, Workflow for development of a multiple reaction monitoring (MRM) method. This is an iterative process, as can be seen from this diagram. B, A final multiplexed MRM method: in this case, for the 45 most abundant proteins in plasma (reprinted from 28 with permission).
A panel of proteins with significant differential concentrations (with a value of different ion fragments, when available, within each sample. A inclusion of a factor to capture the variation due to the measurement (triplicates) of each patient sample. This model also allows the with and without CAD, while accounting for repeated measurements variance model to test for differences between groups of patients proteins were analyzed using a mixed between-within analysis of Figure 3. A, Schematic of the study design for determining biomarkers of coronary artery disease (CAD), using the multiple reaction monitoring (MRM)-based method for the quantitation of the 45 most abundant proteins in plasma. Three peptide/fragment ion pairs were used to determine the plasma concentration of each protein by comparison with the corresponding stable isotope-labeled standard (SIS) peptides and SIS-labeled fragment ions. B, Statistical treatment of the multiple reaction monitoring (MRM) data from these experiments.

scheduled MRM transitions, and the retention times are given in reference28 online-only supplemental material.

All MRM data were processed using MultiQuant 1.1 (Applied Biosystems), with the MQL (Method Quantification Limit) algorithm for peak integration, using a 2-minute retention time window, with “report largest peak” enabled and a 3-point smooth with a peak-splitting factor of 2. All data were manually inspected to ensure correct peak detection and accurate integration. Linear regression of all calibration curves was performed using a standard I/x (x=concentration ratio) weighting option, to aid in covering a wide dynamic range.

Statistical Workflow

The statistical workflow is shown in Figure 3a and 3b. All 44 proteins were analyzed using a mixed between-within analysis of variance model to test for differences between groups of patients with and without CAD, while accounting for repeated measurements (triplicates) of each patient sample. This model also allows the inclusion of a factor to capture the variation due to the measurement of different ion fragments, when available, within each sample. A panel of proteins with significant differential concentrations (with a P value of <0.05) between patients with and without CAD was identified for further analysis and validation. A classifier was constructed by Support Vector Machines that reflects the contribution of each identified protein marker to the hyperplane that represents the largest separation between the two groups of patients. A single number, called the biomarker score, was generated by the resulting Support Vector Machines classifier that aggregates the combined contribution of the selected proteins. In this multivariate analysis, a single protein measure was calculated for each patient consisting on the average concentration of all replicates and all ion fragments.

A leave-1-out cross-validation was used to estimate the sensitivity and specificity of the identified proteins in classifying new samples. Briefly, 1 patient is left out as a test sample, and the remaining samples are used as a training set. To obtain unbiased performance measures, the panel of differentially expressed protein biomarkers needs to be identified, based on the resulting training set. The identified protein biomarkers are used to construct a Support Vector Machines classifier, based on their measurements on the training set. The resulting classifier is then used to classify the test sample left out. This procedure is repeated until all patients have been used as a test set. The sensitivity and specificity are then estimated by the average classification performance.

Results

All of the plasma samples were processed, and the absolute concentrations of each of the 44 target proteins were measured in triplicate, relative to the known concentrations of the SIS peptides.

Using mixed within-between analysis of variance, 5 proteins were found with differential concentrations (with a P value of <0.05) in patients with and without CAD (Figure 4a). Although 3 of these proteins were already known to be associated with CAD,9 2 additional proteins that were not previously known to be correlated with CAD were found to be correlated with this disease. A principal component analysis was used to show how the protein biomarkers together discriminate the groups of patients with and without CAD (Figure 4b). The first 3 principal components of the MRM protein measurements of the 5 biomarker proteins demonstrate the ability of the panel to separate the analyzed groups. An initial validation was performed using the MRM measurements of the candidate proteins. A biomarker score based on a panel of proteins identified from a training set was calculated to classify test samples from a leave-1-out cross-validation using Support Vector Machines. As a result, 74% of the samples with CAD (sensitivity) and 74% of the samples without CAD (specificity) were correctly classified, demonstrating the potential of using a panel of proteins measured by MRM to classify new samples.

Conclusions

We previously created an MRM-based method for the absolute quantitation of 45 of the most abundant proteins in human plasma.28 In this new study, a panel consisting of 44 of these proteins was used as a panel to screen plasma from normal and diseased patients for potential biomarker proteins. The accuracy and reproducibility of this MRM-based assay allowed a subset consisting of 5 of these proteins to be distinguished from the other 39 proteins for use as a biomarker panel for CAD. Two of these proteins were not previously known to be correlated with CAD. We are currently in the process of expanding the number of proteins in this multiplexed assay. Using scheduled MRM assays, hundreds of proteins should be able to be analyzed in a single run, with atomole limits of detection on column.31,32 We are also converting these MRM methods to high-flow MRM methods suitable for use on the Agilent 6490 mass spectrometer and other next-generation electrospray instruments.
preliminary experiments indicate that the use of electrospray sources compatible with high-flow high-performance LC will shorten the analysis time by approximately a factor of 5, with an increase in sensitivity of the analysis.33

In conclusion, this current study demonstrates that a panel of multiplexed MRM assays can be successfully used as a tool for discovery proteomics. With regards to this particular CVD application, we intend to expand this preliminary study with a larger patient cohort and a larger array of MRM assays.

Acknowledgments
We would like to acknowledge John S. Hill for providing the samples, and Derek S. Smith and Dominik Domanski for performing the MRM analyses described in this paper. We would also like to acknowledge Carol E. Parker for assistance in preparing the manuscript.

Sources of Funding
The University of Victoria Genome British Columbia Proteomics Centre is partially funded by platform grants from Genome-British Columbia and Genome Canada. We are also grateful for support from the Networks of Centres of Excellence Centre of Excellence for Commercialization and Research Prevention of Organ Failure Centre of Excellence. Support for this research was also provided, in part, by the British Columbia Proteomics Network through a Small Projects Health Research Grant Award.

References


Key Words: biomarker, coronary artery disease, mass spectrometry, plasma, proteomics
Multiple Reaction Monitoring (MRM): Principles and Application to Coronary Artery Disease

Gabriela V. Cohen Freue and Christoph H. Borchers

doi: 10.1161/CIRCGENETICS.111.959528
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2012 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/5/3/378

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Genetics can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Genetics is online at:
http://circgenetics.ahajournals.org/subscriptions/