nature biotechnology

Options and considerations when selecting a quantitative proteomics strategy

Bruno Domon[1] & Ruedi Aebersold[1,2]

The vast majority of proteomic studies to date have relied on mass spectrometric techniques to identify, and in some cases quantify, peptides that have been generated by proteolysis. Current approaches differ in the types of instrument used, their performance profiles, the manner in which they interface with biological research strategies, and their reliance on and use of prior information. Here, we consider the three main mass spectrometry (MS)-based proteomic approaches used today: shotgun (or discovery), directed and targeted strategies. We discuss the principles of each technique, their strengths and weaknesses and the dependence of their performance profiles on the composition of the biological sample. Our goal is to provide a rational framework for selecting strategies optimally suited to address the specific research issue under consideration.

The proteome is more than the mere translation of the protein-coding regions of a genome. Processes such as alternative splicing, protein processing and post-translational modification are key to providing the full complexity of life. Moreover, because the abundances of proteins are often of great biological significance, these must often be tightly controlled[1]. Although the ultimate goal of proteomics is to both identify and quantify the full complement of proteins and their variants in any cell type under conditions of interest, neither the composition of a proteome nor the quantity of its constituents can be reliably predicted by computation or determined by experimentation. Nonetheless, several proteomic strategies now effectively support a range of strategies for biological experimentation.

Owing to limited availability and accessibility of suitable reagents, the majority of proteins in any species cannot be detected and quantified by affinity-based assays. Therefore, essentially all proteomic studies have used mass spectrometric discovery techniques, which are now capable of unambiguously identifying and quantifying thousands of protein components of complex samples (see reviews by B.D and R.A. and others; refs. 2,3). Of the several such discovery methods that have been developed, all involve digesting the protein sample into peptides, typically by trypsin, and then fractionating the resulting peptide mixture before it is subjected to mass spectrometric analysis. MS involves ionizing the peptides and selecting specific precursor ions from the pool of detected peptide ions for fragmentation. The resulting product-ion mass spectra, commonly generated by collisional activation, are recorded and used to determine the amino acid sequence of the selected peptides. Finally, the proteins present in the sample are inferred from the ensemble of identified peptides. In the most common implementation of the method, the precursor ions are selected automatically from the ions detected in a survey scan immediately preceding the ion selection, a process referred to as data-dependent analysis (DDA). When combined with an appropriate stable-isotope labeling strategy, these protein identification methods also permit relative (involving comparison to a reference sample) or absolute quantification of the identified proteins[2,4]. The various methods differ in the requirements for sample preparation, the extent of sample fractionation and the level (protein or peptide) at which fractionation is performed, the type of mass spectrometer used and their needs in data processing tools[5]. Most common implementations rely on electrospray ionization directly coupled to an instrument for liquid chromatography. However, alternative or complementary methods based on matrix-assisted laser desorption ionization (MALDI) have also been proposed[6,7]. MALDI enables repetitive, sequential interrogation of the same sample.

Multiple incremental improvements at each level of this fundamental process have substantially increased the number of proteins typically identified in proteomic studies, the confidence with which fragment-ion spectra resulting from collision-activated dissociation (CAD) are assigned to peptide sequences and the confidence with which protein identities are inferred. Overall, liquid chromatography tandem mass spectrometry (LC-MS/MS) with DDA is now a robust and powerful technology to detect and quantify proteins and their post-translational modifications, as exemplified by some recent large-scale studies[8–10]. As outlined elsewhere in this issue[11], however, the perpetual *de novo* discovery of the proteome or fractions thereof in every proteomic study may not be the most suitable and effective strategy to interface proteomics with biological research.

First, even though the fraction of the proteome identified in discovery proteomic studies has increased over time, the analysis of the complete proteome of even moderately complex cells remains challenging, expensive and slow. Second, the heuristics based on signal intensity used for the precursor-ion selection in DDA MS result in an irreproducible and incomplete sampling of the peptide mixture generated to represent the proteome. Consequently, different subsets of the proteome are identified and quantified after repeated analyses of identical or substantially similar samples. Therefore, partially overlapping proteome data sets are generated even when the parameters for measurement are carefully controlled[12,13]. Third, in each discovery experiment that is focused on a specific biological question (e.g., the proteomic changes induced by the stimulation of a cell sample with a

[1]Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland. [2]Faculty of Science, University of Zurich, Zurich, Switzerland. Correspondence should be addressed to R.A.:(aebersold@imsb.biol.ethz.ch) or B.D. (bruno.domon@crp-santa.lu).

Published online 9 July 2010; doi:10.1038/nbt.1661

© 2010 Nature America, Inc. All rights reserved.

particular drug), a large number of proteins with no relevance to the particular question are identified and quantified. Conversely, some of the relevant proteins are missed. Consequently, a model of the studied biological process has to be assembled from data sets that are noisy, incomplete and contain large amounts of irrelevant data. Moreover, the absence of a protein from the list of identified proteins does not indicate the absence of the protein from the sample. This complicates the comparison of lists of proteins identified in different studies. Fourth, no prior information on the system studied is used to design the experiment, even when it is being conducted in an area with the benefit of vast amounts of prior biological knowledge. Finally, there is a large discrepancy between the size and quality of reported proteomic data sets. Whereas a few highly specialized laboratories now routinely and reliably identify thousands of proteins per study[8–10,14], a more representative selection of proteomic laboratories identified the components of a sample consisting of 20 equimolar proteins only with considerable difficulty and after optimization of their methods[15].

In light of these challenges, we have argued that proteomics will make a larger and more immediate impact on progress in biology if reproducible and quantitatively accurate data can be generated for all the proteins that constitute a particular system or process (R.A. and colleagues[16]). In such a scenario, sets of proteins are defined from prior biological knowledge and then identified and quantified by targeted MS to generate complete, accurate and reproducible data sets that represent the whole system studied under different conditions. Such a proteomic strategy supports the standard way of biological inquiry, where specific hypotheses (proposed explanations for observable phenomena) are generated from the available knowledge and then tested. In contemporary proteomics, the hypothesis is almost invariably that unique and tight regulation of a group of proteins underlies the function or process of interest.

Recent years have witnessed the emergence of several MS-based methods that advance such targeted proteomic strategies. The consistent and reproducible detection of complete sets of proteins in multiple samples and their accurate quantification is important for a wide range of biological studies. In particular, this applies to biomarker research and systems biology, where quantitative data of a system in multiple perturbed states are critical for the mathematical modeling of the process in question.

Three MS strategies

The generic overall process by which peptides are identified and quantified in MS-based proteomics follows the sequence of events indicated in the core section of **Figure 1**. First, the ionized peptides present in the sample solution are transferred into the gas phase, most commonly using the (nano)electrospray technique. Alternatively, peptides deposited on a solid surface are ionized by MALDI. Second, the mass-to-charge ratio (m/z) of the generated peptide ions (precursor ions) is measured, and the mass of the peptide is determined implicitly. Third, selected precursor ions are isolated sequentially in the gas phase. Fourth, the selected precursor ions are fragmented, most commonly through CAD. Fifth, fragment-ion masses are analyzed and recorded as product-ion spectra. From these, the peptide sequence is inferred from the ensemble of fragment-ion masses. The quantity of a peptide is determined from the signal intensity of the precursor ion, most commonly by comparing this to the signal intensity of an isotopically labeled reference peptide of identical sequence. The MS steps in this technique are preceded by sample preparation protocols (that generate suitable protein samples, proteolyze the proteins into peptides and separate the peptide samples) and followed by post-acquisition data processing and analysis.

Over the past decade, three main MS-based strategies have emerged, which we subsequently refer to as shotgun (or discovery), directed and targeted proteomic strategies. These are distinguished by the way in which the individual steps are performed and connected. The shotgun (discovery) approach has been the most widely used method and has generated the vast majority of proteomic data available today. Directed and targeted MS methods, which are still emerging, support proteomic strategies in which prior information is used to define sets of peptides or proteins to be analyzed selectively. We next describe the principles of each method.

Shotgun (or discovery) proteomics

The hallmark of the shotgun proteomic method is the selection of peptide ions detected in a particular sample and their fragmentation by simple heuristics, based on signal intensity. The principles and information pertinent to the technique are summarized in **Box 1** (see also **Fig. 2**) and some of its potential pitfalls are discussed elsewhere (R.A. and others[17]). In a shotgun experiment, the masses (more precisely, m/z) of the ions produced in the ion source at a particular time are recorded to generate a mass spectrum, often referred to as a survey scan. The mass spectrometer then automatically selects one of the detected peptide ions, called a precursor ion, isolates it, subjects it to fragmentation by CAD and records the resulting fragment-ion mass spectrum. This process is called product-ion scanning. Because a cycle that comprises a survey scan and a product-ion scan is fast (~100 ms) compared with the chromatographic elution time of a particular peptide (~30 s), and because many precursor ions are typically

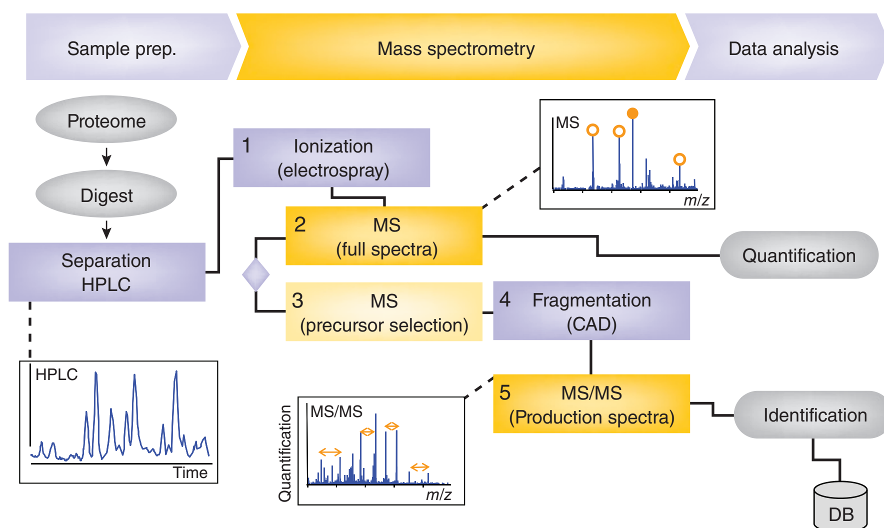

Figure 1 Workflow of a typical proteomic experiment. Proteins are digested to produce a complex mixture of peptides, which are separated by HPLC before analysis by MS. The overall process consists of a number of steps, specifically the ionization of the peptides, acquisition of a full spectrum (survey scan) and selection of specific precursor ions to be fragmented, fragmentation, and acquisition of MS/MS spectra (product-ion spectra). The data are processed to either quantify the different species and/or determine the peptide amino acid sequence through a database search.

© 2010 Nature America, Inc. All rights reserved. npg

Box 1 Principles of shotgun (or discovery) proteomics

Discovery or shotgun proteomics is a universally and successfully used proteomic method that is used almost exclusively in a configuration in which an LC system is connected online to a tandem mass spectrometer operated in electrospray ionization mode (LC-MS/MS; **Fig. 2**). Less frequently, LC systems are used offline to deposit samples on a sample plate for analysis by a MALDI-MS (LC-MALDI-MS). In LC-MS/MS the peptide components present in the sample are separated by reversed phase liquid chromatography and analyzed by MS in the full-scan and MS/MS modes. The method is uniquely suited for the identification of the protein components of samples, including their post-translational modifications. If used with stable isotope-based labeling, it is also suitable for protein quantification.

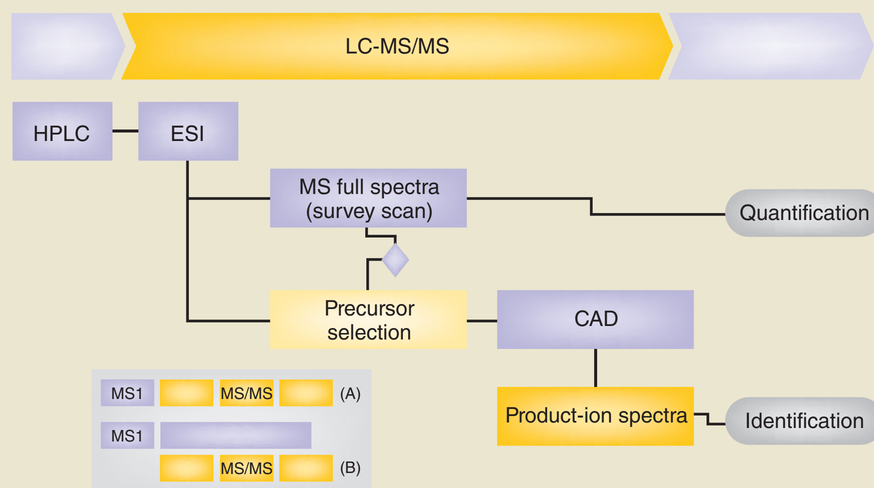

Figure 2 Workflow of a discovery proteomic experiment. The peptide mixture is separated by HPLC and analyzed by MS in full-scan mode. Using simple data-dependent acquisition heuristics based on signal intensity, peptide ions are selected for fragmentation and dissociated by collisional activation. The resulting MS/MS spectra permit determination of the amino acid sequence of the fragmented peptide. The intensity of the precursor-ion signal in the survey scan is used for quantification. The insert indicates the different modes of acquisition; either sequential MS and MS/MS analysis as performed using a quadrupole/time-of-flight instrument (A), or parallel analysis as performed on a linear ion trap/orbitrap mass spectrometer (B).

Instrumentation. The most common instrument types used are ion trap, hybrid quadrupole/TOF and hybrid ion-trap/orbitrap mass spectrometers. Most modern tandem mass spectrometers are compatible with the method, although their respective performances vary considerably.

Workflow. A proteolytic digest of the protein sample is analyzed by LC-MS/MS while the mass spectrometer is operated in DDA mode. In this mode, the system continuously acquires series of survey scans (MS1 mode) and a set of subordinated MS/MS scans, generating fragment-ion spectra of selected peptide ions. These fragment-ion spectra, combined with information on the precursor ions, are then analyzed to determine the amino acid sequence of the fragmented peptides, and to infer the proteins from which the peptides originate.

Survey scan. The survey scan is critical in a shotgun experiment. It detects the peptide ions that are selected on the fly for CAD, using a simple heuristic method. Typically, a subset of three to eight ions per survey scan is selected for fragmentation. The resolution and mass accuracy achieved in the survey scans affect the subsequent database search to assign the amino acid sequences to the generated fragment-ion spectra. The capacities of the FT-ICR, orbitrap and recently developed QTOF instruments for accurate mass measurement have considerably increased the confidence with which peptides can be identified.

MS/MS mode. Most experiments are performed using CAD. Two parameters affect the quality of MS/MS spectra, and thus the results of a shotgun measurement. The first is the mass window used for precursor-ion selection, that is, purity of the signal for sequencing. Typically, a broad window of 2–3 Th ensures sufficient sensitivity. The second is the analyzer performance in the MS/MS mode. Spectra obtained using ion-trap instruments are typically of low resolution and have limited mass accuracy (>0.1 Da), whereas TOF mass spectrometers and the orbitrap instruments provide high mass accuracy measurements for fragment ions. This facilitates the assignment of sequences to the spectra. Accurate mass determination of the precursor ion adds a discriminating constraint in sequence database searching.

Selection of precursor. Precursor-ion selection is performed automatically by the spectrometer on the fly, based on the information detected in the survey scan.

Quantification. Quantification is coupled to protein identification. Because quantification is performed on the 'sparse' survey scan, data precision is limited.

Informatics. All data processing and data analysis occurs after the completion of the mass spectrometric analysis. The tasks of assigning the correct peptide sequence to each acquired fragment-ion spectrum and of inferring the correct set of proteins represented by the identified peptides is computationally challenging and represents a large overhead, especially considering the volume of data acquired during shotgun experiments. This issue and the computational tools developed to address it have been reviewed recently[4,5].

Applications. The method is often used qualitatively, aiming at identifying large sets of proteins in complex samples. More recently, it has been used for differential quantification of the identified proteins. It is almost exclusively applied for discovery experiments. Because no prior knowledge is required, the method is ideally suited for open discovery experiments. The main limitation is its bias in the precursor selection process toward the more abundant component present in the sample, in particular for samples of very high complexity where the number of analytes exceeds the peak capacity of the LC-MS analytical system. It results in an irreproducible replication of the DDA experiment, as simple heuristics sample a different pool of peptides in each experiment[11,12].

© 2010 Nature America, Inc. All rights reserved.

detected in a survey scan, one survey scan can be followed by several product-ion scans. The instrument selects the specific precursor ions of each fragment-ion spectrum on the fly by DDA. State-of-the-art instruments permit data acquisition at a rate of a fraction of a second, enabling thousands of fragment-ion spectra to be collected during a typical reversed phase LC-MS/MS experiment. Although impressive,

Box 2 Principles of proteomic strategy based on directed MS

This method contrasts with the shotgun strategy in that protein identification (based on fragment-ion spectra) and protein quantification (based on survey scans) are decoupled and performed in two distinct experiments (**Fig. 3**). In fact, the two steps happen in the reverse order, and unlike shotgun proteomics, each sample is analyzed twice. A variant of directed sequencing, termed AIMS (accurate inclusion mass screening[22]), has been proposed to expedite the qualification of candidates and overcome some of the limits on an uncontrolled discovery experiment. LC-MALDI–based strategies also have the capability of performing inclusion list–driven peptide identification.

Instrumentation. This type of experiment is typically performed on high performance instruments, such as QTOF or LIT-OT instruments, to leverage their high mass resolution and mass accuracy capabilities.

Workflow. A directed MS experiment includes at least two LC-MS or LC-MS/MS analyses. The first is focused primarily on collecting survey scans, which are processed offline, to detect the features that will be selected for the inclusion list. This step creates an inventory of all detected peptide ions. This information is then used to design a second measurement of the same sample that aims at sequencing the analytes of interest, such as those that show differential expression between two conditions. The second LC-MS/MS run is performed in product-ion mode to generate tandem mass spectra used to identify specific targets listed on the inclusion list.

Survey scan. The survey scan remains mandatory in the second measurement, because the detection of a signal is required to trigger the MS/MS acquisition for an ion that is present in the inclusion list. As in the shotgun strategy, the resolution and accuracy of the survey scan are critical for the selection of the species of interest. The high mass accuracy of high-performance mass spectrometers, coupled with their low tolerance for the detection of the precursor ion needed to trigger an MS/MS event, ensure more effective exclusion of contaminant species that have similar *m*/*z* as the target peptide.

MS/MS mode. The MS/MS acquisition is performed in data-dependent mode, but the precursor mass selection takes into account the additional constraints of the inclusion list. To trigger a CAD event, an ion has to be observed in the survey scan with an intensity above a preset threshold, and it has to be present in the inclusion list.

Precursor selection. As in a shotgun experiment, a broad selection window ensures sensitivity. However, accurate masses are taken into account for the selection of the precursor and for database searching.

Quantification. As mentioned above, this method provides high quality LC-MS data and precise quantification, using the chromatographic dimension. The quantification of any analyte present in the sample is independent of the sequencing events. Therefore, differential analyses can be performed on all detected analytes. Even low-intensity signals in noisy survey spectra that would not be selected in a shotgun experiment can be detected and identified. The method is compatible with stable isotope–based and label-free quantification schemes.

Informatics. The database-searching overhead to perform peptide identification is substantially reduced as the redundancy of the acquired data decreases. There is, however, a large additional cost in processing the LC-MS data to detect and inventory all the ions and their attributes (mass, charge, elution time and signal intensity), and to possibly align and compare data from multiple measurements for the selection of the precursor set that constitutes the inclusion list. Several commercial and open source software tools for feature detection and alignment have recently been developed.

Applications. The method is primarily used in discovery experiments with an emphasis on less abundant species. Directed MS/MS approaches improve the efficiency of peptide identification in complex samples. This strategy has significant advantages over a conventional LC-MS/MS experiment in that the bias in favor of the most intense signals is partially removed, thus providing a deeper penetration into a proteome. In addition, decoupling the quantification and the identification steps provides more reliable quantitative measurements than can be accomplished in shotgun experiments. Triggering an MS/MS acquisition is contingent on the presence of signals corresponding to the peptide of interest in the survey spectrum. Nonetheless, the inclusion list allows the experiment to be tailored toward a specific set of ions.

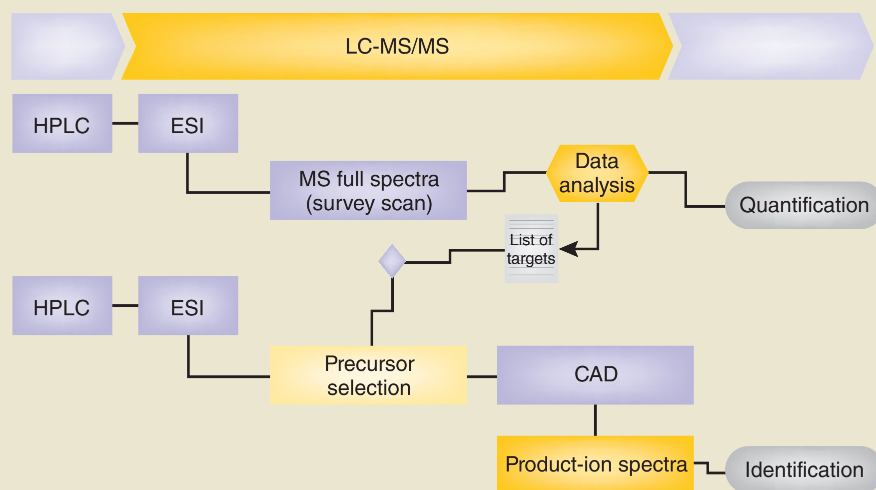

Figure 3 Workflow of a directed proteomic experiment. The sample is first analyzed in LC-MS mode, and the results are analyzed using a suite of bioinformatic tools to quantify the peptides. Typically, peptides that are of particular interest (e.g., those that are regulated by comparing multiple samples) are included in a list of targets for MS/MS sequencing. In a second step, the sample is reanalyzed to sequence exclusively the peptide ions present on the target list. The resulting MS/MS spectra enable the amino acid sequence to be determined.

© 2010 Nature America, Inc. All rights reserved. npg

this number is small in relation to the number of peptides generated by tryptic digestion of a proteome. The substantial discrepancy between the number of peptides present in a digest of a proteome and the analytical capacity of the LC-MS/MS analytical system (that is, the number of components that can be separated, detected and identified) prevents a perfectly reproducible set of peptides from being identified in repeat analyses of the same sample. This arises because a different subset of the available precursor ions is sampled in each subsequent analysis. Proteome coverage and data reproducibility can be improved by increasing the fraction of available precursor ions selected for CAD. This can be accomplished by repeated analysis of the same sample or fractionating the sample for subsequent analysis of each fraction[9,10,14]. With extensive sample prefractionation and the LC-MS/MS analysis of tens to hundreds of fractions per sample, the fraction of a proteome identified can be increased, presumably along with the reproducibility of the proteome patterns generated. These gains are, however, offset by the cost and time required to carry out such extensive proteome discovery experiments.

Developments of MS instrumentation and software engineering have enabled substantial advances in shotgun proteomics over the past decade. Although initially performed on low-resolution ion-trap instruments, the technique is now commonly implemented on last-generation, high-performance, hybrid mass spectrometers (e.g., linear ion trap orbitrap (LIT-OT) or quadrupole time of flight (Q-TOF)

Box 3 Principles of proteomics based on targeted MS

This technique distinguishes itself from shotgun or directed MS in that it uses prior information to generate validated mass spectrometric assays for the detection and quantification of predetermined analytes in complex samples (**Fig. 4**). It is most frequently implemented on triple quadrupole instruments operated in the selected reaction monitoring mode (SRM, often also called MRM).

Instrumentation. This type of experiment is preformed on triple quadruple instruments in which the second analyzer (third quadrupole) is used in nonscanning mode, which concentrates the available measurement time on the targeted analytes. This signal accumulation translates into an improved limit of detection.

Workflow. The method is exclusively hypothesis driven, that is, it requires a priori information at the level of both assay design and target selection. For each peptide, the *m/z* of the precursor ion, its retention time and a set of high-intensity fragment ions unique to the targeted peptide need to be defined, and these values constitute a definitive assay for the detection of the targeted peptide in any sample. The generation of validated SRM assays can be performed at high throughput through the use of synthetic peptide libraries[34].

Survey scan. No survey scan is performed in this mode.

MS/MS mode. As the SRM method is characterized by the measurement of only a few fragment ions of each targeted peptide, the second analyzer will 'jump' to a set of preset values, rather than scan across the entire *m/z* range. The parameters required for each measurement (precursor and fragment-ion *m/z* values, collision energy, elution time, dwell time per transition) have to be defined in the analytical method uploaded to the instrument.

Selection of precursor. As the precursor ions are monitored by default, regardless of their presence in the sample or their detection as a precursor ion, the method is not data dependent. Because of its intrinsically improved limit of detection, lower-mass selection windows (≥1 Th) can be used. This substantially reduces co-eluting interferences, thus increasing the overall selectivity.

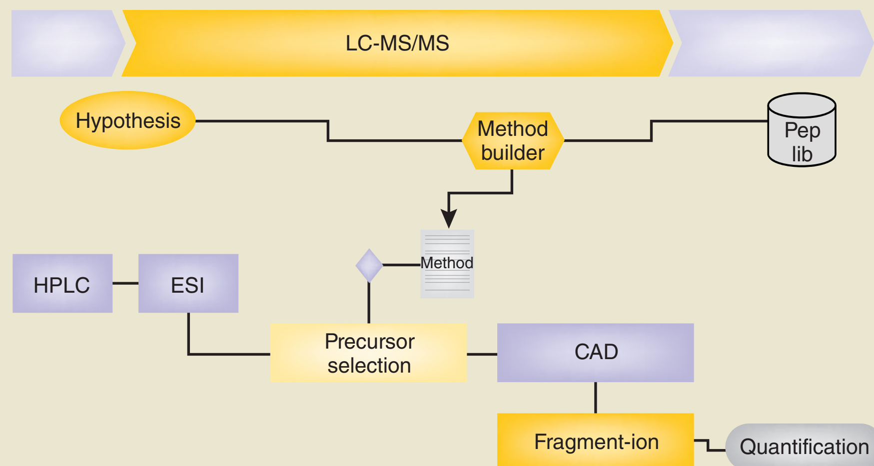

Figure 4 Workflow of a targeted proteomic experiment. As the experiment is hypothesis-driven, it targets a very specific subset of peptides uniquely associated with the proteins of interest. An instrument method is built using existing proteomic resources (peptide spectral libraries) required for a target analysis and is typically performed using a triple-quadrupole instrument. For each peptide, a series of transitions (pairs of precursor and fragment ion *m/z* values) are monitored during a time that specifically corresponds with its predicted elution time. This enables hundreds of peptides to be analyzed in a single experiment.

Quantification. SRM is the prototypical mass spectrometric quantification method, yielding precise measurements, with very low coefficients of variation and high reproducibility[28]. The limits of detection and quantification are typically two orders of magnitude lower than in conventional LC-MS experiments, especially if complex samples are being analyzed.

Informatics. Most of the informatics effort is performed upfront. In essence, SRM exploits existing information from the proteomics databases, such as specific SRM assays stored in either MRM-Atlas[25], or previous discovery information present in a repository such as PeptideAtlas (www.peptideatlas.org/).

Applications. The technique is exclusively hypothesis driven. It is focused on the detection and quantification of peptide candidates that are explicitly included in the experiment. The identity of the analytes relies on the elution time, and sometimes isotopically labeled internal standards are used for accurate quantification and for gaining confidence in the detected transition traces. Developments in instrument-control software that schedule the measurement of targeted peptides in predetermined time windows allow >1,000 transitions to be analyzed in a single LC-MS experiment, without compromising sensitivity[25].

© 2010 Nature America, Inc. All rights reserved.

Table 1 Mass analyzers commonly used in proteomics

Analyzer	Implementation	Type	Resolving power	Mass accuracy	Limit of detection	Dynamic range
Quadrupole	TQ-QTOF	In-beam	1,000–2,000	Low	Very low	4–5
Ion trap	IT	Trapping	1,000–2,000	Low	Very low	2–3
TOF	Q-TOF	In-beam	>25,000	High	Low	3
OT/ICR	Hybrid	Trapping	>50,000	Very high	Low	3

TQ, triple quadrupole.

instruments), resulting in dramatically increased data quality and faster rates of data acquisition. Recent studies have demonstrated dramatic increases in the proteome coverage achieved and the ability to identify large numbers of modified peptides[8,9]. Furthermore, the recent implementation of alternative fragmentation techniques, such as electron transfer dissociation[18], has further increased the range of peptide analytes accessible to mass spectrometric analysis. Specifically, large peptides and peptides subject to post-translational modifications show favorable electron transfer dissociation fragmentation patterns[19]. Therefore, shotgun proteomics is the method of choice for the a priori identification of the protein components of complex samples and the characterization of their post-translational modifications.

Directed proteomics

The hallmark of directed MS is the selection and fragmentation of a predetermined set of peptide ions detected in a survey scan[20–22]. The principles and information pertinent to directed MS are summarized in **Box 2** (see also **Fig. 3**). In this method, the precursor ions that are of interest for a particular study (e.g., peptides that are differentially expressed between samples) are compiled into a master list, along with relevant attributes such as the precursor-ion charge state, m/z ratio and retention time. This list is the basis for the generation of one or several inclusion lists that are loaded into the computer controlling the mass spectrometer to ensure that the instrument exclusively selects for CAD those features that are detected in a survey scan and are present on the inclusion list. Selection of multiple precursors from a survey scan and tight scheduling of retention times have now increased the number of precursors selected in a 60 min or 90 min LC-MS/MS run to several thousand. Because the generation of the master list and its use for measurements are uncoupled in time, feature selection can be optimized according to the quality of the sample and the biological question at hand. A variant of the approach, referred to as LC-MALDI, involves spotting the column effluent on the solid surface of a sample plate and then sampling the contents of sequential spots by MALDI-MS/MS.

Different types of input data have been used to compile master lists (R.A. and colleagues[23]). They include, for example, prior quantitative proteome measurements by differential stable isotope labeling or by comparative analysis of LC-MS feature maps generated from different samples. Compared with a discovery proteomic experiment using DDA, precursor ions of lower abundance can be selected, especially if highly complex samples are being analyzed and the identification rate is increased. Selection of the same set of precursor ions for fragmentation in repeat analyses of the same or substantially similar samples increases reproducibility between data sets. Finally, peptides with detectable features, such as distinctive isotopic signatures or mass defects, or peptide patterns indicating structurally related peptides (e.g., differentially modified peptides) can be detected in LC-MS feature maps and specifically selected for analysis in subsequent LC-MS/MS runs driven by inclusion lists.

Targeted proteomics

The hallmark of targeted MS is the detection of a set of predetermined fragment ions from precursor ions that are anticipated, but not necessarily detected, in a survey scan. Currently, the main implementation of this concept is selected reaction monitoring (SRM) using triple quadrupole instruments. SRM is a quantitatively accurate technique that has been well established in small-molecule MS[24]. The principles and information pertinent to targeted MS are summarized in **Box 3** (see also **Fig. 4**). In this approach, the fragment-ion spectrum of the targeted peptide is determined in prior measurements. The precursor-ion mass, the charge state, elution time and characteristic high-intensity fragment ions represent a definitive assay for the targeted peptide used to detect and quantify the targeted peptide in a sample. The relationship between a precursor ion and a specific fragment ion is referred to as a transition. Quantification is accomplished by relating the fragment-ion intensities of the targeted peptide to the corresponding signals of isotopically labeled reference peptides of identical sequence. If the elution times of the targeted peptides are used as a measurement constraint (that is, specific subsets of the targeted peptide are only detected in a narrow time window spanning a few minutes around their anticipated elution time), several hundred peptides can be targeted in a single LC-MS/MS analysis[25].

The precursor ion of the targeted peptide does not need to be explicitly detected within the matrix of the sample, and background noise is filtered out sequentially at the precursor- and fragment-ion levels. These considerations make targeted MS the most sensitive mass spectrometric strategy and the one least affected by interference effects when analyzing complex samples. The optimal transitions (precursor- and fragment-ion pairs), retention time and collision energy that constitute a definitive assay need to be established once for a particular instrument type and can then be used perpetually. They can therefore be made accessible in public databases[26].

Implementation of MS strategies

Each of the three strategies we have described relies on tandem MS. Each presents unique characteristics that determine its suitability for tackling a specific proteomic or biological research question. The strategies also differ in the way the mass spectrometers are used. The types of mass spectrometers commonly used in proteomics, along with some of their distinctive traits, are summarized in **Table 1**. The instrument characteristics pertinent to proteomics are the selectivity of measurement to avoid cross-talk from other analytes (resolving power), the linear dynamic range, the limits of detection and quantification and the mass accuracy (**Box 4** and **Fig. 5**).

Shotgun proteomics depends on the ability of the instrument to reliably detect precursor ions in a survey scan, to select an optimal set of detected precursor-ion signals for CAD and to generate and acquire fragment-ion spectra with ion series sufficient for the unambiguous assignment of the correct peptide sequence to the spectrum. Additionally, these operations should be carried out at a high cycle frequency to maximize the number of peptide identifications, and the measurements should have high sensitivity, large dynamic range and high mass accuracy. These requirements are best matched by ion trap hybrid instruments such as ion trap–Fourier transform ion cyclotron resonance (FT-ICR) and ion trap–orbitrap, and Q-TOF instruments, respectively. Currently, shotgun proteomic measurements are most frequently carried out using LIT-OT instruments.

© 2010 Nature America, Inc. All rights reserved. npg

The main difference between shotgun and directed sequencing experiments is the method used to select precursor ions detected in survey scans for CAD. Although this process is instrument driven, it is controlled by a time-constrained inclusion list in the directed method and is no longer intensity-dependent. Therefore, shotgun and directed sequencing differ at the level of instrument control rather than at the level of the instrument type, and the same considerations related to instrument performance and characteristics apply to both methods.

Targeted experiments, which are based on SRM (**Box 4**), depend on the effective and sequential filtering of noise at the precursor-ion and fragment-ion level, which increases the signal-to-noise ratio and therefore the limit of detection. Targeted strategies are characterized by a dynamic range of concentrations spanning four to five orders of magnitude, high sensitivity and a relatively small number of analytes detected per unit time. To achieve precise quantification, measurements need to be performed to ensure that enough data points are acquired over the chromatographic elution range of a peptide to reconstruct the chromatographic peak. This limits the number of peptides detected per unit time. For instance, at a 2-s cycle time, 100 transitions using a 20-ms dwell time for each measurement would be acquired. Presently, the characteristics for SRM can be fulfilled only by triple quadrupole mass spectrometers. An interesting variant, useful for the development of SRM assays, is the capability of acquiring full fragment-ion spectra driven by an SRM transition[25]. An advantage of quadrupole/linear ion trap instruments

Box 4 Key considerations when planning quantitative proteomics experiments

When conducting any proteomics experiment, several factors are key to the characterization of MS measurements. These are summarized in **Figure 5** and described below.

Selectivity. The selectivity of a method is its ability to discriminate and quantify a particular analyte in a mixture or matrix without interferences from the other components. The reliability of measurements depends on the selectivity of the analytical device. An increased selectivity is achieved by analyzers with higher resolving power, which separate near-isobaric ionic species and determine their respective accurate mass. High selectivity is particularly critical in the LC-MS analysis of complex mixtures, in which multiple components co-elute from the column. Analyzers such as FT-ICR, orbitrap or the last-generation TOF analyzers present high-resolution capabilities and thus increased selectivity. Alternatively, the selectivity of quantitative analyses can be improved by using a second level of mass selection, as in the SRM mode.

Limit of detection (LOD). The intrinsic LOD of an instrument or a method, which is often incorrectly called the sensitivity, is defined as the minimal quantity of an analyte that can be confidently detected. The related term, limit of quantification (LOQ), is defined as the minimal amount of an analyte that can be confidently quantified (**Box 5**). The instrument LOD is usually specified by measuring the components of a simple mixture or individual analytes in dilution series. In such samples, the chemical background is minimal. The limit of detection and dynamic range, which are pertinent in the context of complex biological samples, are modulated by the background and the interferences associated with it. The components of a complex sample will affect the detected signal-to-noise ratio and may affect the ionization efficiency through suppression effects. Under the conditions encountered when using biological samples, the chemical background is significant and poor signal-to-noise ratios are observed for analytes present at very low concentrations. Although state-of-the-art instruments have LODs and LOQs for single compounds or simple mixtures in the low amol range, matrix and ion-suppression effects considerably reduce the practical ability to detect species of low abundance in complex samples, especially in cases in which the respective precursor-ion signal needs to be detected in a survey scan. Thus, the sample preparation (that is, reduction of the sample complexity) cannot be dissociated from the entire analytical protocol.

Dynamic range. The dynamic range of an instrument denotes the range between the highest signal and the lowest amount of an analyte detected in a single analysis. Often, the linear range of the response is also specified. The dynamic range is determined by performing dilution series of specific analytes, either by themselves, or added to a matrix. The highest dynamic range is currently obtained on in-beam instruments such as quadrupoles, where ions are continuously monitored. Sample overloading is possible in such instruments. This leads to saturation of the major components, whereas minor species emerge from the background. In-beam systems are often preferred for quantitative analyses over trapping devices. Matrix and ion suppression effects occur if multiple components eluting concurrently from the high-performance (HP)LC column are ionized together. As mentioned above, complexity of the samples and chemical background affect the dynamic range, in particular for trapping devices.

Data density. The data density is defined as the number of measurements acquired during one experiment. In a conventional shotgun experiment, the value indicates the number of MS/MS sequencing events. In a targeted experiment, it reflects the number of peptides analyzed, including multiple measurements for each peptide. Obviously, the volume of data acquired is closely related to the sensitivity and the acquisition rate of the instrument.

Repeatability. The repeatability of a measurement refers to the ability of the method to generate identical results if identical test samples are processed with the same procedure under the same conditions (instrument settings, operator, apparatus and laboratory) within a short interval of time.

Reproducibility. The reproducibility of a method refers to the ability to replicate the measurement accurately by someone else working independently, that is, the ability to generate identical results obtained with the same method on identical test material, but under different conditions (different operators, different apparatus, different laboratories and/or after different intervals of time).

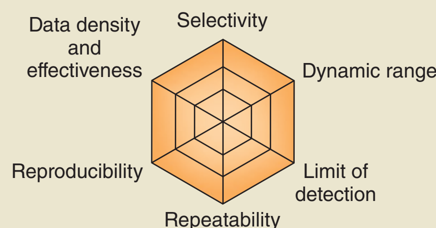

Figure 5 A representation of the desired characteristics of a proteomic experiment. The actual performances for each of the approaches can be compared visually by representing the individual characteristics on each of the six axes (**Fig. 6**).

© 2010 Nature America, Inc. All rights reserved. npg

is that they can be operated alternatively in triple quadrupole and LIT operating mode to acquire MS/MS spectra.

In summary, proteomics researchers have yet to develop the ideal universal mass spectrometer for proteomics. The type of experiment performed and the method chosen for data acquisition determine the optimal type of instrument for each application. Moreover, every instrument and data acquisition mode presents a series of compromises that affects the performance of a given proteomic strategy.

Performance profiles of the three strategies

There is currently no single method capable of routinely identifying and quantifying all the components of a proteome. Each method is therefore a compromise that maximizes the performance at some levels, while reducing it at others. For example, in a SRM-based targeting experiment, the recorded signal-to-noise ratio is related to the dwell time (that is, the time the spectrometer takes to record the signal of a given transition). The lower limit of detection achieved by longer dwell times negatively affects the number of transitions and therefore the number of peptides that can be analyzed during a time segment. Similarly, an increase of the resolving power of a quadrupole mass analyzer reduces sensitivity. As another example, in quantitative shotgun proteomics in trapping instruments, the limit of detection for precursor ions and therefore the quantitative accuracy achieved, depends on the trapping time. Longer trapping times improve the limit of detection but reduce the number of different analytes measured per unit time. Furthermore, many of the performance characteristics depend on the source of the sample and its complexity. For example, the shotgun and directed MS methods, where the precursor ion has to be explicitly detected in a matrix of background ions before selection for CAD, are more strongly affected by background noise than the targeted methods where the precursor ion does not need to be explicitly detected. A comprehensive discussion of the benefits and trade-offs of each strategy is beyond the scope of this account. We therefore summarize the trade-offs inherent to each method with respect to the main factors characterizing proteomic measurements: selectivity, dynamic range, limit of detection, repeatability, reproducibility, data density and effectiveness. These terms are defined in **Box 4** and the performance characteristics of each method are summarized in **Figure 6**. The above discussion of a few of the trade-offs that apply to proteomic measurements already suggests that there is no single best implementation of a particular strategy. The performance profiles discussed below therefore apply to implementation parameters that are commonly applied in proteomics.

Performance profile of shotgun proteomics

Shotgun MS typically involves using a hybrid mass spectrometer with a fast cycle time to analyze complex sample mixtures comprising potentially hundreds of thousands of peptides with abundances that span up to ten orders of magnitude. The combination of intensity-based heuristics for precursor-ion selection, limited cycle speed, high sample complexity and lack of input of prior data for precursor selection contribute to the performance profile indicated in **Figure 6a**.

The high acquisition frequency (1–10 Hz range) of modern spectrometers ensures that shotgun measurements produce a high data density (**Box 4**). Even so, extensive proteome coverage can be achieved only if the samples are fractionated before MS analysis and the individual fractions are sequentially analyzed. This is because the precursors are selected based on their signal intensity, and, even in the fastest available instruments, the number of precursors in a proteome digest exceeds the number of sequencing events available in a LC-MS/MS run. Multiple, repeated selection of the same precursor in the same or sequential fractions results in the redundant identification of the same peptides and

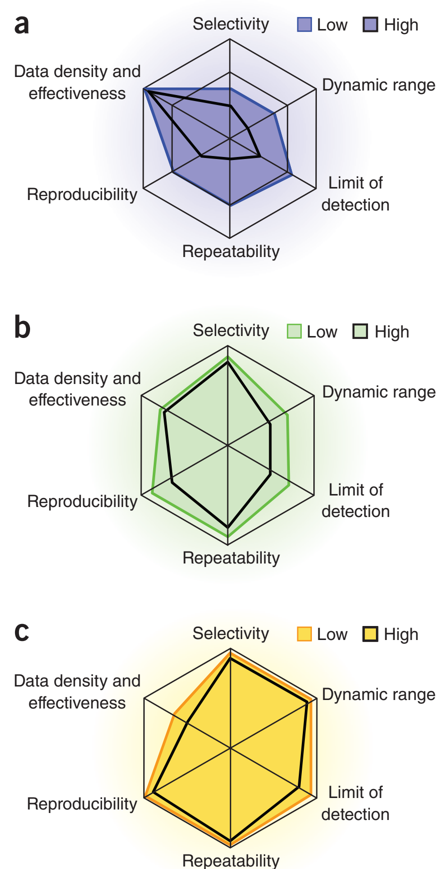

Figure 6 Performance profiles of the shotgun or discovery (**a**), directed (**b**) and targeted (**c**) proteomic methods. The characteristics are defined and discussed in **Box 4**. The terms 'high' and 'low' refer to sample complexity.

proteins. This also reduces the yield of newly identified peptides and proteins and limits the repeatability of the results from replicate analyses of identical or substantially similar samples, especially for proteins of lower abundance.

Another striking feature of the performance profile for shotgun proteomics is the strong dependence of most parameters on sample complexity. In particular, the limit of detection, the dynamic range and sample reproducibility—three of the most critical parameters for proteome analysis—are negatively affected by increasing sample complexity. These considerations significantly affect the experimental strategy of shotgun proteomics, especially when repeat analyses of substantially similar samples are being analyzed, as is the case, for example, in clinical, time-series or dose-response studies.

In conclusion, the shotgun proteomic strategy has a unique potential to discover new proteins and to determine relative protein abundance of proteins identified in different samples. However, the extensive or complete analysis of complex samples, such as those representing whole proteomes, comes at a very high cost in measurement and computational time. Moreover, the performance of the method may vary substantially between samples. Therefore, the shotgun strategy is most frequently applied in cases when samples of unknown composition are being analyzed to identify the largest number of proteins possible; shotgun proteomics is a uniquely powerful method to generate protein inventories. If combined with stable isotope labeling, shotgun proteomics is also commonly used for quantitative comparison of related subsets of the proteins in complex samples. The factors discussed above limit the number of samples that can

© 2010 Nature America, Inc. All rights reserved. npg

be compared, the number of proteins that can be consistently identified and quantified in multiple samples and the range of protein abundances that can be accommodated. Typical applications of this strategy include the quantitative comparison of the proteomes of differentially perturbed cells, the comparison of protein extracts from diseased and healthy tissues and the analysis of specific subproteomes. More recently, shotgun proteomics has been expanded to the systematic analysis of post-translational modifications, specifically protein phosphorylation[8,27]. The implementation of electron transfer dissociation is expected to further advance the potentials of discovery-based analyses, especially for modified subproteomes[19].

Performance profile of directed proteomics

Shotgun and directed MS measurements are usually performed using identical instruments. The two methods are essentially identical, except that in directed sequencing, precursor-ion selection no longer follows abundance-dependent heuristics, but is instead directed by a time-constrained inclusion list that is compiled based on prior information. It is apparent from **Figure 6b** that this seemingly simple difference has several important implications for the performance profile of the method. First, the high cycle time is maintained but the same precursor ion is analyzed with dramatically reduced redundancy (ideally once), even if multiple fractions are being analyzed. This significantly increases the repeatability and the reproducibility of the method. Second, the control of the sequencing events reduces the rate of futile repeated identifications, and the associated computational overhead for data analysis is reduced. Third, because the precursor-ion signal of the selected precursor still needs to be detected, the dynamic range and limit of detection of directed MS depend on the sample complexity, albeit less so than in shotgun methods. Finally, the overall dependency of the performance profile on sample complexity is reduced.

In summary, by virtue of its focus on sequencing only those peptides of particular interest, directed MS offers an effective way to characterize a proteome. It can also be used in a hypothesis-driven mode involving identification of a predetermined set of precursor ions in multiple samples[28]. The data sets generated by directed MS are generally of much higher information content and lower redundancy than those generated by DDA. It can be expected that the recent development of software tools to generate optimized inclusion lists will catalyze wider application of the method. Directed MS can be used for quantitative measurements in conjunction with stable isotope labeling or with label-free quantification, whereby peptide quantities are estimated from their precursor-ion current.

A wide range of typical applications of directed MS has been discussed recently[23]. They include the directed measurement and quantification of proteins that are of different abundance across samples, the directed measurement of modified peptides and the analysis of protein biomarkers in clinical samples. Because identical sets of peptides can be measured in multiple samples with a high degree of repeatability, the method used in the context of predefined peptide lists is well suited to generate reproducible, quantitative data sets.

Performance profile of targeted proteomics

Much like directed MS, the targeted method also depends on lists of peptides deemed important for detection and quantification in a sample based on prior information. However, in contrast to directed sequencing, the targeted precursor ions are not detected in a survey scan and a full fragment-ion spectrum of the selected precursor is not generated. Instead, the targeted precursor is selected 'blindly' in an anticipated chromatographic time window, and the only signals detected are fragment ions that are derived from the targeted peptide (transitions). In the targeted method, an initial effort is required to determine the optimum fragmentation conditions and, thus, generate optimized assays for each peptide. However, the benefits of this one-time investment are apparent from the performance profile of the method (**Fig. 6c**).

Several important features are readily apparent. First, the targeted method is less affected by sample complexity and background, as noise signals are filtered out both at the precursor level by a narrow (<1 Da) mass-selection window and at the fragment-ion level, where the targeted precursor and background ions are expected to lead to distinct fragment ions for detection and quantification. Second, of the available mass spectrometric methods, the targeted method has the lowest limit of detection and the widest dynamic range, especially for complex samples. This is a result of the nonscanning nature of fragment-ion signal acquisition, which allows the integration of the respective signal over extended periods (dwell time). Third, repeatability and reproducibility are excellent because of the nonredundant, targeted data acquisition[29]. Finally, the data density is lower than that of the shotgun or directed methods because of the increased time needed to measure each peptide to lower the limit of detection.

In targeted MS, accurate quantification is achievable by any of the commonly used stable isotope labeling techniques[4]. Of course, no new proteins are detected by the targeted method; the approach depends substantially on the prior measurement of the targeted proteins by discovery proteomics. This strategy therefore is an excellent choice for those studies in which sufficient prior information on a system has been acquired, and the types of research question have shifted from identifying the full complement of proteins associated with a process or location to issues related to characterizing coordinated changes in the abundances of these proteins. Targeted MS will therefore likely become a key technology to test biological hypotheses, reproducibly generate complete quantitative data sets for systems biology or validate biomarkers by scoring changes in their abundances in large set of clinical samples.

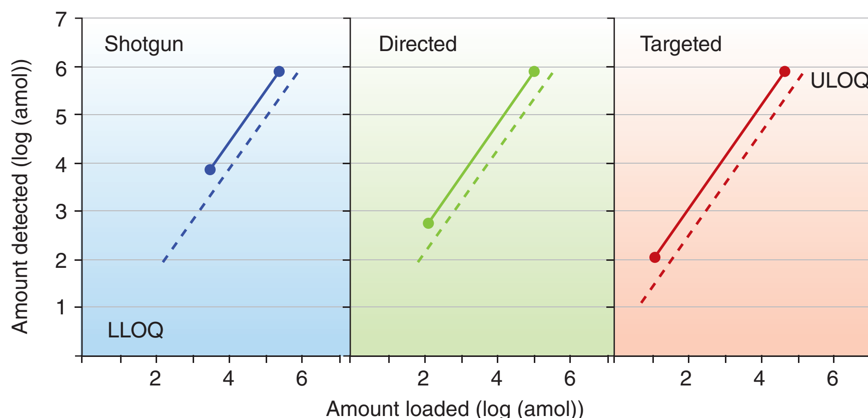

Figure 7 Effect on biochemical background on quantification by the shotgun (discovery), directed and targeted proteomics strategies. Whereas dotted lines indicate a low-complexity background, full lines represent a complex background, such as a full cell lysate. LLOQ, lower limit of quantification; ULOQ, upper limit of quantification.

Differential and quantitative analyses

The term quantitative proteomics usually refers to measuring the changes in the level of abundance of proteins in different samples. Typical studies include the quantitative comparison

© 2010 Nature America, Inc. All rights reserved. npg

of samples from 'different biological conditions,' with the underlying assumption that the proteins showing different abundance are functionally related to the processes affected by the applied conditions. Commonly, comparative studies use isotopic-labeling approaches such as isotope-coded affinity tagging (ICAT), isotope-coded protein labeling or stable isotope labeling with amino acids in cell culture (SILAC)[2,4]. All involve labeling the peptides in a sample before the LC-MS analysis with different reagents or labels that are chemically identical but differ in their isotope composition. The relative abundance of a specific peptide across the samples tested is then computed from the precursor-ion signals of the heavy and light forms of the peptide, respectively. It should be noted that other quantification methods, based on isobaric tagging reagents, or tandem mass tags, exemplified by isobaric tagging for relative and absolute quantification (iTRAQ)[30], are also used for quantitative proteomics. In these methods, the determination of the relative abundance of peptides is performed on reporter fragment ions measured in the MS/MS mode. In such studies, four to eight samples are compared and analyzed concurrently.

Such relative abundance measurements, often also erroneously referred to as being semiquantitative, contrast with the definition of quantification used in analytical chemistry. This denotes the precise determination of the concentration of specific analytes present in the sample with a coefficient of variation typically <20%. Such measurements require calibrants (internal standards) and have to be performed in the linear dynamic range of the analytical system[25]. Quantification performed by mass spectrometric techniques is usually performed by stable isotope dilution, that is, by adding the analyte of interest in which some stable isotopes have been incorporated. Internal standards, prepared by incorporation of stable isotopes, such as ^{13}C, ^{15}N and ^{18}O, are most commonly used in proteomics. The use of deuterium is less desirable as it changes the peptides' physicochemical properties such that the corresponding heavy and light compounds no longer co-elute under reversed phase conditions. Dilution series of a limited number of the reference compounds are usually measured to ensure that measurements are performed within the linear dynamic range.

Differential analysis and precise quantification differ in their scope, the experimental design and the platform used for such studies. Whereas comparative studies often deal with a limited number of samples (typically no more than a dozen), quantitative studies may include hundred of samples. Quantitative analysis thus requires a rugged high-throughput platform to generate reproducible data sets. Differential analyses can in principle be performed using any platform, whereas precise quantification is routinely performed on triple quadrupole instruments in the targeted mode. The design of quantitative experiments

Box 5 Limit of detection and dynamic range

The limit of detection is defined, in a first approximation, as three times the signal-to-noise ratio. Correspondingly, the limit of quantification (LOQ) is defined as nine times the signal-to-noise ratio. The LOD/LOQ value of a measurement and its dynamic range are ultimately determined by the nature and the complexity of the analyzed sample (**Fig. 8**).

Although it is relatively straightforward to detect low concentrations of peptides in simple mixtures (not taking into account possible losses during the sample handling and the HPLC separation), the LOD of the same peptide in very complex samples is considerably higher, often by several orders of magnitude. The main factor for this apparent sample dependency of the achieved limit of detection is the low ratio of analyte to the total amount of peptide. It results in the co-elution of chemically similar species which, in turn, can cause ion suppression during the ionization process.

In addition, some mass analyzers, more specifically the trapping devices, show a decreased signal if the analyte in question is in a complex sample compared to the signal recorded if the same nominal amount of a pure analyte is injected. This is due to the limited total number of ions that can be stored in the trap without affecting its performance by space charging. In pure samples, a large fraction of the available ion capacity is occupied by the ion in question, whereas in complex samples the majority of the available ion capacity can be occupied by ions representing background signals.

In contrast, in-beam analyzers such as quadrupoles can handle large ion fluxes. Even if the most abundant components saturate the detector and can no longer be quantified accurately, low abundance species can be detected as long as their signal exceeds the signal-to-noise ratio of the background. These dependencies are schematically illustrated in **Figure 8**.

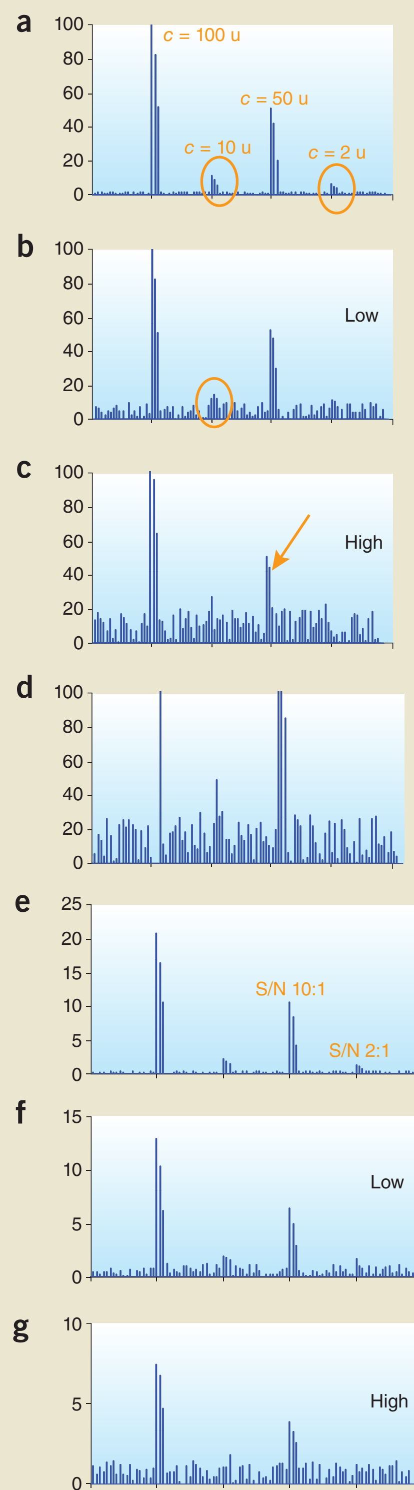

Figure 8 Effect of background on the detection of a peptide mixture in various concentrations (arbitrary units). (**a**) In-beam, no background; (**b**) low/moderate background; (**c**) high background; (**d**) increased amount of sample loaded; (**e**) trapping instrument, no background; (**f**) moderate; (**g**) high background; note the changes in the ordinate as ion counts decrease. *c*, concentration; u, arbitrary units; S/N, signal-to-noise; the *x* and *y* axes correspond to m/z values and signal intensities, respectively.

© 2010 Nature America, Inc. All rights reserved.

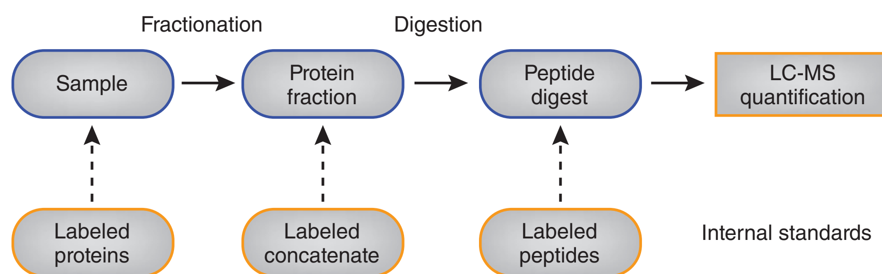

Figure 9 Isotopically labeled internal standards can be added at various stages in quantitative proteomics experiments. Full-length reference proteins are added at the beginning of analyses, concatenated peptides are added prior to digestion and synthetic reference peptides are added prior to the LC-MS analysis.

needs to consider the sample preparation and the mass spectrometric measurements. A detailed discussion of the different isotope labeling methods is beyond the scope of this account and can be found elsewhere[4]. Overall, quantification is based on the first principle claiming a direct relationship between signal measured and amount of analyte present in a sample (Signal = $F \times$ Amount, where F is the response factor specific to each analyte).

Quantification in shotgun proteomics

Historically, shotgun proteomics has focused on identifying a large set of peptides and proteins. In this method, only precursor signals assigned to a sequence are quantified, it has generally been used less frequently for protein quantification than for protein identification. Consequently, in such measurements, the majority of the data acquisition time is spent on analytes of unchanged abundance that are likely of no relevance to the biological question of the study. The focus has been on detecting relative changes in the concentrations of peptides (and indirectly the associated proteins) across samples. Such experiments, by analogy to genomics arrays, typically focus on changes of at least twofold and usually analyze all signals observed in an LC-MS experiment. The objective of such measurements is to maximize the number of sample components measured. Because relative changes are of primary interest, the precision of measurements becomes of secondary importance. There is an obvious trade-off between the comprehensiveness and the precision of the measurements.

Differential shotgun proteomics is a typical discovery technology. In addition to the technical limitations discussed above, the user is routinely faced with the problem of assigning biological significance and meaning to hundreds of proteins whose abundance is regulated. In principle, such comparative analyses can be performed on any platform capable of LC-MS analysis. However, the overall analytical precision depends heavily on the choice of the instruments and the characteristics of the analyzer in terms of resolving power, mass accuracy, limit of detection and dynamic range. For instance, ion-trap instruments commonly used in qualitative proteomics experiments have limited resolving power and dynamic range. Other platforms with increased resolving power or dynamic range might be more suited for such experiments (**Table 1**). Furthermore, as illustrated in **Figure 7**, the biochemical background inherent to complex samples greatly affects performance. Moreover, the lower limits of detection or quantification are significantly compromised in complex samples (**Box 5** and **Fig. 8**).

Quantification in directed proteomics

In contrast to shotgun proteomics, which is used exclusively as a discovery method and where only the identified analytes are subject to comparative quantification, the directed MS method offers an interesting range of strategic possibilities. The method can be used in a discovery mode, whereby the differentially abundant peptides can be detected first and sequenced second, typically in sequential LC-MS runs. This simple reversal of steps compared to the shotgun method has profound implications: the available sequencing cycles can be focused on the differentially abundant peptides and wider array of comparative quantification strategies can be used to initially identify differentially abundant analytes that are then subsequently subjected to directed MS/MS sequencing. The directed approach is also compatible with accurate quantification by means of stable isotope dilution, using isotopically labeled reference compounds. Quantification is based on the comparison of the precursor-ion currents of the heavy and light forms of the analytes, and quantitative accuracy might be compromised by contaminant signals interfering with either or both isotopic forms. The potential for such interferences increases with increasing sample complexity.

In addition, apart from the differential isotope labeling methods commonly used for shotgun comparative studies, differentially abundant peptides are also frequently detected by comparative precursor-ion pattern analysis without isotope labeling[5,23]. In such cases, also referred to as label-free quantification, the different samples are analyzed sequentially under rigorously controlled instrument conditions. The signal intensity of each of the peptides is then used to assess the amount present in each sample. If advanced software tools for pattern comparison are used, the precursor-ion patterns of many samples can be overlaid, compared and statistically analyzed for the selection of the most significant differentially abundant species. There are underlying assumptions or approximations in this approach that, if violated, impinge on the precision of the measurements. More specifically, it is assumed that the response factor is not affected by the 'micro-environment'; in other words, ionization suppression or enhancement effects are negligible. Furthermore, the amount of material injected in each experiment has to be carefully controlled, or a correction factor has to be applied after the analyses.

Quantification in targeted proteomics

Targeted proteomics can be used for comparative quantitative analysis or for accurate, absolute quantification of the targeted peptides. Owing to its exquisite selectivity and very low limit of detection, SRM has *de facto* become the reference method for quantification in complex samples. As discussed above, quantitative, targeted MS is most frequently used in studies in which the same, predefined set of proteins is quantified in multiple samples. Accurate quantification depends on the addition of isotope-labeled reference molecules to the samples[31–33]. In a proteomic experiment, internal standards can be added at various stages, from the crude protein isolate such as cell lysate or plasma sample to the fractionated peptide sample immediately before injection into the LC-MS system (**Fig. 9**). By adding the reference samples at the step closest to the origin of the biological sample, results of higher precision are generated because progressive losses and variability induced by sample processing are compensated for.

Conclusions

The development and commercialization of better mass spectrometer software tools for data analysis have driven tremendous advances in proteomics over the past decade. This progress has translated into larger and more reliable data sets, mostly generated using the shotgun

© 2010 Nature America, Inc. All rights reserved.

(or discovery) approach. Concurrently, new proteomic strategies have emerged that are accurately quantitative, support rigorous testing of biological hypotheses and show improved reproducibility. The advent of these new strategies has been primarily driven by the need for comprehensive and reproducible data sets for applications such as biomarker validation or modeling processes of interest to systems biologists. In either case, large sets of peptides (used as surrogate for the proteins of interest) have to be analyzed precisely and reliably in each of many samples in the study.

Like any other analytical process, the three proteomic strategies discussed above have limitations that set the boundaries of their respective performance and define the biological or biomedical research questions that best match the performance profile of each method. It can be expected that the emergence of approaches, such as directed and targeted MS, which are built on the use of often vast amounts of prior knowledge, will increase the impact of proteomics in biomedical research. These techniques will increasingly augment more common types of experimentation, especially as they provide the capacity of generating data sets that can be compared across studies and laboratories[29], and because quantitative proteomics data are generated with unprecedented sensitivity, accuracy and reproducibility.

ACKNOWLEDGMENTS
We thank P. Picotti, A. Schmidt and N. Selevsek for the preparation of figures. The Swiss National Science Foundation (grant no. 3100AO-107679), SystemsX.ch, the Swiss initiative for systems biology and the European Research Council are acknowledged for financial support.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Published online at http://www.nature.com/naturebiotechnology/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Schrimpf, S.P. *et al.* Comparative functional analysis of the *Caenorhabditis elegans* and *Drosophila melanogaster* proteomes. *PLoS Biol.* **7**, e48 (2009).
2. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).
3. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science* **312**, 212–217 (2006).
4. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* **389**, 1017–1031 (2007).
5. Listgarten, J. & Emili, A. Statistical and computational methods for comparative proteomic profiling using liquid chromatography-tandem mass spectrometry. *Mol. Cell. Proteomics* **4**, 419–434 (2005).
6. Shevchenko, A., Loboda, A., Ens, W. & Standing, K.G. MALDI quadrupole time-of-flight mass spectrometry: a powerful tool for proteomic research. *Anal. Chem.* **72**, 2132–2141 (2000).
7. Medzihradszky, K.F. *et al.* The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal. Chem.* **72**, 552–558 (2000).
8. Beausoleil, S.A., Villen, J., Gerber, S.A., Rush, J. & Gygi, S.P. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* **24**, 1285–1292 (2006).
9. de Godoy, L.M. *et al.* Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455**, 1251–1254 (2008).
10. Denny, P. *et al.* The proteomes of human parotid and submandibular/sublingual gland salivas collected as the ductal secretions. *J. Proteome Res.* **7**, 1994–2006 (2008).
11. Parag Mallick, P. & Bernhard Kuster, B. *Nat. Biotechnol.* **28**, 695–709 (2010).
12. Tabb, D.L. *et al.* Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **9**, 761–776 (2010).
13. Paulovich, A.G. *et al.* Interlaboratory study characterizing a yeast performance standard for benchmarking LC-MS platform performance. *Mol. Cell. Proteomics* **9**, 242–254 (2010).
14. Washburn, M.P., Wolters, D. & Yates, J.R. III. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19**, 242–247 (2001).
15. Bell, A.W. *et al.* A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. *Nat. Methods* **6**, 423–430 (2009).
16. Kuster, B., Schirle, M., Mallick, P. & Aebersold, R. Scoring proteomes with proteotypic peptide probes. *Nat. Rev. Mol. Cell Biol.* **6**, 577–583 (2005).
17 Duncan, M., Aebersold, R. & Caprioli, R. *Nat. Biotechnol.* **28**, 659–664 (2010).
18. Syka, J.E., Coon, J.J., Schroeder, M.J., Shabanowitz, J. & Hunt, D.F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. USA* **101**, 9528–9533 (2004).
19. Coon, J.J. Collisions or electrons? Protein sequence analysis in the 21st century. *Anal. Chem.* **81**, 3208–3215 (2009).
20. Schmidt, A. *et al.* An integrated, directed mass spectrometric approach for in-depth characterization of complex peptide mixtures. *Mol. Cell. Proteomics* **7**, 2138–2150 (2008).
21. Domon, B. & Broder, S. Implications of new proteomics strategies for biology and medicine. *J. Proteome Res.* **3**, 253–260 (2004).
22. Jaffe, J.D. *et al.* Accurate inclusion mass screening: a bridge from unbiased discovery to targeted assay development for biomarker verification. *Mol. Cell. Proteomics* **7**, 1952–1962 (2008).
23. Schmidt, A., Claassen, M. & Aebersold, R. Directed mass spectrometry: towards hypothesis-driven proteomics. *Curr. Opin. Chem. Biol.* **13**, 510–517 (2009).
24. Baty, J.D. & Robinson, P.R. Single and multiple ion recording techniques for the analysis of diphenylhydantoin and its major metabolite in plasma. *Biomed. Mass Spectrom.* **4**, 36–41 (1977).
25. Stahl-Zeng, J. *et al.* High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Mol. Cell. Proteomics* **6**, 1809–1817 (2007).
26. Picotti, P. *et al.* A database of mass spectrometric assays for the yeast proteome. *Nat. Methods* **5**, 913–914 (2008).
27. Pan, C., Olsen, J.V., Daub, H. & Mann, M. Global effects of kinase inhibitors on signaling networks revealed by quantitative phosphoproteomics. *Mol. Cell. Proteomics* **8**, 2796–2808 (2009).
28. Malmstrom, J. *et al.* Proteome-wide cellular protein concentrations of the human pathogen Leptospira interrogans. *Nature* **460**, 762–765 (2009).
29. Addona, T.A. *et al.* Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nat. Biotechnol.* **27**, 633–641 (2009).
30. Ross, P.L. *et al.* Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3**, 1154–1169 (2004).
31. Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W. & Gygi, S.P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. USA* **100**, 6940–6945 (2003).
32. Rivers, J., Simpson, D.M., Robertson, D.H., Gaskell, S.J. & Beynon, R.J. Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. *Mol. Cell. Proteomics* **6**, 1416–1427 (2007).
33. Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D.A. & White, F.M. Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proc. Natl. Acad. Sci. USA* **104**, 5860–5865 (2007).
34. Picotti, P. *et al.* High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nat. Methods* **7**, 43–46 (2010).

© 2010 Nature America, Inc. All rights reserved.
npg