Liquid Chromatography-Mass Spectrometry-based Quantitative Proteomics*

Published, JBC Papers in Press, June 1, 2011, DOI 10.1074/jbc.R110.199703
Fang Xie, Tao Liu, Wei-Jun Qian, Vladislav A. Petyuk, and Richard D. Smith

From the Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99352

LC-MS-based quantitative proteomics has become increasingly applied to a wide range of biological applications due to growing capabilities for broad proteome coverage and good accuracy and precision in quantification. Herein, we review the current LC-MS-based quantification methods with respect to their advantages and limitations and highlight their potential applications.

LC-MS-based quantitative proteomic approaches have become increasingly popular over the past decade (1–7). In general, discovery-based proteomic efforts lend themselves to global analyses whereby a broad survey of the proteome is performed across various samples, and the quantitative differences among them are estimated. In discovery-based efforts, the breadth of analysis is emphasized more than the precision and accuracy of quantification. For studies in which these qualities are crucial such as verification efforts, the tactic switches to sensitive, precise, and accurate analysis of a few targeted proteins in relatively large sets of samples and internal standards are often used.

Fig. 1 illustrates an LC-MS-based global proteomic workflow in which proteins are converted into peptides for identification and quantification (i.e., “bottom-up” proteomics). Typically, methods are applied in conjunction with enzymatic digestion of proteins and subsequent measurement of one or more peptides from each protein that serve as effective measurement surrogates. We note that direct measurement of intact proteins (i.e., “top-down” proteomics) is another analytical option but is beyond the scope of this minireview and therefore not discussed herein. In global analyses, relative quantification of peptides usually involves either label-free or stable isotope labeling techniques (1) to discern differences in protein abundances among different biological conditions, and results are often expressed as “-fold changes.” Overall, label-free approaches have wider dynamic range and broader proteome coverage, whereas stable isotope labeling approaches offer higher quantification precision and accuracy (8). Another common approach is absolute quantification, which determines the exact amount or concentration of a peptide/protein in a given sample and requires the use of an appropriate “internal” standard. All of these approaches have considerably different pros and cons that must be weighed before deciding which one is best for a specific course of study.

Challenges affecting quantification in a bottom-up proteomic workflow stem from the wide range of peptide and protein physicochemical properties that give rise to large differences in MS responses (8). Sample handling, digestion efficiency, and separation also can have an impact on results. As such, relative peptide intensities may not directly reflect the relative abundances of different proteins. A major factor that influences LC-MS-based quantification via electrospray ionization is ion suppression (9). Peptide intensity depends on the quantity of the peptide being ionized as well as on ionization efficiency and, under some conditions, on the properties of co-eluting peptides. The use of lower flow rates (e.g. <100 nl/min) (9, 10) or internal standards (11, 12) can help alleviate ion suppression. Other issues in LC-MS-based quantification include the separation peak capacity and reproducibility of the chromatography and the mass measurement accuracy and resolving power of the mass spectrometer. Significant technological advances such as the development and commercialization of ultra-performance LC and high-mass accuracy/resolution mass spectrometers have substantially overcome these issues, making LC-MS-based quantification more reliable and accessible to biologists.

Basically, there is no recognized “one-size-fits-all” method that fulfills every quantitative need, and available options for quantification can make it difficult for an investigator to choose the most appropriate approach to answer particular biological questions. This minireview presents the advantages and limitations of commonly used LC-MS-based protein quantification approaches and provides guidelines for researchers who may not be familiar with but would like to benefit from quantitative proteomic measurements.

Label-free Quantification

Straightforward and inexpensive, label-free quantification is being increasingly applied to proteomic measurements. Without the need to modify peptides/proteins with stable isotope-containing compounds or to add heavy isotope-labeled internal standards to the sample, label-free approaches require minimal manipulation of the sample and can be used on any type of biological materials. Conceptually, label-free quantification allows for the comparison of an unlimited number of samples; however, each sample has to be analyzed individually (no sample multiplexing). This type of approach usually offers wide dynamic range, which is especially advantageous when rela-
tively large-abundance changes are expected. The two major label-free quantification approaches are spectral counting and MS ion intensity (or peak area) measurement.

Spectral counting, i.e. counting the number of MS/MS spectra identifying a given peptide or protein, represents a simple approach for relative quantification without stable isotope labeling (13, 14). The rationale behind this method, which has been experimentally validated (13), is that the frequency for which a peptide is selected for MS/MS fragmentation is positively correlated to its quantity in the data-dependent acquisition operation mode. In this mode, a survey MS scan is acquired, followed by selection of typically **10** of the most abundant ions in the survey MS scan for subsequent MS/MS analysis (i.e. undersampling). Spectral counts for different peptides from a given protein can be summed up for relative quantification of the protein; however, the linearity and the number of quantified proteins will depend on the data-dependent acquisition setting details (15). Although spectral counting is straightforward, small numbers of spectral counts for proteins present in low abundance provide less robust quantitative measurements of these proteins due to statistical limitations and other factors such as the need to limit the false discovery rate so as not to “count” lower quality spectra (16).

Label-free quantification can also be performed based on MS ion intensity (or peak area) for the peptides in a given sample. These LC-MS-based approaches alleviate the undersampling issue inherent in typical LC-MS/MS analysis (17). An example of this type of approach is the accurate mass and time tag strategy (18–20), which utilizes high-resolution high-mass accuracy LC-MS to analyze individual samples. Peptides are identified by matching accurate mass and normalized elution time features to those stored in a previously established reference database. Following accurate alignment of detected LC-MS features (i.e. LC retention time and \( m/z \) values) across different analyses, the areas under chromatographic elution profiles of the identified peptides can be compared among different samples for relative quantification. Similar quantitative approaches that rely on direct LC-MS measurements, feature alignment, and peak identifications have also been reported (21, 22).

In label-free intensity-based quantification, any variations in sample preparation, LC-MS reproducibility, ionization efficiency, and other sources of “instrument drift” can lead to increased measurement error. Therefore, it is important to normalize the data to correct (as much as possible) for systematic variations (23). Software tools for performing LC-MS feature alignments, peak matching, data normalization, and statistical analyses of label-free quantification data are available (20, 22, 24). Although label-free intensity-based quantification has been broadly applied, its accuracy and reliability for quantification are inherently limited by the reproducibility related to
Stable Isotope Labeling Approaches for Quantification

Stable isotope labeling approaches enable accurate quantification based on a stable isotope dilution concept (1). Because a stable isotope-labeled peptide has the same chemical properties as its native equivalent, the two peptides within a mixture should exhibit identical behaviors during LC and electrospray ionization processes and be separated by their differences in mass, thus enabling accurate peptide (but not necessarily protein) quantification. The relative abundance differences measured by MS between the two peptide forms are taken to quantitatively reflect true differences in abundance within the mixture. $^{13}$C, $^{15}$N, and $^{18}$O are commonly used stable isotopes (25) that can be incorporated into proteins or peptides metabolically, enzymatically, or chemically. Although sometimes used, $^2$H is less desirable because it changes the physicochemical properties of peptides such that the heavy form elutes slightly earlier than the light form in reverse-phase LC (26). In stable isotope labeling methods, samples are combined after labeling and analyzed by LC-MS or LC-MS/MS, which essentially avoids the uncertainty induced by variations in instrument performance between measurements. As a result, quantification precision and accuracy are markedly improved compared with label-free approaches.

The most popular metabolic labeling approach is stable isotope labeling with amino acids in cell culture (SILAC)$^2$ (27). Briefly, one or several amino acids (typically Arg or Lys) are labeled with “heavy” isotopes (e.g. $^{13}$C and/or $^{15}$N atoms) and added to the growth medium. Then, the heavy isotope-labeled amino acids are incorporated into all the proteins after several cell doublings. Equal amounts of the heavy isotope-labeled proteome and the normal-labeled proteome can be mixed at the level of either intact cells (i.e. equal numbers of heavy and normal isotope-labeled cells) or cell lysates (i.e. equal amounts of heavy and normal isotope-labeled protein contents) and analyzed. Peptide abundance ratios are determined in the MS mode by comparing the intensities of the labeled and unlabeled peptides within a survey mass spectrum. SILAC can be used to quantify in vivo changes (e.g. protein turnover) as the isotopes are introduced during the natural biosynthesis process. Another advantage is that multiplexing is possible due to the availability of several labels (e.g. $^{12}$C$_6$$^{15}$N$_4$-Arg, $^{13}$C$_6$$^{14}$N$_4$-Arg, and $^{13}$C$_6$$^{15}$N$_4$-Arg) (1, 12). A shortcoming is that SILAC is limited to cell cultures and cannot easily be applied to tissues or biofluids.

With enzymatic labeling, $^{18}$O atoms are incorporated into the C terminus of every proteolytic peptide during or after protein digestion catalyzed by trypsin (28), Glu-C (29), or some other proteases (30). Previously, two main issues, incomplete labeling and $^{18}$O back-exchange to $^{16}$O, limited the application of $^{18}$O labeling for global quantitative proteomics. Typically, $^{18}$O labeling is performed when a sample is digested in H$_2^{18}$O solution with trypsin; however, different peptides often have different labeling efficiencies. We and others demonstrated more effective labeling by directly incubating peptides in H$_2^{18}$O water with trypsin as a catalyst (31, 32), and a recently improved protocol that incorporates a boiling step after labeling now prevents oxygen back-exchange (33). $^{18}$O labeling coupled with the accurate mass and time tag approach has been applied in pairwise global quantitative proteomics (31, 34), and more large-scale applications have also been reported (35, 36). The main advantages of $^{18}$O labeling are its simplicity in labeling, relatively low cost, and applicability to all types of samples (e.g. tissues, cells, and biological fluids). Additionally, many different software tools are available to facilitate the $^{18}$O/$^{18}$O-based quantification (31, 37, 38).

Another commonly applied stable isotope labeling strategy is based on incorporating isotope-containing tags into proteins or peptides via chemical reactions, a strategy that can be applied to any type of biological materials. The first reported chemical labeling proteomic approach was isotope-coded affinity tag (ICAT), which reacts specifically to the sulfhydryl group of Cys residues (39). Relative protein abundance is determined in the MS mode, and the mass shift is 9 Da per ICAT-labeled Cys (provided $^{13}$C is used in the heavy form) (40). More recently, novel chemically reactive tags have been developed for labeling peptides that facilitate quantitative analysis of multiple samples simultaneously (i.e. multiplexing), which is particularly useful for following a biological system over multiple time points (8, 41). The most commonly used tags for peptide labeling are isobaric tags for relative and absolute quantification (iTRAQ) (42–44) and tandem mass tags (TMT) (45, 46). iTRAQ and TMT labels react specifically with primary amine groups of tryptic peptides, e.g. N-termini and the side chains of Lys residues, and the tagging reactions are largely complete without major side reactions. The same peptides labeled with different isobaric tags have exactly the same mass and co-elute precisely in LC separations. Although tags remain indistinguishable in the MS scan, they fragment into reporter ions of different masses in the MS/MS scan. The intensities of the different reporter ions are then used to determine the relative abundance of the corresponding peptides and proteins in different samples. It is important to note that both iTRAQ- and TMT-based quantifications require the ability to observe low m/z fragment ions (i.e. the reporter ions), which limits the type of mass spectrometers that can be used. The iTRAQ reagent can be utilized to analyze up to four (4-plex iTRAQ) or eight (8-plex iTRAQ) samples simultaneously, whereas TMT can address two to six samples. Because iTRAQ and TMT label-based quantifications are measured at the MS/MS level, potentially higher signal/noise ratios may be obtained for quantification compared with those obtained at the MS level. Additionally, the accuracy of MS/MS level quantification depends on the isolation window for selected precursors in the first stage MS (8, 47), which is typically 3 Thomson, as all ions within that window will fragment, and potential interferences could skew the quantification results.

$^2$The abbreviations used are: SILAC, stable isotope labeling with amino acids in cell culture; ICAT, isotope-coded affinity tag; iTRAQ, isobaric tag for relative and absolute quantification; TMT, tandem mass tag; SRM, selected reaction monitoring.
“Universal” Reference for Relative Quantification

A challenge for large-scale quantitative applications involving label-free and/or labeling-based proteomics is maintaining platform reproducibility for large studies that may extend over several months or years. The use of a stable isotope-labeled whole proteome as a universal internal standard offers a solution to this challenge. Briefly, a pooled reference sample can be generated, and the digested peptides can be labeled with $^{18}$O (as an example) to serve as the universal reference. The labeled reference sample can be added to each unlabeled biological sample so that each unlabeled peptide will have a corresponding $^{18}$O-labeled version of the peptide from the universal reference. Relative peptide and protein abundances in many different samples can be compared based on their ratios to the universal reference in each analysis. The concept is similar to the use of synthetic isotope-labeled peptides as internal standards (48) that are added to each biological sample. The $^{18}$O-labeled universal reference can be generated from any type of biological sample and requires only a single step to label a digest of the pooled sample. This strategy has been applied to two-dimensional proteome mapping of mouse brain (49) and in human plasma proteome studies to discover biomarkers (36, 50).

A “super-SILAC” approach has been applied to quantify proteins in human tumor tissues (51). In this approach, equal amounts of SILAC-labeled proteins from several previously established cancer-derived cell lines were combined (i.e. a super-SILAC mixture). This mixture served as a “global internal standard” to quantify relative protein abundance from multiple tissue samples of the same tumor type. Unlike the $^{18}$O-labeled universal reference, the super-SILAC strategy is limited by the availability of appropriate cell lines for generating SILAC-labeled samples and cannot be applied to biofluids.

Although the universal reference or global internal standard approach offers great flexibility for large-scale relative quantification studies, it is difficult to produce a “true” universal reference that contains all the proteins and protein forms of interest and that can be reproducibly generated. The nature of adding a reference/standard makes it inevitable to dilute every sample 1:1 with the standard and thus makes it more difficult to detect and quantify lower abundance proteins.

Targeted Quantification

Global quantitative proteomics inherently suffers some limitations as a result of missing data in individual analyses, false identifications, reproducibility issues, and computational challenges. Targeted quantification approaches using selected reaction monitoring (SRM) or multiple reaction monitoring for accurate quantification of selected analytes are gaining in popularity as a means of overcoming these limitations. SRM is typically performed using a “triple quadrupole” tandem mass spectrometer that consists of a selection quadrupole for the precursor ion, a collision quadrupole, and another selection quadrupole for the fragments. Specifically, an SRM assay defines a list of precursor m/z values associated with specific retention times. The critical rules for precursor selection are defined in a comprehensive review (52). For each m/z value, the assay defines one or several fragment ions that are predicted to have good MS response and are readily distinguished from interference. By setting the MS platform to exclusively monitor predefined precursor-to-fragment ion transitions in rapid succession, a specific ion can be detected and quantified when it has the expected m/z value and produces fragments of the expected m/z. Recently, an interesting variant, intelligent SRM, was developed to confirm the precursor identity without significantly perturbing the SRM quantification (53). In this method, additional transitions are acquired in a data-dependent fashion (triggered when all the primary transitions exceed a predefined threshold), which increases the specificity of the analysis.

Among the benefits of SRM-based quantification is the excellent reproducibility attained by using labeled synthetic peptides as internal standards (54). Additionally, the approach is less affected by sample complexity, as noise signals are filtered out at both the precursor and fragment levels (55). As a result, SRM has the highest sensitivity and a wide dynamic range that extends 4–5 orders of magnitude, which makes SRM well suited for targeted quantification experiments such as biomarker verification. SRM can also be applied to quantify particular peptide modifications (e.g. ubiquitination and phosphorylation) (41, 56).

An important application of SRM is absolute quantification, which is achieved by using known concentrations of synthetic isotope-labeled peptides as internal standards (48). In this method, a target peptide is synthesized in an isotope-labeled form and added to the protein digest at a known concentration. The concentration of the native peptide is determined by comparing the ion intensities between the labeled and unlabeled forms. The internal standard can correct for ion suppression and matrix effects (11). For an experiment with multiple samples, the use of an internal standard ensures a fair comparison across all samples. However, the protein amount determined by absolute quantification may not reflect the true expression levels in the original sample because the internal standard is added after digestion and thus cannot correct for the variable losses or enrichments that occur during sample preparation (57), i.e. not true “absolute quantification.” Using an isotope-labeled protein as the internal standard may alleviate this problem because the surrogate protein can be combined with the target protein at the very beginning of the analytical process. The standard can be either a full-length isotope-labeled equivalent of a target protein (58) or an isotope-labeled artificial concatamer of proteotypic peptides from several proteins of interest, which allows multiplexing (59). Regardless, identifying and synthesizing suitable internal standards are not trivial tasks, which limits the absolute quantification approach to a restricted set of prescreened proteins (1), making it suitable for applications such as validating potential protein candidates of interest or quantifying particular post-translational modifications such as ubiquitination (60).

Factors to Consider When Choosing a Quantification Strategy

When it comes to quantitative proteomics, many options are available, each with their own set of advantages and disadvantages. Some factors to be considered when making a choice
include the types of samples, the number of samples to be compared, the biological sources and complexity of the samples, the analytical needs (e.g. quantification precision, accuracy and whether an absolute concentration is necessary) of the biological problem, and the cost of the experiment (12).

Table 1 summarizes the performance characteristics for each of the common LC-MS-based proteomic quantification strategies. Among the stable isotope labeling techniques, SILAC offers the best accuracy for quantification because it labels at the cell culture level. Both 

\[ ^{18} \text{O} \] labeling, SILAC, and ICAT fall into this category. B, multiple comparison is used to compare up to four, six, or eight samples depending the isobaric tags used (e.g. 4-plex iTRAQ, 6-plex TMT, or 8-plex iTRAQ). C, large-scale comparison employs an internal standard: 

\[ ^{18} \text{O} \] labeled universal reference, super-SILAC mixture, or synthetic isotope-labeled peptides. AQUA, absolute quantification.

**TABLE 1**

<table>
<thead>
<tr>
<th>Overview of the characteristics of different LC-MS-based quantification approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proteome coverage</strong></td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>Spectral counting</td>
</tr>
<tr>
<td>AUC/ion intensity</td>
</tr>
<tr>
<td>SILAC</td>
</tr>
<tr>
<td>[ ^{18} \text{O} ]</td>
</tr>
<tr>
<td>ICAT</td>
</tr>
<tr>
<td>iTRAQ/TMT</td>
</tr>
<tr>
<td>[ ^{18} \text{O} ] universal reference</td>
</tr>
<tr>
<td>Super-SILAC</td>
</tr>
<tr>
<td>SRM</td>
</tr>
</tbody>
</table>

\[ ^{18} \text{O} \] labeled universal reference employs an internal standard: 

\[ ^{18} \text{O} \] labeled universal reference, super-SILAC mixture, or synthetic isotope-labeled peptides. AQUA, absolute quantification.

\[ ^{18} \text{O} \] labeled universal reference employs an internal standard: 

\[ ^{18} \text{O} \] labeled universal reference, super-SILAC mixture, or synthetic isotope-labeled peptides. AQUA, absolute quantification.

**Conclusion and Perspective**

In conclusion, proteomic quantification is a multifaceted term encompassing global and targeted measurements that can involve relative and/or absolute abundance determinations across large sets of proteins. No single “gold standard method” can resolve all of the analytical problems associated with pro-
MINIREVIEW: LC-MS-based Quantitative Proteomics

VOLUME 286 • NUMBER 29 • JULY 22, 2011
32. Fenselau, C., and Yao, X. (2009) J. Proteome Res. 8, 2140–2143
49. Petyuk, V. A., Qian, W. J., Chin, M. H., Wang, H., Livesay, E. A., Monroe,


