Use of DNA Ladders for Reproducible Protein Fractionation by Sodium Dodecyl Sulfate−Polyacrylamide Gel Electrophoresis (SDS−PAGE) for Quantitative Proteomics

Guoan Zhang,[†] David Fenyö,[‡] and Thomas A. Neubert*,[†]

Department of Pharmacology and Skirball Institute of Biomolecular Medicine, New York University School of Medicine, New York City, New York 10016, and The Rockefeller University, New York City, New York 10021

In proteomics, one-dimensional (1D) sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) is widely used for protein fractionation prior to mass spectrometric analysis to enhance the dynamic range of analysis and to improve the identification of low-abundance proteins. Such protein prefractionation works well for quantitation strategies if the proteins are labeled prior to separation. However, because of the poor reproducibility of cutting gel slices, especially when small amounts of samples are analyzed, its application in label-free and peptide-labeling quantitative proteomics methods has been greatly limited. To overcome this limitation, we developed a new strategy in which a DNA ladder is mixed with the protein sample before PAGE separation. After PAGE separation, the DNA ladder is stained to allow for easy, precise, and reproducible gel cutting. To this end, a novel visible DNA-staining method was developed. This staining method is fast, sensitive, and compatible with mass spectrometry. To evaluate the reproducibility of DNA-ladder-assisted gel cutting for quantitative protein fractionation, we used stable isotope labeling with amino acids in cell culture (SILAC). Our results show that the quantitative error associated with fractionation can be minimized using the DNA-assisted fractionation and multiple replicates of gel cutting. In conclusion, 1D PAGE fractionation in combination with DNA ladders can be used for label-free comparative proteomics without compromising quantitation.

Keywords: DNA • quantitation • fractionation • SILAC • mass spectrometry • proteomics

Introduction

In shotgun proteomics, protein digests are usually analyzed by liquid chromatography−tandem mass spectrometry (LC−MS/MS) in a data-dependent manner, in which the most intense peaks in mass spectra are selected for sequencing by MS/MS. In such experiments, it has been proposed that analysis of complex protein mixtures is not limited by the sensitivity but rather by the dynamic range and sequencing speed of MS.[1,2] A simple and efficient way to alleviate this problem is to separate the sample into multiple fractions prior to LC−MS. This can improve protein identification in two ways: (1) by reducing sample complexity, which means less demand for a wide dynamic range and high sequencing speed of the mass spectrometer, and (2) by decreasing the number of components in each fraction so that a larger amount of each component can be analyzed without overloading the LC−MS system. Therefore, sample fractionation can dramatically improve protein identification. Sample fractionation can be performed at either the protein or peptide level. For protein fractionation, gel-enhanced liquid chromatography–mass spectrometry (GeLC−MS) is generally considered the method of choice.[1,3,4]

In this approach, proteins are separated by one-dimensional (1D) sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) prior to LC−MS analysis. For peptide fractionation, the most commonly used approach is two-dimensional (2D) LC, where the peptides are separated by ion-exchange liquid chromatography followed by reverse-phase liquid chromatography.[5] While both strategies are widely used, peptide fractionation by 2D LC may suffer from the limitation that peptides from high-abundance proteins dominate the available analytical space in both chromatographic dimensions, which can limit the dynamic range for protein identification. Protein fractionation has a better chance to partition high- and low-abundance proteins into different fractions and thus can improve the identification of low-abundance proteins.[6,7]

While sample fractionation improves protein identification, it can complicate protein quantitation when the samples to be compared have to be processed in parallel. For quantitative approaches based on protein isotope labeling, differentially labeled protein samples can be mixed before fractionation; thus, quantitation is not influenced by fractionation. For peptide labeling and label-free approaches, however, the impact of protein fractionation on quantitation has not been thoroughly investigated. It can be argued that, in principle, quantitation is not affected by fractionation as long as all peptide signals from a protein are summed from all fractions of the sample for quantitation, assuming all fractions are

* To whom correspondence should be addressed. Telephone: (212) 263-7265. Fax: (212) 263-8214. E-mail: neubert@saturn.med.nyu.edu.
† New York University School of Medicine.
‡ The Rockefeller University.

analyzed in the same way. However, in real experiments, the quantitation is complicated because (1) different fractions may contain very different amounts of protein and thus have different levels of signal suppression effect in MS (i.e., the same amount of a given peptide in different fractions can produce different amounts of ion current) and (2) there are technical difficulties in combining signals from different fractions. In experiments where these two factors cannot be ignored, measures have to be taken to ensure the high reproducibility of fractionation.

As an alternative to isotope-labeling approaches, label-free quantitation based on LC–MS techniques has become increasingly popular. The last several years have seen especially rapid progress in this technique. There are basically two types of LC–MS-based label-free approaches. The first type is based on counting the number of matched peptides or MS/MS spectra.[2,8–10] Although this approach is powerful, the quantitation is generally crude, especially when only a few peptides are observed for a protein. The second type is based on the intensity of the same peptide ion in different LC–MS runs,[11–22] and it is therefore more accurate than the first type. In this paper, we will use label-free to refer to the approach where the ion intensity is used for quantitation. Recent advances in label-free techniques have made it possible to obtain reasonable accuracy compared to isotope-labeling approaches. However, in most label-free studies, proteins are not fractionated before LC–MS analysis, possibly because of the concern that sample fractionation may compromise quantitation. Similarly, protein fractionation is also a problem and is generally not used when the isotope labeling occurs at the peptide level; therefore, fractionation is usually limited to the peptide level. Thus, there is a need for a reproducible approach for sample fractionation at the protein level for label-free or peptide-labeling quantitative applications.

For GeLC–MS, the major difficulty for reproducible protein fractionation seems to be the gel-cutting step, which is typically carried out manually. To aid in precise gel cutting, marker bands are needed and two types of markers can be used. The first type of potential marker is stained protein bands from samples as "internal" markers. However, this requires that (1) samples contain enough material to allow for visualization of a good number of clear bands after staining and (2) samples have bands that are common to all samples and that are well-distributed across the whole molecular-weight range of the SDS–PAGE. In our experience, these requirements are often difficult to satisfy when working with low levels of proteins, such as are available when studying intracellular signal transduction processes. When the amount of sample is limited, sensitive staining methods, usually silver staining, have to be used to visualize bands. However, it has been shown that silver staining can cause cross-linking of proteins inside the gel, resulting in decreased sequence coverage by MS analysis after in-gel digestion compared to Coomassie Brilliant Blue (CBB) staining.[23] Fluorescent proteins stains[24] have good sensitivity and are compatible to MS. However, they need UV light to visualize protein and thus are inconvenient to use. Moreover, in some cases, even silver staining may not produce enough protein bands, even in cases when proteins can readily be identified by MS. Even when sufficient number of bands can be visualized, it may be difficult to find a set of bands as markers for cutting: when protein bands are smeared because of high sample loading or high sample complexity or when no suitable bands can be found for a specific molecular-weight

range. The second type of potential marker is protein molecular-weight standards. The advantage of this approach is that it gives a predictable set of protein bands across almost any given molecular-weight range. The disadvantage is the markers are used as "external standards"; i.e., markers and samples are run in parallel lanes and not added into samples because the markers themselves are proteins and can affect subsequent MS analysis.

To address this difficulty in gel cutting for quantitative analysis, we developed a new strategy in which DNA ladders are mixed with protein samples before SDS–PAGE separation. After electrophoresis, the DNA ladders instead of proteins are stained to allow for easy but precise gel cutting. To this end, a novel visible DNA-staining method was developed. This method is fast, sensitive, and compatible to MS and conventional protein-staining methods, including CBB and zinc staining. We used stable isotope labeling with amino acids in cell culture (SILAC) to evaluate the feasibility of using this DNA-assisted gel-cutting method for reproducible protein fractionation.

Materials and Methods

Cell Culture and Metabolic Labeling. Two populations of NG108 cells (mouse neuroblastoma and rat glioma hybrid) stably overexpressing ephrinB1 were maintained in Lys- and Arg-depleted Dulbecco's modified Eagle's medium (Specialty Media) supplemented with 10% dialyzed fetal bovine serum (Invitrogen), hypoxanthine–aminopterin–thymidine (Sigma), 100 units/mL penicillin/streptomycin (Invitrogen), 0.4 mg/mL G418 (CalBiochem), and either normal or $^{13}C_6$ Lys and $^{15}N_4$ Arg (Cambridge Isotope Laboratories), respectively. Cells were grown for at least six divisions to allow for full incorporation of labeling amino acids. After the cells were metabolically labeled, they were lysed in buffer containing 1% Triton X-100, 150 mM NaCl, 20 mM Tris at pH 8, 0.2 mM ethylenediaminetetraacetic acid (EDTA), 2 mM Na_3VO_4, 2 mM NaF, and protease inhibitors (Complete tablet; Roche Applied Science). Lysates were clarified by centrifugation at $14000g$ for 20 min.

Gel Electrophoresis and Staining. Samples containing a DNA ladder (1 μg/μL) (1 kb plus, Invitrogen) were mixed with equal volumes of Laemmli sample buffer (Bio-Rad) before loading onto 8.6 × 6.8 cm precast Tris-HCl gels (Bio-Rad). For DNA staining with indoine blue (IB) (Sigma), gels were fixed in 7% acetic acid/40% ethanol for 20 min before they were stained with 0.025% IB in 7% acetic acid/40% ethanol for 25 min. Finally, gels were washed with 7% acetic acid/40% ethanol for 5 min. While IB is considered to be relatively safe (Sigma-Aldrich, Materials Safety Data Sheets), its toxic effects have not been studied as thoroughly as fluorescent dyes, such as ethidium bromide. Because of its ability to associate strongly with DNA, thus suggesting the possibility of mutagenesis, standard precautions should be observed when using IB. For CBB staining, gels were stained with 0.1% CBB-R250 in 7% acetic acid/40% ethanol for 30 min and destained with 7% acetic acid/40% ethanol until the background was clear. For silver and zinc staining, the silver-staining kit SilverQuest (Invitrogen) or the reversible stain kit E-Zinc (Pierce) were used. Gel cutting was performed manually with a scalpel.

In-Solution and In-Gel Digestion. For in-solution digestion, bovine serum albumin (BSA) was incubated in 25 mM NH_4HCO_3 with trypsin (Promega) at a ratio of 1:50 (enzyme/protein) for 4 h at 37 °C after heat denaturation of target proteins at 95 °C for 5 min. In-gel digestion was performed using a modified version of the protocol developed by Shevchen-

ko et al.[25] Briefly, excised gel bands were cut into small pieces and destained in 25 mM NH_4HCO_3 and 50% acetonitrile, dehydrated with acetonitrile, and dried. Then, the gel pieces were rehydrated with 12.5 ng/μL trypsin solution (in 25 mM NH_4HCO_3) and incubated overnight at 37 °C. Peptides were extracted twice with a solution containing 5% formic acid and 50% acetonitrile followed by a final extraction with acetonitrile. Samples were dried with vacuum centrifugation before further preparation or analysis.

Mass Spectrometry. For LC−MS/MS analysis, a linear quadrupole ion-trap (LTQ)-Orbitrap hybrid mass spectrometer (ThermoFinnigan) equipped with a nano-electrospray ionization (ESI) source (Jamie Hill Instrument Services) was used. A Nano-Acquity UPLC system (Waters) equipped with a 100 μm × 15 cm reverse-phase column (Symmetry C18, Waters) was coupled to the ion-trap instrument via a 10 μm inner diameter PicoTip emitter (New Objective). Samples were loaded onto a trap column (180 μm × 2 cm Symmetry C18, Waters) with 3% acetonitrile in 0.1% formic acid for 5 min at 4 μL/min. After sample loading, the flow rate was reduced to 0.4 μL/min and directed through the analytical column and peptides were eluted by a gradient of 7–50% acetonitrile in 0.1% formic acid over 120 min. Mass spectra were acquired in data-dependent mode with one 60 000 resolution MS survey scan by the Orbitrap and four concurrent MS/MS scans in the LTQ for the most intense four peaks selected from a preliminary 15 000 resolution spectrum from each survey scan. Automatic gain control was set to 500 000 for Orbitrap survey scans and 10 000 for LTQ MS/MS scans. Survey scans were acquired in profile mode, and MS/MS scans were acquired in centroid mode. Mascot generic format files were generated from the raw data using DTASuperCharge (version 1.01) and Bioworks (version 3.2, ThermoFinnigan) for database searching.

For matrix-assisted laser desorption ionization−time of flight (MALDI−TOF) analysis, a Micromass (Manchester, U.K.) TOF Spec-2E mass spectrometer equipped with a nitrogen laser (337 nm) was used. Recrystallized 2,5-dihydroxybenzoic acid (160 mg/mL, Sigma) in 1% trifluoroacetic acid (TFA)/30% acetonitrile (ACN) was used as the MALDI matrix. Typically, 150 laser shots were summed into each MS spectrum. MS spectra were processed by Masslynx 4.0 software to generate peak lists for peptide mass fingerprinting (PMF).

Database Searching. Mascot software (version 2.1.0, Matrix Science, London, U.K.) was used for database searching. For BSA data, a Swissprot Mammalia database (downloaded September 19, 2006) was used. Otherwise, an IPI database containing mouse and rat protein sequences (downloaded November 17, 2006) was used. For PMF searching, peptide mass tolerance was 100 ppm. Trypsin specificity was applied with a maximum of one missed cleavage. For LC−MS/MS data, the peptide mass tolerance was 0.03 Da, fragment mass tolerance was 1 Da, trypsin specificity was applied with a maximum of one missed cleavage, and variable modifications were $^{13}C_6$ Lys and $^{15}N_4$ Arg. To control the false-positive rate for protein identification, a decoy database was created by reversing the protein sequences of the original database. On the basis of the decoy database searching, three filters for protein identification were applied: (1) Peptide score threshold was 20. (2) Protein score threshold was 60. (3) Each protein was identified on the basis of at least two peptides. After these filters were applied, no false-positive protein hits were found when using our LC−MS/MS data to search the reversed database.
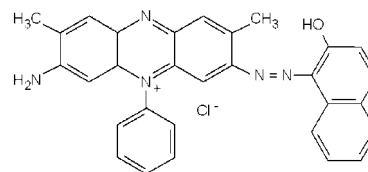

Figure 1. Chemical structure of IB.

Protein Quantification. SILAC ratios were determined using the open-source software MSQuant (version 1.4.0a16) developed by Matthias Mann, Peter Mortensen, and colleagues at the University of Southern Denmark. Protein ratios from automated MSQuant analysis were subjected to manual inspection.

Results and Discussion

Novel Visible DNA-Staining Method. Traditionally, DNA separated by gel electrophoresis is stained by fluorescent dyes, such as ethidium bromide. However, the use of fluorescent dyes as guides to reproducibly cut gels is inconvenient because it requires UV light to visualize DNA bands. Prolonged exposure to UV light can also be dangerous. Therefore, we set out to find a visible staining method for DNA that could be used with SDS−PAGE for the separation of proteins. A variety of staining methods based on visible dyes have been reported previously, such as methylene blue[26] and Nile blue.[27,28] However, these methods require long staining time and have poor sensitivity because of high background staining. This was confirmed by our own result from methylene blue staining (data not shown). More recently, Choi et al. developed the counterion−dye staining method,[29,30] which showed low background staining and hence improved sensitivity. We tried this method,[29] but the sensitivity was poor in our hands, possibly because this method is not compatible with SDS−PAGE.

In view of the problems with existing visible staining methods, we developed a new method that is sensitive, fast, and fully compatible with SDS−PAGE. In this method, IB was used as the dye to stain DNA. The structure of IB is shown in Figure 1. Most likely, IB is able to bind to DNA molecules in the gel through electrostatic interaction[27] or possibly through intercalation with the nucleobases. We used 40% ethanol and 7% acetic acid in both staining and washing buffer. The use of ethanol is to enhance the solubility of IB, which does not dissolve well in aqueous solutions. We found that the addition of acetic acid into the staining solution can significantly decrease background staining, possibly by reducing nonspecific interactions between the dye and polyacrylamide gel matrix: DNA bands were usually visualized even without destaining. In contrast, the presence of acetic acid and ethanol apparently did not affect DNA−IB interaction, and once stained, the gel can be stored in the washing solution for months without detectable fading. The new staining method is very simple, and the whole procedure takes less than 1 h. The sensitivity of the new method was measured to be around 10–15 ng/band as shown in Figure 2. This level of sensitivity is better than most published visible DNA-staining methods.

It is also worth noting that the use of SDS−PAGE for DNA separation has only been rarely reported.[31] Agarose gel electrophoresis and nondenaturing PAGE are the most commonly used approaches. However, our experiments showed that SDS−PAGE is a good alternative for DNA separation, yielding higher resolution than agarose gel electrophoresis.
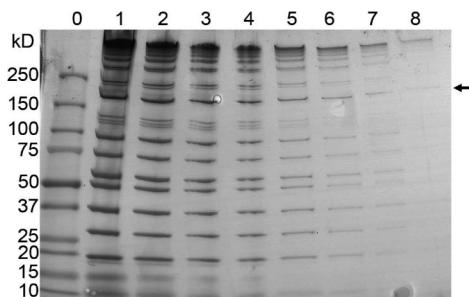
Figure 2. Sensitivity of IB staining of DNA separated by 4–15% SDS−PAGE. The gel was stained by IB as described in the Materials and Methods. Prestained protein molecular-weight markers were loaded into lane 0. Total amounts of DNA loaded into lanes 1–8 were 10, 5, 2.5, 1.25, 0.62, 0.31, 0.16, and 0.08 µg, respectively. The DNA band indicated by the arrow accounted for 8% of the total amount of DNA according to the manufacturer.

For the new staining method to be used to assist protein fractionation, it is critical that the staining is selective, i.e., stains DNA but not protein, so that the DNA ladder is not complicated by protein bands from the sample. To test the selectivity of the staining method, NG108 cell lysate containing about 100 µg of total protein was separated by SDS−PAGE and stained using IB and CBB, respectively. While CBB staining revealed dark bands, no protein bands were observed with the IB-staining method (data not shown). In another test, 2 µg of the DNA ladder was separated by SDS−PAGE and stained using zinc and CBB staining, respectively. Neither staining method was able to visualize the DNA (data not shown), suggesting that the addition of DNA to protein samples would not complicate protein band patterns when these two staining methods are used. Silver staining was not tested, because it is known to stain DNA.

For fractionation purposes, because the DNA-staining method makes it unnecessary to stain proteins, it allows for the avoidance of protein staining methods, such as silver staining, that may affect subsequent MS analysis. In case protein staining is needed for the estimation of protein content, zinc and CBB staining can be used together with IB staining. Because zinc staining is reversible, fairly sensitive (1–10 ng/band), and fully compatible with MS,[32] it can be used before DNA staining without affecting the latter procedure. We have found that the migration rates of DNA markers in SDS−PAGE gels relative to protein molecular-weight markers were consistent, which means that, under the experimental conditions used in this study, the DNA ladder can be used as an indicator of the protein molecular weight. We are not sure, however, whether this correlation remains constant between different buffering systems or different types of gels.

Compatibility of the New DNA-Staining Method with In-Gel Digestion and MS. After having established the new DNA-staining procedure, we further investigated whether the procedure is compatible with in-gel digestion and MS analysis.

First, to make sure that the commercial DNA ladder we used did not contain any contaminating protein that might interfere with subsequent MS analysis, 3 µL of the 1 µg/µL DNA was digested in-solution with trypsin. As a control, the same experiment was also carried out without DNA. Then, the resulting digests were purified by ZipTip before analyzed by MALDI−TOF. A comparison of the MS spectra from the two experiments indicated that no extra peaks were detected from
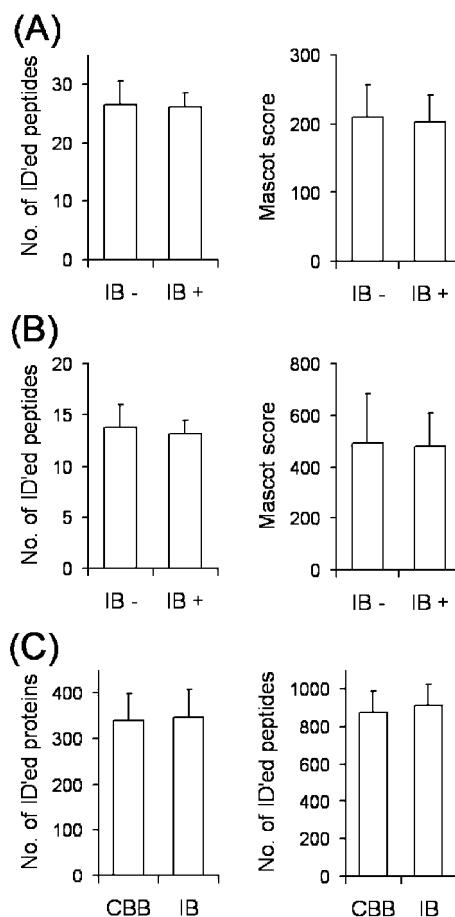

Figure 3. Effect of IB staining on in-gel digestion and MS analysis. A total of 800 fmol of BSA in-gel either with or without IB staining (IB+/IB−) was digested with trypsin. A total of 300 fmol of each resulting digest was analyzed by both MALDI−TOF (A) and LC−MS/MS (B). Each analysis was carried out in five replicates. PMF and MS/MS data were used respectively for Mascot database searching for protein identification. In C, the same amount of the whole cell lysate in-gel was stained with IB or CBB before tryptic digestion and LC−MS/MS analysis. Each analysis was carried out in four replicates. The resulting MS/MS data were used for protein identification by Mascot.

the DNA sample, suggesting that the DNA ladder preparation used in our experiments did not contain any detectable contaminating proteins (data not shown). Next, we tested whether the IB staining affects in-gel digestion and subsequent MS analysis. To this end, 10 aliquots of 800 fmol of BSA were visualized by zinc staining after SDS−PAGE and excised from the gel. After removal of the zinc stain, five of the BSA gel bands were stained with IB. Then, all of the BSA bands were digested in-gel with trypsin. The resulting digests were analyzed by both MALDI−TOF after ZipTip cleaning and LC−MS/MS, with 300 fmol of digest used for each analysis. The peptide mass maps from MALDI−TOF and the MS/MS data from LC−MS/MS were used for protein identification by searching protein sequence databases (Figure 3). As shown in parts A and B of Figure 3, digestion with or without the presence of IB resulted in similar Mascot scores and numbers of matched peptides for both MALDI−MS and LC−MS/MS.

We went on to use a more complex protein mixture, the unlabeled NG108 whole cell lysate, to make a comparison between the reliability of protein identification after the IB and
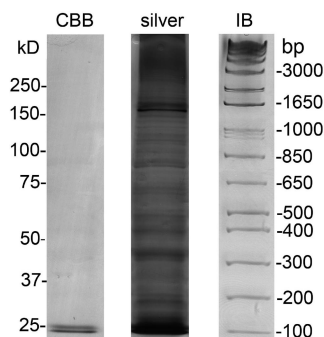
Figure 4. Comparison of CBB-, silver-, and IB-stained protein/DNA bands as markers for SDS−PAGE gel cutting. The cell lysate containing about 5 μg of total protein was separated by SDS−PAGE and subjected to different staining methods. For DNA staining, 2 μg of the DNA ladder was mixed with the lysate before SDS−PAGE.

CBB stainings. To obtain gel pieces of the same volume, a SDS polyacrylamide gel was made using the cell lysate: first, the lysate was mixed with the DNA ladder (final concentration: 0.5 mg/mL protein and 0.02 mg/mL DNA). Then, the gels (7.5%) were casted by mixing this mixture with 30% acrylamide/0.8% bisacrylamide at a 3:1 ratio (v/v) before the addition of ammonium persulfate and N,N,N',N'-tetramethylethylenedi- amine (TEMED). Gel disks of equal sizes were excised from the gel by pressing the open end of a 1 mL pipet tip (7.5 mm i.d., Fisher Scientific) against the gel. Each gel disk contained about 12 μg of protein and 0.5 μg of the DNA ladder. Four gel disks were stained with IB, and another four were stained with CBB before in-gel digestion and LC−MS/MS analysis. The resulting MS/MS data were used for protein identification by Mascot. The results from the two staining methods were very similar in terms of the number of identified proteins and peptides (Figure 3C).

When these results are taken together, they suggest that the IB-staining method does not affect in-gel digestion or MS analysis.

Comparison of Different Staining Methods for Gel Cutting. To compare the applicability of protein or DNA as markers for gel cutting, NG108 cell lysate containing ap- proximately 5 μg of total protein was separated by SDS−PAGE and subjected to CBB, silver, and IB staining. For IB staining, 2 μg of the DNA ladder was mixed with the lysate before SDS−PAGE separation (Figure 4). For protein staining, CBB and silver staining were used because they are the most widely used methods. As shown in Figure 4, with the amount of sample loaded, CBB staining barely produced any visible protein bands because of its low sensitivity. Silver staining revealed many bands. However, too many bands can also be troublesome for gel cutting because of smeared bands and a complicated background. Zinc staining of the protein sample was also performed with a number of bands visualized (data not shown), but because the staining is reversible and usually fades ap- proximately 15 min after staining, it is not suitable for gel cutting. In contrast, the addition of DNA to the protein sample prior to DNA staining produced a regular pattern of clear bands that divided the gel lane nicely into more than 10 sections and thus made it very easy for gel cutting. DNA markers are especially good for comparing a large number of samples because, unlike "endogenous" protein bands that may differ
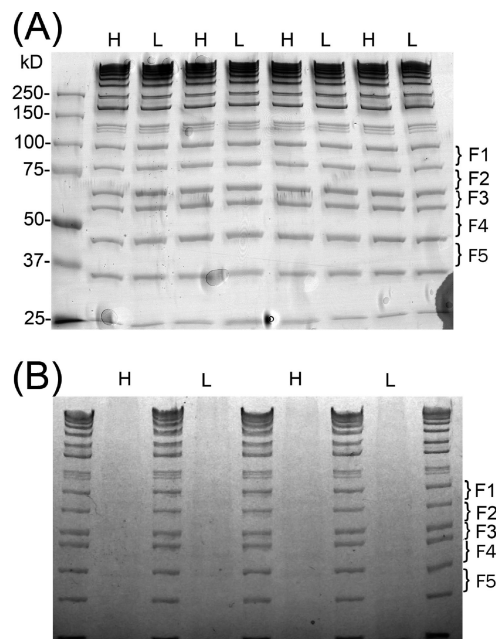

Figure 5. Experimental scheme to evaluate the accuracy of protein fractionation by SDS−PAGE using SILAC. (A) DNA ladders as internal markers. An equal amount of heavy ($^{13}C_6$ Lys/ $^{15}N_4$ Arg labeled) and light (normal Lys/Arg) cell lysate containing about 5 μg of total protein was loaded onto each SDS−PAGE gel lane, with the heavy (H) and light (L) lysate in alternate lanes. A total of 2 μg of the DNA ladder was added to each sample just before sample loading. IB staining was used to visualize the DNA ladders. Five gel slices (F1−F5) from each lane were excised on the basis of the DNA ladders. For each fraction, the heavy and light gel slices were combined for in-gel digestion and LC−MS/ MS. A representative of three identical gels used for the experi- ment is shown. (B) DNA ladders as external markers. The same experiments as the internal experiments were carried out, except that the lysates (without DNA) were loaded onto SDS−PAGE gel lanes with DNA loaded in alternate lanes, and the gels were stained with IB and CBB. A representative of five identical gels used for the experiment is shown.

from sample to sample, DNA ladders are predictable and hence provide consistent and reliable markers for fractionation.

Reproducibility of Protein Fractionation by SDS−PAGE: Internal versus External Markers. We used SILAC to study the reproducibility of protein fractionation by SDS−PAGE. Two types of markers were used as guides for gel cutting: the internal markers (the DNA ladder added into protein samples) and external markers (the DNA ladder from adjacent lanes). For the internal marker experiment, equal amounts of heavy ($^{13}C_6$ Lys/ $^{15}N_4$ Arg) and light (normal Lys/Arg) cell lysates were separated by SDS−PAGE in separate lanes. A total of 2 μg of the DNA ladder was added to each sample just before sample loading. For the external marker experiment, instead of being mixed with the lysates, the DNA ladder was run on lanes adjacent to each lysate lane from both sides. After SDS−PAGE separation, the gels (three identical gels for the internal experiment and five for the external experiment) were stained with IB to visualize the DNA ladders. In the external marker experiment, the gels were stained with CBB after IB staining (Figure 5 shows one gel from each group). On the basis of the DNA ladders, five gel slices from each sample lane were excised as indicated in Figure 5. When the heavy and light gel slices were pooled for each MW fraction and in-gel digestion and LC−MS/MS

analysis were performed, the SILAC ratios of identified proteins reflect the reproducibility of fractionation. To investigate the effect of the replicate number of gel cutting (N) on fractionation reproducibility, different N values ($N = 1$, 3, and 6) were tested: for each of the five fractions, N (1, 3, or 6) pieces of light gel slices were pooled with N heavy gel slices, digested in-gel, and analyzed in a single LC−MS/MS experiment. The same amount of tryptic peptides were analyzed in each run (i.e, one-third and one-sixth of the peptides from the $N = 3$ and 6 pools were analyzed, respectively). All identified proteins were quantified by taking the average ratios for their peptides. The peptide ratios were calculated from the ratios of the peak intensities of the heavy and light peptides (Figure 6).

When the quantitation results from the internal (Figure 6A) and external (Figure 6B) markers were compared, the use of the external markers resulted in much larger errors compared to the internal markers as indicated by more proteins with ratios different from 1 ($\log_2 = 0$), suggesting that the external marker is less accurate for gel cutting. The increased error could be due to slight differences between migrations of samples from different lanes; as one can see from Figure 5, the migration fronts were not always perfectly even. In experiments where limited protein samples are used, it is difficult to detect aberrant migration in sample lanes if no clear protein bands are visualized by protein staining. With internal DNA markers, aberrant sample running can be easily recognized and handled accordingly. Another observation from Figure 6 is that with increasing replicates of gel cutting, protein ratio distributions become more compact, suggesting that the error in quantitation can be reduced by increasing the number of replicates for gel cutting. When these results are taken together, they suggest that the reproducibility of protein fractionation by SDS−PAGE is improved by employing multiple gel cutting and using internal DNA markers as opposed to external markers.

Because the quantitation of proteins at the borders of gel slices would suffer from imperfect gel cutting, the number of affected proteins is dependent upon the size of gel slices. The larger the gel slices are, the fewer proteins suffer from these border effects. In our experiments, 10−12 fractions could be obtained on the basis of the DNA ladder pattern and the heights of gel slices were 3.5–7 mm (distance between DNA bands). This level of fractionation is typical for most GeLC−MS applications. More fractions can be obtained using other suitable DNA ladders if necessary.

Implication of Fractionation Reproducibility on Relative Protein Quantitation. In Figure 6, the quantitation was "fraction-wise"; i.e., proteins were compared between pairs of similar MW fractions (gel slices) of two samples. Those proteins that were differentially partitioned into two or more gel slices in different samples would show large errors. As mentioned in the Introduction, in theory, protein fractionation would not affect quantitation measurements if all peptide signals from a protein were summed from all gel slices containing the protein, thus eliminating the deleterious "edge effects" of protein fractionation. To test to what extent this holds true for practical experiments, "total sample-wise" quantitation was carried out: the abundance of a protein in a sample was represented by the sum of its peptide ion intensities from all of the fractions of the sample, and protein ratios were calculated by comparing protein abundances in the heavy and light samples. To this end, a perl script was developed to parse the entire set of MSQuant result files corresponding to one sample for automated ratio calculation. The results for both the internal and external experiments are shown in Figure 7. A comparison between "fraction-wise" (Figure 6) and "total sample-wise" (Figure 7) quantitation indicated that, although the "total sample-wise" calculation did improve quantitation as compared to "fraction-wise", the improvement was limited (Figure 6 versus Figure 7). Even when peptide signals from different gel slices were summed, improving fractionation accuracy using DNA ladders improved the accuracy of quantitation (Figure 7). We believe improving the accuracy of fractionation helps because of (1) run-to-run variation of LC−MS sensitivity, (2) the different peptide ionization efficiency between fractions as a result of different protein amounts and composition, and (3) variations in peptide identification for different fractions by LC−MS/MS as a result of the use of data-dependent acquisition.[2] In label-free and peptide-labeling experiments, some of these obstacles can be partially overcome by repeating the LC−MS/MS analysis of each sample, but this is at the cost of sample consumption and analytical time. Improvements can also be obtained by applying more sophisticated data processing, including LC retention time alignment, MS intensity normalization, and peptide matching between adjacent fractions, but this can only be performed in cases when there are a sufficient number of peptides observed in both fractions. With the high numbers of fractions in typical GeLC−MS experiments, the large number of repetitive LC−MS runs puts great pressure on LC−MS reproducibility and subsequent data processing. Moreover, some in silico data-processing procedures are not always feasible; for example, it is difficult to do peptide matching between fractions when low-resolution MS instruments are used.[22,33] For a particular type of peptide-labeling approach, isobaric labeling (such as iTRAQ),[34] peptide matching between fractions would not work because quantitation is based on reporter ion signals from MS/MS spectra. Therefore, it is desirable to have accurate fractionation in the first place to minimize the detrimental effect of fractionation on quantitation.

To better illustrate the influence of fractionation on quantitation, cumulative probability plots were generated (Figure 8) for "total sample-wise" quantitation. To produce a standard control sample for quantitation, SILAC-labeled samples labeled with light and heavy isotopes, respectively, were mixed at a 1:1 ratio before SDS−PAGE separation and subsequently analyzed by LC−MS/MS. As shown in Figure 8, the curves became closer to the control when the accuracy of gel cutting was improved using DNA ladders as internal MW markers. With six replicates and internal DNA markers, the curve was almost identical to the control curve, suggesting that six replicates of gel cutting were sufficient to nearly eliminate the error in quantitation because of sample fractionation (gel cutting). On the basis of the probability curves, the quantitative error caused by fractionation can be easily calculated (Table 1). Dependent upon the degree of quantitative accuracy needed for an experiment, the proper N value can be chosen. In most cases, three or more replicates are necessary to ensure accurate quantitation.

Because of the high reproducibility of the new fractionation method described here, it can be used in comparative applications in which parallel protein fractionation is needed to enhance proteome coverage without compromising quantitation. Two major potential applications are label-free quantitative approaches and methods based on peptide labeling, such as iTRAQ[34] and O[18] labeling during protease digestion.[35] For peptide-labeling-based approaches, peptides rather than proteins often are separated extensively, for example, by 2D LC.
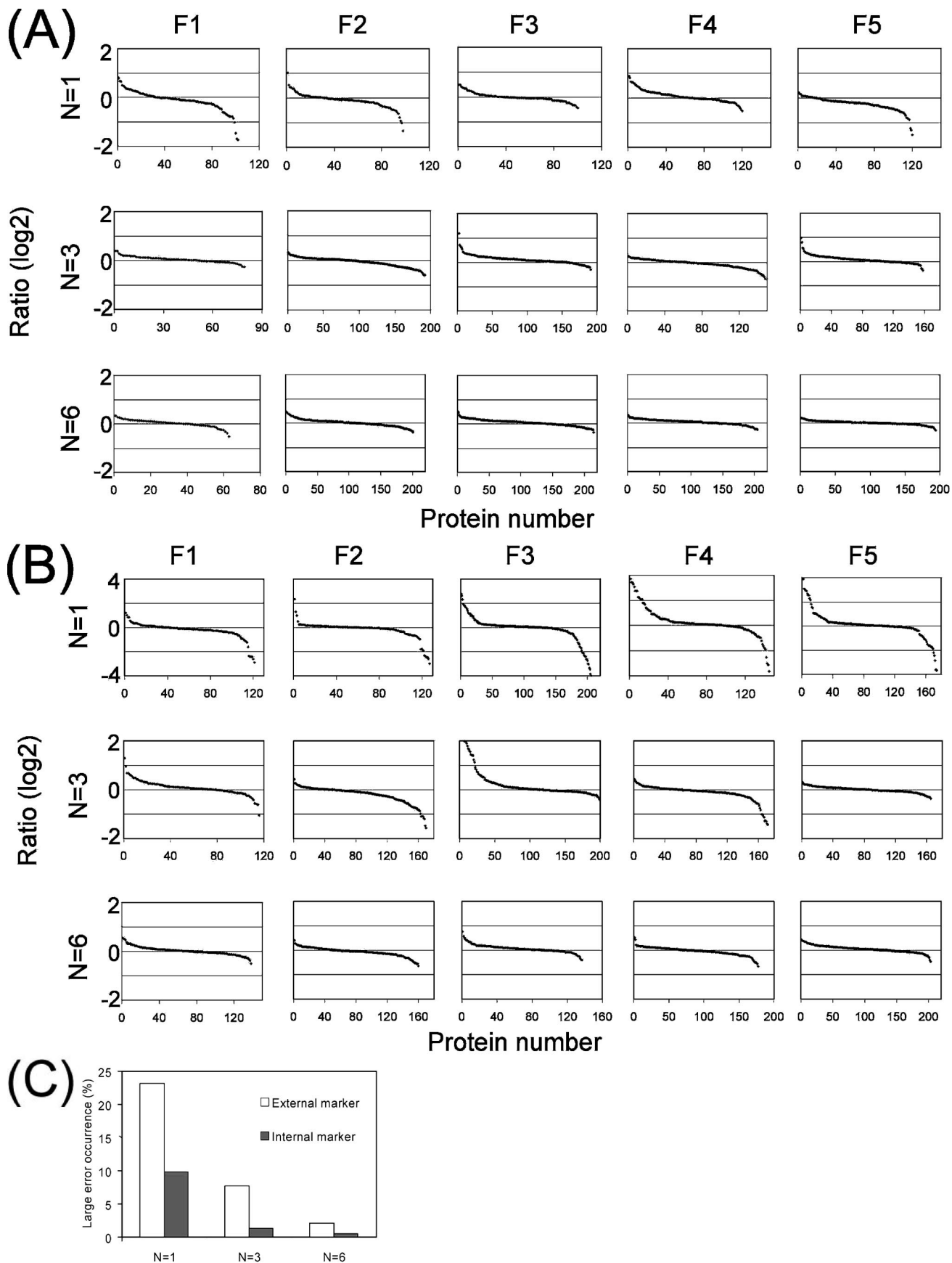
Figure 6. SILAC ratios of proteins with different numbers of replicates for gel cutting using (A) internal and (B) external markers. Proteins from each fraction (F1−F5) were analyzed by LC−MS/MS and quantified. Three replicate numbers for gel cutting were performed ($N = 1, 3,$ and 6). The sorted protein ratios (heavy/light) were plotted. (C) Percentage of proteins with large ratio errors (relative error > 30%) when different replicate numbers and markers were used.
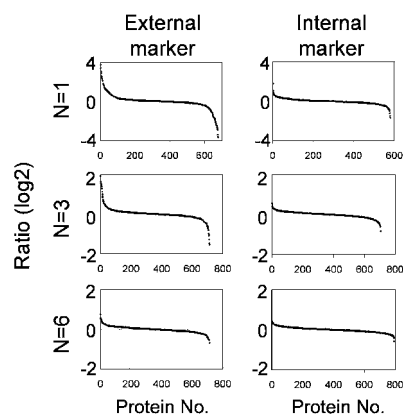
Figure 7. "Total sample-wise" protein quantitation after protein fractionation. For each identified protein, the signal intensities of identified peptides from all fractions of the sample were summed for quantitation. The protein ratio was calculated as the summed intensities of heavy peptides over light peptides. This was carried out for each replicate group (N = 1, 3, and 6) for both external and internal marker experiments.
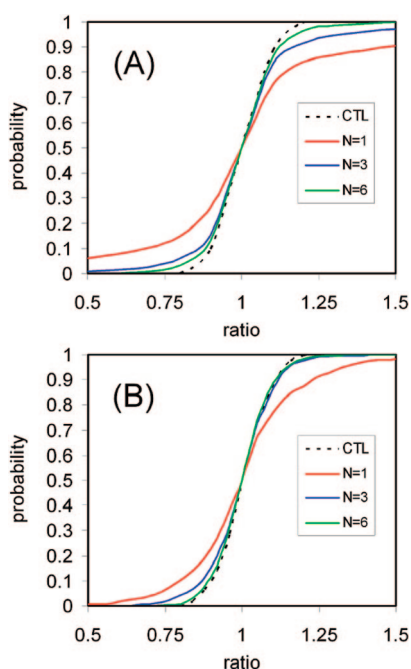

Figure 8. Cumulative probability curves for protein quantitation after fractionation using (A) external and (B) internal markers. Quantified proteins from the "total sample-wise" results of each replicate group (N = 1, 3, and 6) were used to plot the curves. For the control sample, a 1:1 mixture of light and heavy isotope-labeled lysates was digested and analyzed by LC–MS/MS.

As discussed in the Introduction, protein fractionation may provide advantages over peptide fractionation and our new method provides a good option for these applications.

In addition, our new method may be used in applications in which peptides are correlated between samples for quantitative purposes. For example, protein correlation profiling has shown great promises in organelle proteome characterization.[36–38] In a typical protein correlation profiling experiment, cellular fractions from sucrose-gradient purifications are analyzed by LC–MS. The abundance of proteins in each sucrose-gradient fraction is then calculated on the basis of the corresponding peptide ion intensities from each fraction. If GeLC–MS is used,

Table 1. Quantitative Errors for Protein Fractionation with Different Replicate Numbers[a]

	N	$E < 25\%$	$E < 50\%$
control		0.998	1.000
	1	0.851	0.975
internal marker	3	0.976	1.000
	6	0.988	1.000
	1	0.738	0.841
external marker	3	0.897	0.964
	6	0.964	0.998

[a] N, replicate number of gel cutting; E, quantitative error. Values in the table represent the probability that the quantitative error is less than 25% and 50% of the actual ratio.

irreproducible gel cutting can result in differential splitting of proteins into different fractions for different samples, which can complicate quantitation. Similarly, in temporal SILAC experiments, results from two GeLC–MS experiments are combined to obtain a complete time-course result.[39] In these cases, it is desirable to have consistent SDS–PAGE fractionation to improve quantitation accuracy and simplify data processing for correlation.

Another application of DNA ladders as a guide for accurate and reproducible gel cutting is when a low-abundance protein is found to be of interest (e.g., a potential disease biomarker) and a large number of samples need to be analyzed to validate the finding. A pair of DNA bands could be selected to precisely mark the location of the protein of interest, and only this narrow band might be analyzed from the large number of samples of the validation set, thus speeding up the analysis considerably.

Conclusions

We have developed a novel visible DNA-staining method and used this method to assist protein fractionation by SDS–PAGE for quantitative proteomics. The new DNA-staining method is sensitive, fast, and easy to use. Unlike protein molecular-weight markers, the DNA bands are used as internal markers to allow for maximum accuracy and reproducibility of gel cutting. In addition, visualization of these markers is completely independent of the protein samples analyzed, making it especially appealing for those applications in which only a limited amount of protein is available. When DNA markers are combined with a sufficient number of replicates of gel cutting, very reproducible fractionation and quantitation can be achieved.

Abbreviations: 1D, one-dimensional; 2D, two-dimensional; CBB, Coomassie Brilliant Blue; GeLC–MS, gel-enhanced liquid chromatography–mass spectrometry; IB, indoine blue; PMF, peptide mass fingerprinting; SILAC, stable isotope labeling with amino acids in cell culture.

References

(1) de Godoy, L. M.; Olsen, J. V.; de Souza, G. A.; Li, G.; Mortensen, P.; Mann, M. Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system. *Genome Biol.* **2006**, *7* (6), R50.
(2) Liu, H.; Sadygov, R. G.; Yates, J. R. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **2004**, *76* (14), 4193–4201.

(3) Pflieger, D.; Le Caer, J. P.; Lemaire, C.; Bernard, B. A.; Dujardin, G.; Rossier, J. Systematic identification of mitochondrial proteins by LC−MS/MS. *Anal. Chem.* **2002**, *74* (10), 2400–2406.

(4) Schirle, M.; Heurtier, M. A.; Kuster, B. Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography−tandem mass spectrometry. *Mol. Cell. Proteomics* **2003**, *2* (12), 1297–1305.

(5) Washburn, M. P.; Wolters, D.; Yates, J. R., III. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001**, *19* (3), 242–247.

(6) Gaucher, S. P.; Taylor, S. W.; Fahy, E.; Zhang, B.; Warnock, D. E.; Ghosh, S. S.; Gibson, B. W. Expanded coverage of the human heart mitochondrial proteome using multidimensional liquid chromatography coupled with tandem mass spectrometry. *J. Proteome Res.* **2004**, *3* (3), 495–505.

(7) Taylor, S. W.; Fahy, E.; Zhang, B.; Glenn, G. M.; Warnock, D. E.; Wiley, S.; Murphy, A. N.; Gaucher, S. P.; Capaldi, R. A.; Gibson, B. W.; Ghosh, S. S. Characterization of the human heart mitochondrial proteome. *Nat. Biotechnol.* **2003**, *21* (3), 281–286.

(8) Ishihama, Y.; Oda, Y.; Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.; Mann, M. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* **2005**, *4* (9), 1265–1272.

(9) Lu, P.; Vogel, C.; Wang, R.; Yao, X.; Marcotte, E. M. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* **2007**, *25* (1), 117–124.

(10) Rappsilber, J.; Ryder, U.; Lamond, A. I.; Mann, M. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* **2002**, *12* (8), 1231–1245.

(11) Callister, S. J.; Barry, R. C.; Adkins, J. N.; Johnson, E. T.; Qian, W. J.; Webb-Robertson, B. J.; Smith, R. D.; Lipton, M. S. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *J. Proteome Res.* **2006**, *5* (2), 277–286.

(12) Fang, R.; Elias, D. A.; Monroe, M. E.; Shen, Y.; McIntosh, M.; Wang, P.; Goddard, C. D.; Callister, S. J.; Moore, R. J.; Gorby, Y. A.; Adkins, J. N.; Fredrickson, J. K.; Lipton, M. S.; Smith, R. D. Differential label-free quantitative proteomic analysis of *Shewanella oneidensis* cultured under aerobic and suboxic conditions by accurate mass and time tag approach. *Mol. Cell. Proteomics* **2006**, *5* (4), 714–725.

(13) Fischer, B.; Grossmann, J.; Roth, V.; Gruissem, W.; Baginsky, S.; Buhmann, J. M. Semi-supervised LC/MS alignment for differential proteomics. *Bioinformatics* **2006**, *22* (14), e132–e140.

(14) Higgs, R. E.; Knierman, M. D.; Gelfanova, V.; Butler, J. P.; Hale, J. E. Comprehensive label-free method for the relative quantification of proteins from biological samples. *J. Proteome Res.* **2005**, *4* (4), 1442–1450.

(15) Le Bihan, T.; Goh, T.; Stewart, I. I.; Salter, A. M.; Bukhman, Y. V.; Dharsee, M.; Ewing, R.; Wisniewski, J. R. Differential analysis of membrane proteins in mouse fore- and hindbrain using a label-free approach. *J. Proteome Res.* **2006**, *5* (10), 2701–2710.

(16) Meng, F.; Paweletz, C. P.; Mazur, M. T.; Deyanova, E. G.; Yates, N. A.; Hendrickson, R. C.; Wiener, M. C.; Sachs, J. R.; Burns, C.; Verma, P. Quantitative analysis of complex peptide mixtures using FTMS and differential mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2006**, *24*, 24.

(17) Old, W. M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K. G.; Mendoza, A.; Sevinsky, J. R.; Resing, K. A.; Ahn, N. G. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol. Cell. Proteomics* **2005**, *4* (10), 1487–1502.

(18) Ono, M.; Shitashige, M.; Honda, K.; Isobe, T.; Kuwabara, H.; Matsuzuki, H.; Hirohashi, S.; Yamada, T. Label-free quantitative proteomics using large peptide data sets generated by nanoflow liquid chromatography and mass spectrometry. *Mol. Cell. Proteomics* **2006**, *5* (7), 1338–1347.

(19) Radulovic, D.; Jelveh, S.; Ryu, S.; Hamilton, T. G.; Foss, E.; Mao, Y.; Emili, A. Informatics platform for global proteomic profiling and biomarker discovery using liquid chromatography−tandem mass spectrometry. *Mol. Cell. Proteomics* **2004**, *3* (10), 984–997.

(20) Wang, G.; Wu, W. W.; Zeng, W.; Chou, C. L.; Shen, R. F. Label-free protein quantification using LC-coupled ion trap or FT mass spectrometry: Reproducibility, linearity, and application with complex proteomes. *J. Proteome Res.* **2006**, *5* (5), 1214–1223.

(21) Wang, W.; Zhou, H.; Lin, H.; Roy, S.; Shaler, T. A.; Hill, L. R.; Norton, S.; Kumar, P.; Anderle, M.; Becker, C. H. Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal. Chem.* **2003**, *75* (18), 4818–4826.

(22) Wiener, M. C.; Sachs, J. R.; Deyanova, E. G.; Yates, N. A. Differential mass spectrometry: A label-free LC−MS method for finding significant differences in complex peptide and protein mixtures. *Anal. Chem.* **2004**, *76* (20), 6085–6096.

(23) Scheler, C.; Lamer, S.; Pan, Z.; Li, X. P.; Salnikow, J.; Jungblut, P. Peptide mass fingerprint sequence coverage from differently stained proteins on two-dimensional electrophoresis patterns by matrix assisted laser desorption/ionization−mass spectrometry (MALDI−MS). *Electrophoresis* **1998**, *19* (6), 918–927.

(24) Patton, W. F. A thousand points of light: The application of fluorescence detection technologies to two-dimensional gel electrophoresis and proteomics. *Electrophoresis* **2000**, *21* (6), 1123–1144.

(25) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **1996**, *68* (5), 850–858.

(26) Flores, N.; Valle, F.; Bolivar, F.; Merino, E. Recovery of DNA from agarose gels stained with methylene blue. *BioTechniques* **1992**, *13* (2), 203–205.

(27) Adkins, S.; Burmeister, M. Visualization of DNA in agarose gels as migrating colored bands: Applications for preparative gels and educational demonstrations. *Anal. Biochem.* **1996**, *240* (1), 17–23.

(28) Yang, Y. I.; Hong, H. Y.; Lee, I. S.; Bai, D. G.; Yoo, G. S.; Choi, J. K. Detection of DNA using a visible dye, Nile blue, in electrophoresed gels. *Anal. Biochem.* **2000**, *280* (2), 322–324.

(29) Hwang, S. Y.; Jin, L. T.; Yoo, G. S.; Choi, J. K. Counterion-dye staining for DNA in electrophoresed gels using indoine blue and methyl orange. *Electrophoresis* **2006**, *27* (9), 1739–1743.

(30) Jin, L. T.; Choi, J. K. Usefulness of visible dyes for the staining of protein or DNA in electrophoresis. *Electrophoresis* **2004**, *25* (15), 2429–2438.

(31) Hendrix, L. R.; Samuel, J. E.; Mallavia, L. P. Differentiation of *Coxiella burnetii* isolates by analysis of restriction-endonuclease-digested DNA separated by SDS−PAGE. *J. Gen. Microbiol.* **1991**, *137* (2), 269–276.

(32) Castellanos-Serra, L.; Hardy, E. Detection of biomolecules in electrophoresis gels with salts of imidazole and zinc II: A decade of research. *Electrophoresis* **2001**, *22* (5), 864–873.

(33) Chelius, D.; Zhang, T.; Wang, G.; Shen, R. F. Global protein identification and quantification technology using two-dimensional liquid chromatography nanospray mass spectrometry. *Anal. Chem.* **2003**, *75* (23), 6658–6665.

(34) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3* (12), 1154–1169.

(35) Mirgorodskaya, O. A.; Kozmin, Y. P.; Titov, M. I.; Korner, R.; Sonksen, C. P.; Roepstorff, P. Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using ^{18}O-labeled internal standards. *Rapid Commun. Mass Spectrom.* **2000**, *14* (14), 1226–1232.

(36) Andersen, J. S.; Lam, Y. W.; Leung, A. K.; Ong, S. E.; Lyon, C. E.; Lamond, A. I.; Mann, M. Nucleolar proteome dynamics. *Nature* **2005**, *433* (7021), 77–83.

(37) Andersen, J. S.; Mann, M. Organellar proteomics: Turning inventories into insights. *EMBO Rep.* **2006**, *7* (9), 874–879.

(38) Andersen, J. S.; Wilkinson, C. J.; Mayor, T.; Mortensen, P.; Nigg, E. A.; Mann, M. Proteomic characterization of the human centrosome by protein correlation profiling. *Nature* **2003**, *426* (6966), 570–574.

(39) Blagoev, B.; Ong, S. E.; Kratchmarova, I.; Mann, M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* **2004**, *22* (9), 1139–1145.

PR700601Y