Journal of proteome research

Article

pubs.acs.org/jpr

An Efficient Approach to Evaluate Reporter Ion Behavior from MALDI-MS/MS Data for Quantification Studies Using Isobaric Tags

Stephanie M. Cologna,[†] Christopher A. Crutchfield,[‡,⊥] Brian C. Searle,[§] Paul S. Blank,[∥] Cynthia L. Toth,[†] Alexa M. Ely,[†] Jaqueline A. Picache,[†] Peter S. Backlund,[‡] Christopher A. Wassif,[†] Forbes D. Porter,[†] and Alfred L. Yergey*[,‡]

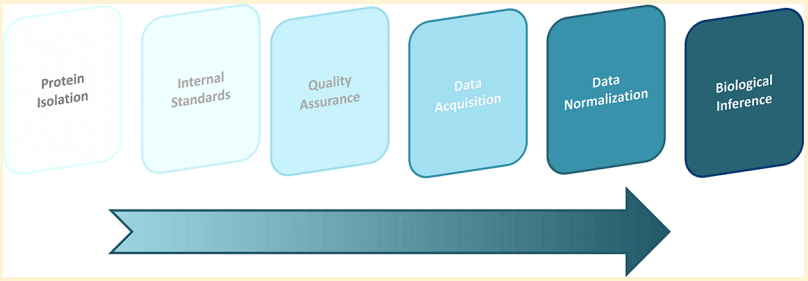
[†]Program on Developmental Endocrinology and Genetics, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, DHHS, 31 Center Drive, Bethesda, Maryland 20892, United States

[‡]Biomedical Mass Spectrometry Facility, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, DHHS, 31 Center Drive, Bethesda, Maryland 20892, United States

[§]Proteome Software, Inc., 1340 SW Bertha Boulevard, Portland, Oregon 97219, United States

[∥]Section on Membrane and Cellular Biophysics, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, DHHS, 31 Center Drive, Bethesda, Maryland 20892, United States

Supporting Information



ABSTRACT: Protein quantification, identification, and abundance determination are important aspects of proteome characterization and are crucial in understanding biological mechanisms and human diseases. Different strategies are available to quantify proteins using mass spectrometric detection, and most are performed at the peptide level and include both targeted and untargeted methodologies. Discovery-based or untargeted approaches oftentimes use covalent tagging strategies (i.e., iTRAQ, TMT), where reporter ion signals collected in the tandem MS experiment are used for quantification. Herein we investigate the behavior of the iTRAQ 8-plex chemistry using MALDI-TOF/TOF instrumentation. The experimental design and data analysis approach described is simple and straightforward, which allows researchers to optimize data collection and proper analysis within a laboratory. iTRAQ reporter ion signals were normalized within each spectrum to remove peptide biases. An advantage of this approach is that missing reporter ion values can be accepted for purposes of protein identification and quantification without the need for ANOVA analysis. We investigate the distribution of reporter ion peak areas in an equimolar system and a mock biological system and provide recommendations for establishing fold-change cutoff values at the peptide level for iTRAQ data sets. These data provide a unique data set available to the community for informatics training and analysis.

KEYWORDS: *MALDI-MS, time-of-flight, iTRAQ, quantitative proteomics, quality control, fold-change*

INTRODUCTION

Quantitative analysis of proteins has always been an important component of biological and biomedical research, but the tools available for the task have not always been generally applicable for reasons of sensitivity or experimental difficulty. Recently a number of mass spectrometry-based strategies have emerged for quantitative analyses of proteins; these strategies include approaches ranging from covalent labeling to label-free experiments (as reviewed in ref 1). Of particular interest to this work are covalent labeling strategies such as those commercially available including isobaric tags for relative and absolute quantitation (iTRAQ, ABSciex) and tandem mass tags (TMT, ThermoScientific), where quantification is performed in the tandem MS (MS/MS or MS^2) experiment. Using isobaric tags to label the peptides does not change the properties of a peptide between samples; however, it generates a series of "reporter ions" that can be used to quantify different biological systems or states. The original iTRAQ report was a 4-plex strategy,[2] which was then extended to an 8-plex system.[3] Similarly, TMT was introduced[4] and is commercially available as a 6-plex quantification kit. TMT has recently been extended[5] with a commercially available kit capable of 10-plex quantification. It should be noted that the 10-plex TMT

Received: March 25, 2015

DOI: 10.1021/acs.jproteome.5b00254
J. Proteome Res. XXXX, XXX, XXX−XXX

platform requires high resolving power instruments (e.g., Orbitrap, FTMS mass analyzer) to take advantage of the mass differences observed by the incorporation of the heavy carbon versus nitrogen. A recent report from Gygi et al., further extends the multiplexing capabilities of TMT chemistry to 54, a development that may promise a substantial improvement in throughput.[6] It is evident that the field of covalent-labeling -based quantitative proteomics is continuing to expand.

Using isobaric tags for quantification during tandem MS experiments is straightforward, and iTRAQ-based relative quantification studies have successfully been achieved on samples ranging from simple model systems to cultured cells, tissue homogenates, and other specimens of clinical importance. Several technical and quantitative aspects of iTRAQ-based studies have recently been reviewed,[7] and much attention has been given to attempt to overcome the inherent analytical and technical challenges.[8] These areas of interest include sample preparation,[9] coeluting peptides,[10] reporter ion suppression,[11] instrumental operating parameters,[12] and proper data interpretation, normalization, statistics, and modeling.[13] Many of the early reports on the evaluation of iTRAQ performance were focused on the use of the 4-plex system, while attention is now on the 8-plex product given the increase in multiplexing capabilities. The chemistry of the labeling is identical between the two products; however, the 8-plex label requires an increased mass label to accommodate the balance and reporter group structures necessary for eight reporter ions to be generated. This advancement allows for more sophisticated experimental designs to include replicates or extensive time-course studies within the same experiment. It should be noted that the approaches presented herein are equally applicable to all isobaric tag labeling strategies.

With regard to quality control and understanding the variability of iTRAQ-based studies, Ow et al. reported that the iTRAQ reagents could provide appropriate quantification over 2 orders of magnitude with Q-ToF instrumentation using a synthetic standard labeled and the 8-plex chemistry.[8c] At the same time, isotopic impurity of the iTRAQ reagents is a significant source of variability, and it has become accepted that an underestimation is present, most noticeably for the m/z 121.1 tag.[10,8c] Robust data analysis in light of this complexity is daunting. In particular, establishing fold-change cut-off values is one area that still lacks consensus.

In the current study, we have designed several large-scale experiments to examine sources of variability, monitor the data quality of the iTRAQ reporter method, and consider these results when making biologically relevant conclusions. The data reported herein have been processed using within spectrum-based normalization using iTRAQ reporter ion intensities rather than the more typical global reporter-based normalization. We also consider the incorporation of an internal standard and approaches to evaluate data integrity. A mock biological sample was generated and analyzed to evaluate feasibility of this approach for biological or clinically significant studies. Finally, we have established a simple and robust method to determine fold-change values with robust statistical significance.

MATERIALS AND METHODS

All chemicals were obtained from Sigma-Aldrich (St. Louis, MO) unless otherwise stated. Chemicals were used as received without further purification. Standard laboratory safety measures were followed.

Cell Culture and Protein Preparation

All Gibco cell culture materials were obtained from Life Technologies (Grand Island, NY) unless otherwise noted. HepG2 cells were cultured in 10% fetal bovine serum (FBS) containing DMEM media, supplemented with PenStrep, and L-glutamine. Cell pellets were washed with 1X phosphate buffered saline a total of three times. The resulting cell pellet (~5 million cells) was then resuspended in 1 mL of 500 mM triethylammonium bicarbonate buffer containing 0.1% SDS (K-D Medical, Columbia, MD), pH 8.5 and lysed by sonication on ice. The cell lysate solution was centrifuged at 20,000g x 30 min at 4 °C. The supernatant was removed and the protein lysate was quantified in triplicate via the Bradford assay (Bio-Rad, Hercules, CA). Protein lysate aliquots were stored at −30 °C until use. Approximate protein recovery was 10 mg of protein per 5 million cells. For the complex biological system, human cerebrospinal fluid (CSF) was collected from consenting adults and obtained from Biochemed (Winchester, VA). A pooled sample consisting of CSF from 8 individuals was generated, and human serum albumin was depleted using an Agilent (Santa Clara, CA) HSA antibody depletion column. The resulting depleted sample was quantified using the Bradford assay, and 10 μg of the CSF proteome was used as a complex biological matrix. Predigested alcohol dehydrogenase was added to the mixture and protein digestion and labeling were carried out as noted below.

Protein Digestion and iTRAQ Labeling

All iTRAQ 8-plex labeling experiments were conducted using reagents from the same lot number for the HepG2 study. (Separate lot numbers were used for the other studies.) For each experiment, the total protein concentration and volume was kept constant. One hundred micrograms (10 μg for CSF experiment) of protein per tube (total of eight tubes per experiment) in a total volume of 20 μL was reduced using tris-carboxyethyl phosphine and alkylated using methanemethyl thiosulfonate using the iTRAQ Chemistry Kit (AB Sciex, Framingham, MA) according to the manufacturer's recommendation. Trypsin (Promega, Madison, WI) was added at a 1:50 (w/w) enzyme to protein ratio, and the solution was incubated at 37 °C overnight. An additional aliquot of trypsin (1:100 ratio) was then added, and the solution was further incubated for 4 h at 37 °C. Protein digestion was verified by removing 0.5 μL of the digest solution diluted in 0.1% trifluoroacetic acid (TFA), desalted and concentrated using C_{18} ZipTips (Millipore, Billerica, MA) according to the manufacturer's protocol. One microliter from each sample was then mixed with 1 μL of α-cyano-4-hydroxycinnamic acid (5 mg mL^{-1} in 70% acetonitrile (ACN)/0.1% TFA, 10 mM ammonium dibasic citrate), and MS data were collected on a model 4800 MALDI-TOF/TOF Proteomics Analyzer (Applied Biosystems, Framingham, MA). For testing, MS spectra were collected in positive ion reflectron mode with external calibration. For the HepG2 experiments, the addition of predigested yeast alcohol dehydrogenase (Waters, Milford, MA) at a concentration of 50 fmol/μg protein was included, and each tube was subsequently labeled with the appropriate iTRAQ label. Prior to mixing, each of the labeling experiments was confirmed by removing 1 μL and testing by MALDI-MS/MS as previously described. The eight samples were then combined equally for HepG2, CSF or in desired ratios for BSA). The samples were dried and resuspended in 10 mM KH_2PO_4 in 25% ACN, pH 3. The mixture was fractionated

DOI: 10.1021/acs.jproteome.5b00254
J. Proteome Res. XXXX, XXX, XXX−XXX

using strong cation exchange either by single elution (BSA), none (CSF), or by step-gradient (HepG2, ten fractions total) using 10 mM KH_2PO_4 in 25% ACN, pH 3 containing KCl. Each fraction was further desalted and concentrated using C_{18} Omix tips (Agilent Technologies, Santa Clara, CA), dried and resuspended in 0.1% TFA, then stored at −30 °C until analysis.

LC−MALDI-MS/MS and Data Analysis

Approximately 1 μg of digested and labeled peptides from each SCX fraction was injected onto a model SYS0201UPLC (Scientific Systems, State College, PA) outfitted with an external splitter and equipped with a C_{18} BioBasic column (ThermoScientific, Waltham, MA, 100 × 0.18 mm, 5 μm particle size) with a resulting flow rate of ∼1.5 μL/min. Solvent A was 0.1% TFA/2% ACN. Solvent B contained 85% ACN, 5% isopropanol, and 0.1% TFA. The gradient conditions were as follows: 2% B for 5 min, 5−35% B for 60 min, 35−60% B for 35 min, 60−85% B for 20 min, 90% B × 5 min, followed by equilibration. The LC eluent was spotted onto a MALDI target every 8 s for a total of 840 spots using a Dionex ProBot (ThermoScientific, Waltham, MA) coupled to a mixing "tee" to introduce the MALDI matrix (5 mg mL^{-1} in 70% ACN/0.1% TFA, 10 mM ammonium dibasic citrate) that was set at a flow rate of 1.8 μL min^{-1}. MALDI-MS and MS/MS spectra were collected in the positive ion reflector mode. MS spectra were collected from 700 to 4500 Da using external calibration. The 10 most intense peaks in each spot were subjected to fragmentation. All MS and MS/MS data were extracted using the T2D Extractor (courtesy of Phillip C. Andrews, University of Michigan). Qualitative and quantitative data analysis was carried out using scripts written for execution in R [R version 3.1.0 (2014−04−10) Copyright 2014 The R Foundation for Statistical Computing Platform: x86_64-apple-darwin13.1.0 (64-bit)] developed specifically for this work. For data analysis of peptide assignments, Mascot[14] was used as the search engine with the SwissProt Database using taxonomy: human, mammalian or yeast; fixed modifications: iTRAQ (N-term, K); and variable modifications: MMTS (C), Oxidation (M), and iTRAQ (Y). Mass error tolerances were set to 100 ppm for MS spectra and 0.4 Da for MS/MS data. Peptide and protein assignments were obtained using Scaffold v4.4.1.1.[15] Data were submitted using the .dat file obtained from Mascot, and searching criteria were kept constant. Scaffold incorporates the use of the Peptide Prophet[16] and Protein Prophet[17] algorithms. Protein assignments were made at the 95% confidence level with a minimum of two peptides at the 99% confidence level. Scripts utilized in this work are available upon request.

RESULTS

The goals of the enclosed work are 3-fold: (i) to generate a robust experimental platform to evaluate sample processing and data quality during biological experiments, (ii) to use a complex system to establish the variability of iTRAQ reporter ion technology on a matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) platform, and (iii) to develop an unbiased modeling approach of experimental data to establish nonarbitrary, fold-change cut off values for biological interpretation. Together, these studies should provide an experimental platform capable of making strong biological conclusions using covalent-labeling strategies for quantitative proteomic studies regardless of the chemistry being used.

Experimental Design Considerations

To execute a robust quantitative experiment, one must rely on a strong and well-conceived experimental design. Therefore, a platform has been established that incorporates a means to verify proper processing at several steps during the experiment. As depicted in Figure 1, the typical bottom-up proteomics approach is used. In addition to the typical sample preparation procedure, a predigested exogenous protein is included to allow for evaluation of iTRAQ labeling during data analysis. Typically our approach is to verify tryptic digestion and iTRAQ labeling in each sample prior to final mixing, fractionation and LC−MS/MS analysis. While other reports have commented and discussed the variability associated with covalent-labeling approaches such as iTRAQ, large-scale evaluations on this analytical platform have been lacking. In this study, we have used a human liver cancer cell line (hepatocellular carcinoma, HepG2) to model a complex system. Figure 1 shows the experimental approach used where a cell lysate is distributed into eight individual samples at equal concentration. The eight samples are then reduced, alkylated and proteolytically digested using trypsin according to standard protocols provided by the manufacturer. The resulting peptides are labeled with each of the eight iTRAQ reagents, then combined, SCX-fractionated, and analyzed via LC−MALDI-TOF-MS/MS. Four replicate experiments (e.g., digestion, labeling, and separation) were carried out in which these data were used for all subsequent analysis.

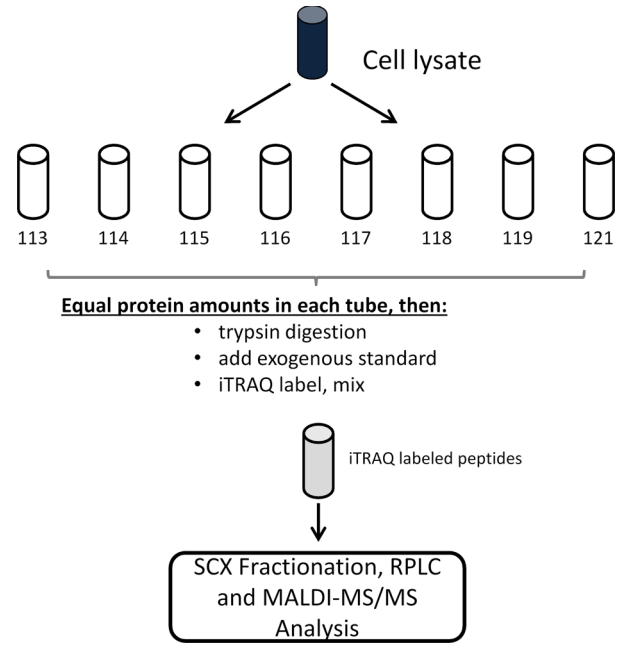


Figure 1. Experimental design for iTRAQ experiments. A single cell lysate supernatant was distributed equally (100 μg/tube) into eight tubes. Reduction, alkylation, and digestion were carried out identically for all eight samples in individual reactions. An exogenous protein standard was added. iTRAQ labeling was carried out using standard procedures; then, the peptide mixture was fractionated by SCX and analyzed via LC−MALDI-MS/MS.

Evaluation of Reporter Ion Intensity Value Distributions for All Reporters

Tandem mass spectra (MS/MS) were extracted as Mascot Generic Files (.mgf) for all SCX fractions from each replicate.

DOI: 10.1021/acs.jproteome.5b00254
J. Proteome Res. XXXX, XXX, XXX−XXX

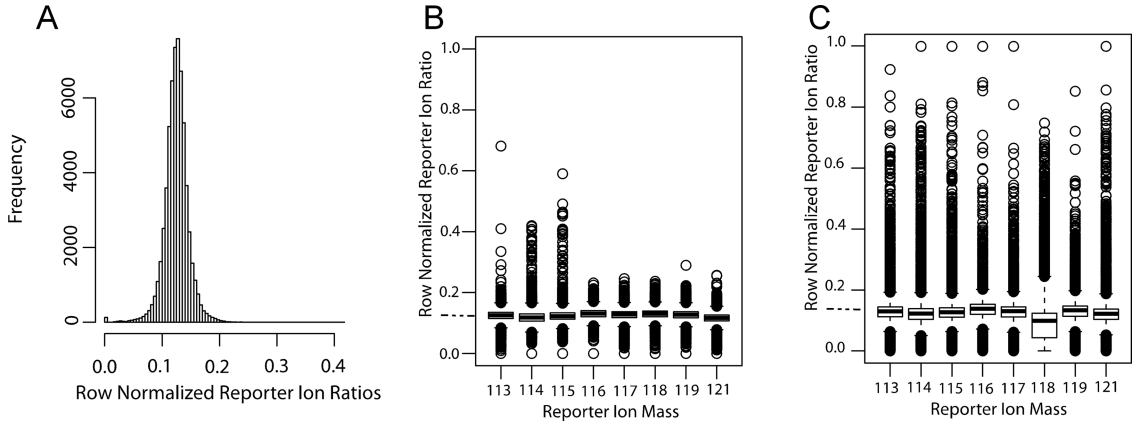


Figure 2. Evaluation of normalized reporter ion intensities from HepG2 cell lysate protein digests. (A) Distribution of iTRAQ reporter ion intensity values for all precursor ions with an area row sum cutoff of 5000. Data presented are from replicate analysis #3. (B) Box and whisker plot of normalized iTRAQ reporter ions broken out by individual reporter ion masses. Again, these data are from replicate analysis #3. (C) Box and whisker plot of row-normalized iTRAQ reporter ion area values (from Replicate #1) broken out by individual reporter ion masses. These data represent the one replicate in which one iTRAQ reporter ion did not display concomitant results with the other reporter ions, representing 1 labeling reaction out of 32 performed in this study.

The iTRAQ reporter ion area values (noted here on as intensity for sake of clarity) for each MS/MS spectrum were parsed from the larger files. All spectra were initially considered regardless of whether a peptide assignment could be made. Data were filtered by removal of MS/MS spectra for which the sums of the intensity of all reporter ions did not exceed 5000 units. The inclusion of a cutoff was determined based on manual inspection of several data sets prior to the initiation of this study. As expected, some spectra were not included due to this cutoff, typically ∼10%. The data matrices were organized as nine columns, the first being the m/z of the precursor and the next eight being the reporter ion intensities. We subsequently refer to "row sums" to mean the sum of these intensities for each precursor m/z. Previous studies have shown that low-intensity iTRAQ reporter ion abundances and poor MS/MS quality can increase quantitative variance.[8a,18] This data-processing approach ensured that only high-quality MS/MS spectra were used for analysis and interpretation.

Once the data sets were filtered, the individual reporter ion intensities for each precursor ion were normalized to the sum of intensities for that precursor m/z. We denote this as "row normalization". This approach allows for comparison within a tandem mass spectrum and does not rely on normalization across an individual reporter ion; that is, what would correspond to a column sum in our matrix. Performing row normalization is consistent with our observations and others[8a] that fragmentation efficiency is peptide-dependent. On the contrary, normalization based on the "column sum", assumes that all reporter ions behave the same during the fragmentation experiment regardless of the peptide ion being fragmented. The row normalization process does not make the assumption that fragmentation of the iTRAQ label is consistent for a given label; rather fragmentation efficiency is expected to be consistent for all reporter ions with a single precursor ion. All data presented subsequently will be noted as "normalized intensity", referring to the row normalization procedure that is by definition a ratio as described later. Figure 2A is a histogram displaying the distribution of normalized reporter ion intensities for all spectra collected in one of the replicate analyses. Data points observed with normalized intensities of zero result from the row sum passing filtering criteria; however, zero intensity remains in at least one reporter ion. The total number of spectra processed for this data set (termed Replicate 3) was 8592, of which 7666 met the filtering criteria (89.2%). Because each analyte in the set of eight replicates is present in equal concentrations, all reporter ion ratios are expected to be one-eighth of the total signal or 0.125. The mean value for this distribution was found to be 0.125 with a standard deviation of 0.022. It is important to note that these data are normally distributed, as determined by the Kolmogorov−Smirnov test (replicate 3, $p = 2.2 \times 10^{-16}$), without log-transformation. The ability to employ a nonlog transformed distribution is an important attribute given that typical parametric statistics are based on the assumption that the data are normally distributed. Recent data suggest that many data sets collected on some instrumental platforms resulted in the acquisition of data that were not normally distributed and, upon log-transformation, led to a truncated Gaussian distribution.[19] To further evaluate the data quality, we chose to extract the normalized intensities for each of the iTRAQ labels. Shown in Figure 2B is a box and whisker plot of the normalized intensities for each of the eight reporter ions from replicate three. The box represents the upper and lower 25% quartiles, whereas the whiskers represent maximum values excluding outliers. The dark black line indicates the median for each group. The median values were as follows for each iTRAQ reporter: m/z 113 = 0.126, m/z 114 = 0.118, m/z 115 = 0.122, m/z 116 = 0.131, m/z 117 = 0.128, m/z 118 = 0.130, m/z 119 = 0.128, and m/z 121 = 0.117. We observe the m/z 121 data to be the lowest of the reporter ions, which is consistent with other reports.[8c] Box and whisker plots were made for each of the four experimental replicates, and we observed a deviation for the m/z 118.1 iTRAQ label for one of the replicates, as displayed in Figure 2C. The median normalized intensity value for the 118.1 m/z reporter ions was found to be 0.120 ± 0.026. The mean and standard deviation for the entire normalized data set were 0.125 ± 0.039, and there were 8901 precursor masses that were successfully filtered from a full set of 9581 precursors (93%). Of the 32 labeling experiments that were carried out for this study, only one failed (termed replicate 1, Figure 2C). We determined this as a failed experiment in that the reporter ion intensities for the 118 label were significantly lower than the other reporter ions within the same spectrum.

DOI: 10.1021/acs.jproteome.5b00254
J. Proteome Res. XXXX, XXX, XXX−XXX

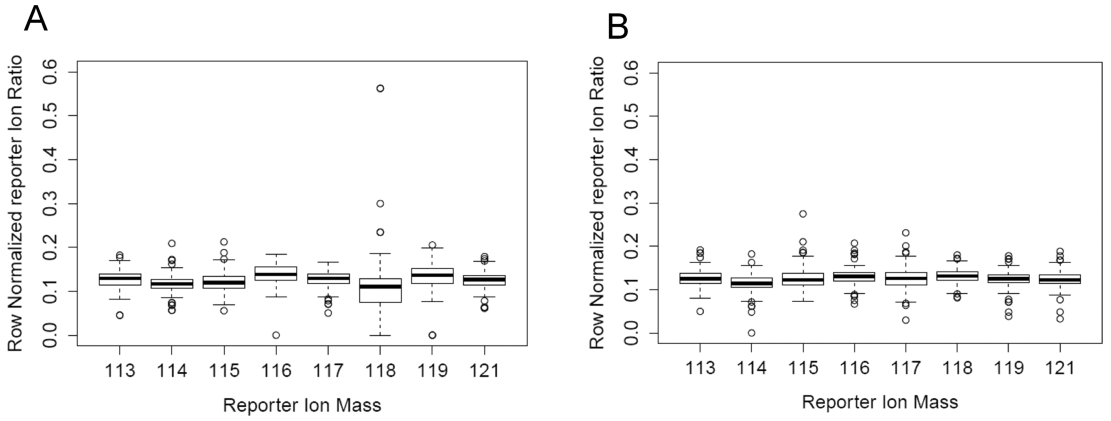


Figure 3. Distributions of normalized iTRAQ reporter ion intensities for internal standard peptides. Alcohol dehydrogenase was considered to be an internal standard to include in these studies. The sample was predigested with trypsin prior to spiking into the complex mixture. Box and whisker plots display the distribution of normalized iTRAQ reporter ion intensities values for the same data sets shown in Figure 2. Data presented from (A) replicate 1 and (B) replicate 3. Data were isotope-corrected prior to additional processing.

That is to say that labeling did occur but not completely. Manual evaluation of the spectra displayed reduced ion intensities at m/z 118 in the majority of the spectra collected. It is important to note that these data were not corrected for isotope impurity or overlap. The commercial reagents are noted to have >98% isotopic purity; however, to evaluate the effect of isotopic impurity on the normalized reporter ion intensities, we utilized the i-Tracker algorithm[20] and applied the corrections to our data set. The results for the isotope-corrected analysis of data from Replicates 1 and 3 are presented in Supplemental Figure 1. The data in Figure 2 and Figure S1 display similar trends with regard to the median row-normalized values. Therefore, the effect of isotope correction, while important and necessary, does not result in significant differences, specifically with regard to the 118 reporter ion values.

Internal Standard Considerations

Given that we have incorporated a predigested exogenous protein into the experimental design, the fragmentation spectra corresponding to peptides from yeast alcohol dehydrogenase were extracted from the .mgf files for each data set. Figure 3A contains the box and whisker plot of 119 spectra (for Replicate 1) and Figure 3B 126 spectra (for Replicate 3) assigned via Mascot and Scaffold to alcohol dehydrogenase. These data were first isotope-corrected and are the internal standard data from the same data sets noted in Figure 2C,B. From these data, the iTRAQ reporter ion displayed a reduced mean of normalized area m/z 118 reporter ions for the first replicate. In this case, a deviation is not observed for any of the reporter ion intensities. A similar approach by Karp et al.,[8a] suggests the inclusion of internal standard peptides to provide a means to implement a correction factor for each label. The most critical aspect of this discussion is to emphasize the importance of an experimenter evaluating their data prior to its use in biological interpretations.

Establishing Biologically Significant Cutoff Values for Protein Fold-Changes

In Figure 2 we have shown that, in an experimental system in which all proteins are expected to be at the same levels for each of the reporter ion mass tags, reporter ion ratios for all precursors behave as expected when normalized. We have also shown that this approach might well be used to detect otherwise unexpected errors in labeling chemistry. While important from the point of view of experimental reliability, that approach does not allow one to predict the sensitivity for detecting changes in reporter ion intensities and thus cannot be used to set limits for detection of significant changes in reporter ion intensities at the peptide or protein level. To accomplish this in an unbiased fashion, we have employed an approach analogous to one used in qRT-PCR experiments.[21]

The data collected from all four replicate runs of the HepG2 lysate were used. The iTRAQ reporter ion intensities were extracted (after normalization), and the resulting matrix contained the iTRAQ reporter ion intensities for each precursor ion that was fragmented. We now use the intensity of a given reporter ion to calculate a ratio of it to the other seven reporter ion area values *for each precursor* ion, (i.e., each row); this is done for each of the eight reporter ions. Prior to this calculation, however, *all* reporter ions having a zero area value must be removed from the data array because forming ratios where zeroes could be in the denominator cannot be allowed; this results in the loss of <10% of the total precursor ions from subsequent analysis, similar to the loss of precursors where area row sums are <5000. Precursors for which no iTRAQ reporter ion intensity values are reported are likely not peptidic in nature or are modified in such a manner that a reactive group is not accessible. The result is an array with eight columns and a number of rows corresponding to eight times the modified number of precursor ions. This array, which we term "EACH to ALL" needs an additional step of processing before it can be interpreted. That is, all of the reporter ion self-ratios, for example, 113/113, 114/114, 115/115, 116/116, and so on, must be removed to not bias the overall distribution of ratios in favor of unity.

Ideally one would expect that all of the ratios, excluding the self-ratios, would be the same and equal to one because each labeled sample consists of the same proteins present in the same amounts. Any variation from unity represents a measure of the variance in the system; however, evaluation of such data results in a non-normal distribution because the ratios are truncated at zero. Once the data are log-transformed, the transformation stabilizes the variance for subsequent analyses. One could also consider fitting the data to a zero-truncated Gaussian as an alternative. Given that the $\log_{10}(1) = 0$, then a $\log_{10}$ transform of the ratios for any replicate will show the range of variance for the system and thus allow the estimation

DOI: 10.1021/acs.jproteome.5b00254
J. Proteome Res. XXXX, XXX, XXX−XXX

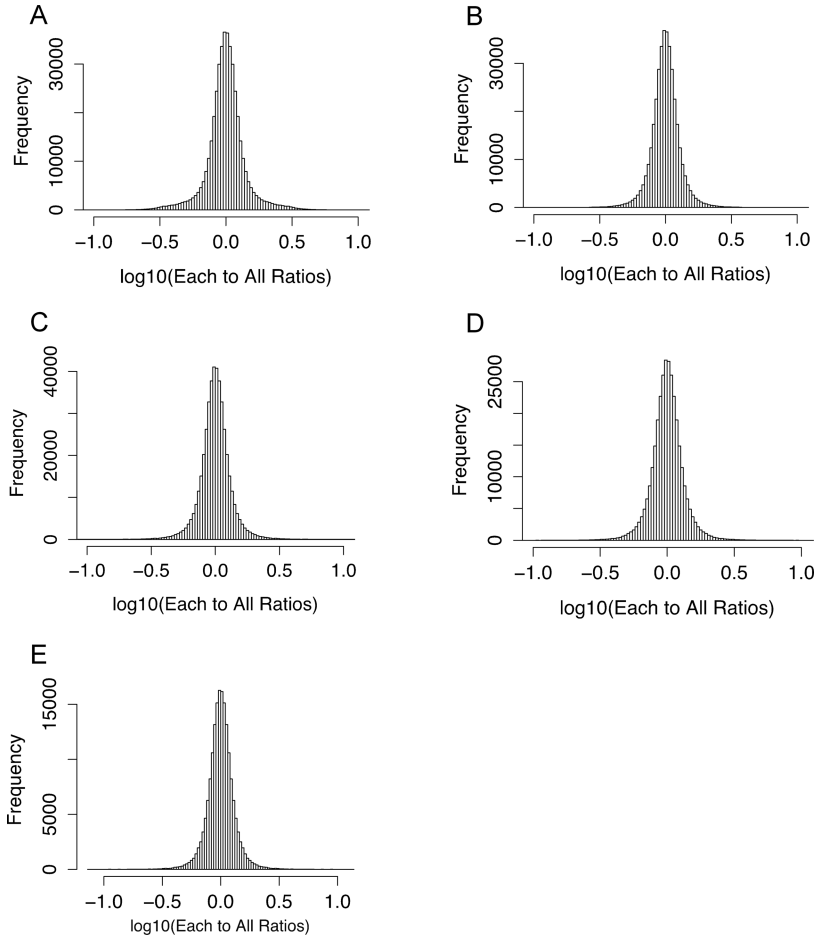


Figure 4. Comparison of all combinations of iTRAQ reporter ions. All pairwise combinations were considered for iTRAQ reporter ion area values collected from all four technical replicate experiments (A−D) and termed EACH to ALL. Data are presented in Log_{10} space; therefore, theoretical mean values for each distribution would be zero. (E) Behavior of iTRAQ reporter ions generated from precursors with peptide assignments. MS/MS spectra collected from the Replicate 3 experiment in which peptide assignments could be made in both Mascot and Scaffold were considered. The iTRAQ reporter ion area values after row-normalization were compared with one another in a pairwise fashion. The distribution is shown in Log_{10} space, where an expected mean would be at 0.

of a 95% confidence interval for the reporter ion ratios. Figure 4A−D displays the distribution of all pairwise comparisons from each of the replicates. These data once again show a normal distribution with mean values of each replicate centered on zero. Further investigation of these data allows calculation of the 95% confidence intervals which for Replicate 3 are 0.25 and −0.51 log_{10} fold-change ratios with similar values observed for the other three technical analyses, Replicate 1:0.26 and −0.54, Replicate 2:0.21 and −0.35, and Replicate 4:0.22 and −0.38. That is to say, in an experiment where proteins are *not* expected to be at equal concentrations, an iTRAQ reporter ion that is observed to be equal to or outside of these limits can be considered a candidate to determine peptide and downstream protein fold-change values that is significant. As such, we argue based on these observations, that this is an unbiased approach that uses spectra to establish the iTRAQ variance regardless of whether a peptide assignment is made.

To further refine and assess the data, we considered only the evaluation of fragmentation spectra in which a peptide assignment could be made. To accomplish this, we further processed data from the third replicate, where all precursor ions were initially considered This data set contained 8596 precursor masses that were selected for fragmentation. The .mgf file was submitted to the Mascot search engine, and the resulting protein identification list was generated then submitted to the Scaffold protein identification software. Protein identifications at the 99% confidence interval were considered in Scaffold, with a minimum of two peptides identified at the 95% confidence interval. From the Scaffold results, 3274 precursors from the Replicate 3 data set were used. These masses were then evaluated for the quality of the iTRAQ reporter ions. We removed spectra in which the sum of all reporter ion areas was below 5000 and those in which no iTRAQ reporter ions were present. This data processing resulted in a total of 3039 precursor ions being evaluated, all of which a peptide and subsequent protein assignment could be made. The distribution of these data is presented in Figure 4E. Using only peptides with high confidence assignments ensures that the MS/MS quality is high, thus yielding the minimal variance that could be obtained in a covalent labeling experiment. In Log_{10} space, the mean of the distribution was $5.81e^{-18}$, that is, essentially zero, as expected. The standard deviation was ±1.29. Therefore, considering two standard deviations of the distribution, or the 95% confidence intervals, the log_{10} fold change values identified in this exercise were 0.22 and −0.22. In smaller data sets, we observed a narrower distribution; however, this may be a consequence of a small number of data points. Within this data set, these data suggest

DOI: 10.1021/acs.jproteome.5b00254
J. Proteome Res. XXXX, XXX, XXX−XXX

that peptides used for establishing quantitative fold-changes of proteins are significant only above and below the noted values. In studies by Gan et al.,[18b] using the 4-plex chemistry and commercial software, a recommendation of ±50% as a cutoff for biological experiments is suggested; however, our data set appears to suggest more stringent requirements. As a general criterion, blanket cutoff values should not be implemented, and these values should be determined for the instrumentation within a specific laboratory.

Mock Biological Sample Analysis

A final study was performed to evaluate the potential of including a new normalization scheme into a biologically relevant analysis. Bovine serum albumin was digested using trypsin and labeled with the iTRAQ 8-plex kit. Aliquots from each labeling reaction were then mixed to generate 1:1, 1:2, 1:3, and 1:4 ratios. The ratios were generated in duplicate, therefore utilizing all eight labels. Additionally, a secondary experiment with independent protein digestion reactions and iTRAQ labeling was carried out. The data are provided in Figure 5A,B.

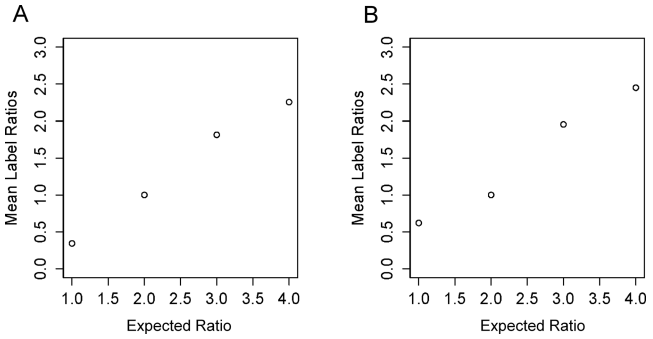


Figure 5. Technical replicate plots of mock biological data using bovine serum albumin tryptic digests. Tryptic digests were labeled with iTRAQ reporter ions and in duplicate mixed to generate 1:1, 1:2, 1:3, and 1:4 ratios. The duplicate median normalized intensities were then used, and the 1:1 normalized intensities were used as the references. Only MS/MS spectra in which a peptide assignment to albumin could be made were used. Plots display the observed versus expected ratios. Raw box plots of reporter ion data can be viewed in Supplemental Figure S1.

Plots were generated by determination of the median normalized intensity for each reporter ion, as provided in Supplemental Figure S2, followed by calculation of the mean. The mean for the two intensities was then used as the reference to generate fold-change values, which are presented in Figure 5. Interestingly, a slight deviation is observed from linearity above 3-fold; it is important to note that at the higher intensities there was no signal saturation observed; that is all raw ion intensities are $<1 \times 10^4$. To expand upon this study, we performed a spike-in study in which a complex matrix background was used (Figure 6). Specifically, albumin-depleted human CSF was spiked with yeast alcohol dehydrogenase at varying concentrations, 1.25, 2.5, 5, and 10 picomoles per reaction, in duplicate. Following a mascot search, 10 peptides were assigned to AHD1 having ion scores of 22 or greater. Figure 6A is a box and whisker plot of the isotope-corrected, row-normalized reporter ion peak area values for the ADH1 peptides. To evaluate the data as a whole, we combined each reporter ion data and plotted the ratio (relative to the 114/117 average) versus the expected ratio, as shown in Figure 6B. An example of a tandem mass spectrum obtained from this experiment is provided in Figure 6C for the ADH1 peptide, YVVDTSK (1419.79 observed m/z). Taken together, these data indicate that incorporation of this new normalization strategy is feasible for large-scale studies where differences in peptide and protein abundance are expected.

DISCUSSION

This study was aimed at understanding the behavior of the iTRAQ reporter ions recorded in the tandem MS experiment in a complex mixture using the raw data without the need of proprietary or sophisticated data processing. Specifically, we focused on using a MALDI-TOF/TOF instrumentation platform. Previous work using the iTRAQ 4-plex reagents and a MALDI-TOF/TOF instrument reports an increase in variance as a function of low signal intensity of the reporter ions,[18a] and this finding has been also noted in subsequent work on electrospray platforms.[8a,18b] Consistent with this work, we sought to evaluate the behavior of the 8-plex reagents by imposing a row-sum threshold of 5000 area units. This roughly corresponds to a peak area of 625 for each reporter ion, which we determined based on manual evaluation of MS/MS spectra collected in the initial data set. A previous study of iTRAQ reporter ion behavior analyzed via MALDI-TOF/TOF also used individual peak area values for characterization; however, individual peak areas >5000 were required, and normalization was carried out over the entire experiment data set based on median values.[18a] It should also be noted that these authors also report an increase in variance with decreasing ion abundance. Related studies that have used the same instrumental setup to analyze spiked standard proteins have relied on peak-centroided intensity-based values rather than peak area values and analyzed using a log transformation.[8b] The thresholding included in this study is used solely to characterize the iTRAQ behavior. One cannot employ such an approach in a biological system because data could be omitted for peptides representing true protein differences between systems or conditions that are at the upper and lower limits of detection for a particular instrument.

The normal distribution of our data is particularly interesting to note. This feature allows for the incorporation of parametric statistical analyses without the need for logarithmic transformation. Other approaches to establish fold-change cut off values have been recently reported including the incorporation of control−control samples using TMT alongside non-parametric evaluation to establish cutoff values.[22] Other approaches include the development of an ANOVA model for multiple iTRAQ studies using simple mixtures,[13b] followed by extensive characterization of complex mixtures.[13a] A completely different approach to evaluate the accuracy within an experiment is to include a decoy sample;[13d] however, this approach reduces the ability to multiplex by consistently occupying one label for the decoy proteome. Yet another attempt to solve this difficult problem was reported in which a data-driven approach was incorporated to generate probability distribution functions and false discovery rates of large data sets.[23] A comprehensive study in which the iTRAQ reporter ion accuracy and precision were evaluated suggests that using a generalized log variance-stabilizing transformation is appropriate.[8a] It should be noted that in this report the experimental variation was determined to be included in fold-change cut off values of ±1.1. Contrary to the current study, those authors included only unique peptides that were identified as belonging to a protein, and, in addition, a different instrumental platform(s) was used. The data presented here represent a

DOI: 10.1021/acs.jproteome.5b00254
J. Proteome Res. XXXX, XXX, XXX−XXX

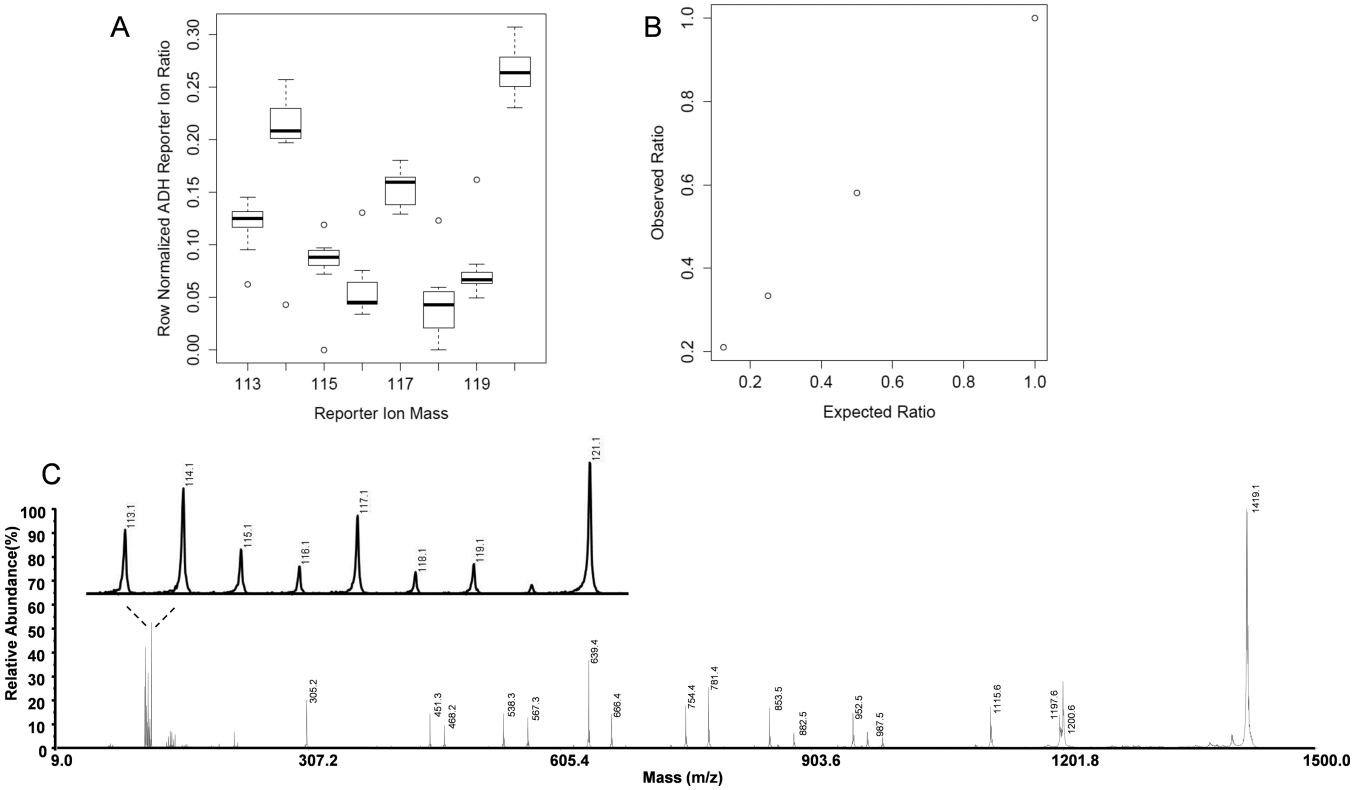


Figure 6. Evaluation of ADH1 peptides spiked into albumin-depleted human cerebrospinal fluid. (A) Box and whisker plot displaying the differing normalized intensity values for ADH1 peptides assigned using Mascot. The internal standard was spiked in duplicate as follows: 10 pmol (114, 121), 5 pmol (113, 117), 2.5 pmol (115, 119), and 1.25 pmol (116, 118). (B) Graph of observed versus expected ratio by taking the mean normalized intensity for each label (using all peptides) and dividing relative to the mean of the 114/121 intensities. A linear response is observed. (C) Example tandem mass spectrum of the peptide YVVDTSK from ADH1. The insert contains a magnified portion of the spectrum that contains the iTRAQ reporter ions.

means to evaluate raw data and variance without the need for a global normalization or logarithmic conversion.[13e] The fold-change values reported here may appear conservative to some; however, in our hands we believe that this strict cutoff will result in large data sets where potential biological markers can truly be validated by other methods.

While not extensively addressed in this work, the idea of isotope impurity or overlap still presents a challenge in robustly making quantitative measurements as previously reported.[8c] At present, the manufacturer does not provide correction factors for the 8-plex iTRAQ reagents. Isotopic purity is noted for each label and in our experience is between 98 and 100%, as reported by the manufacturer. Several commercial software platforms do provide isotopic purity correction as default settings; however, it remains arbitrary at this point to apply corrections in a blanket fashion. Work from Mahoney and coworkers suggests that isotope correction does not correct bias seen between observed and expected fold-change values.[8b] The data presented in Figures 3, 5, and 6 and Supplemental Figure S1 represent data that have been isotope-corrected using default values. That said, our data represent a minimum threshold for evaluating raw data and making cutoff decisions. Finally, the mock biological study using tryptic digests of bovine serum albumin and spiked human CSF suggests that the normalization strategy presented is appropriate for large-scale studies.

Several issues remain, including migration from peptide-level data to inferred protein-level quantitative results, evaluation of peptides that do not behave coordinately within a protein assignment, and the minimum number of peptides needed to make a quantitative determination at the protein level. On the basis of the data collected on this instrumental platform, a recommendation of strict fold-change required is suggested, although similar evaluations are needed on the platform specific to an individual laboratory. Furthermore, optimization on the instrumentation platform is also required. This study demonstrates the need for individual laboratories to perform complete characterization and optimization for iTRAQ studies. We report here our approach to evaluate the behavior of iTRAQ reporter ion utilizing a MALDI-TOF/TOF instrumentation platform.

CONCLUSIONS

This report demonstrates a simple yet thorough protocol to evaluate protein digestion, covalent labeling steps, and analysis of quantitative mass spectrometry data obtained on a MALDI-TOF/TOF instrumental platform. Using a cell lysate sample as a true complex mixture, we were able to confirm digestion and labeling steps in our protocol. Furthermore, using a "row-based" normalization strategy for each precursor ion, we are able to evaluate the distribution of iTRAQ reporter ion area values for an entire data set in which normal distributions were observed without the need for log-based transformation. Additionally, by extracting the individual reporter ion area values for each label, the behavior of an individual reporter ion is observed. This is important when considering problems associated with labeling, as was observed for one labeling reaction out of the total of 32 that were performed for this

DOI: 10.1021/acs.jproteome.5b00254
J. Proteome Res. XXXX, XXX, XXX−XXX

work. Incorporation of a predigested exogenous protein as an internal standard was considered and may be useful to evaluate performance of a specific iTRAQ label. Current efforts are focused on the incorporation of internal standards for this type of evaluation, specifically, understanding the power needed to observe poor iTRAQ reporter ion signals. By evaluating the comparison of all reporter ions to one another we are able to generate a distribution of ratios that represent the normal variance of reporter ion signals. This distribution can then be used to evaluate a fold-change cut-off value for an entire data set regardless of peptide assignments. Alternatively, as was demonstrated in this work, using a subset of high-quality data with peptide assignments can be implemented to define a significant ratio at the peptide level based on experimentally determined variance, and the normalization protocol described is relevant for biological studies. Taken together, this report demonstrates the need to understand the iTRAQ reporter ion behavior, or any other quantitative approach, on the instrumental platform of choice and provides an opportunity for informaticians to utilize as a training data set.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.5b00254.

Supplementary Figure 1. Box and whisker plot analysis of representative replicate data sets (#1, #3) of the HepG2 iTRAQ labeled proteome analysis including isotope correction, row-sum cutoff, and normalization. Supplementary Figure 2. Box and whisker plot analysis of a mock biological sample using bovine serum albumin tryptically digested and labeled with iTRAQ tags. Data was isotope-corrected and normalized. (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: aly@helix.nih.gov.

Present Address

⊥C.A.C.: The Johns Hopkins University School of Medicine, Baltimore, MD.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has been supported by the intramural program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services. Partial support was provided by the National Niemann-Pick Disease Foundation (Peter Pentchev Fellowship to S.M.C.). Funding for this work was also provided in part by the RSH/Smith Lemli Opitz Research Foundation (to S.M.C.)

■ REFERENCES

(1) (a) Bantscheff, M.; Schirle, M.; Sweetman, G.; Rick, J.; Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* **2007**, *389* (4), 1017−1031. (b) Bantscheff, M.; Lemeer, S.; Savitski, M. M.; Kuster, B. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal. Bioanal. Chem.* **2012**, *404* (4), 939−965. (c) Treumann, A.; Thiede, B. Isobaric protein and peptide quantification: perspectives and issues. *Expert Rev. Proteomics* **2010**, *7* (5), 647−653.

(2) Ross, P. L.; Huang, Y. L. 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.

(3) Williamson, B. L.; Ross, P. L.; Pillai, S.; Purkayastha, B.; Daniels, S.; Pappin, D. Protein quantitation using a novel 8-plex set of isobaric peptide labels. *Mol. Cell Proteomics* **2006**, *5* (10), S55−S55.

(4) Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Hamon, C. Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **2003**, *75* (8), 1895−1904.

(5) Werner, T.; Becher, I.; Sweetman, G.; Doce, C.; Savitski, M. M.; Bantscheff, M. High-Resolution Enabled TMT 8-plexing. *Anal. Chem.* **2012**, *84* (16), 7188−7194.

(6) Everley, R. A.; Kunz, R. C.; McAllister, F. E.; Gygi, S. P. Increasing Throughput in Targeted Proteomics Assays: 54-Plex Quantitation in a Single Mass Spectrometry Run. *Anal. Chem.* **2013**, *85* (11), 5340−5346.

(7) Evans, C.; Noirel, J.; Ow, S. Y.; Salim, M.; Pereira-Medrano, A. G.; Couto, N.; Pandhal, J.; Smith, D.; Pham, T. K.; Karunakaran, E.; Zou, X.; Biggs, C. A.; Wright, P. C. An insight into iTRAQ: where do we stand now? *Anal. Bioanal. Chem.* **2012**, *404* (4), 1011−1027.

(8) (a) Karp, N. A.; Huber, W.; Sadowski, P. G.; Charles, P. D.; Hester, S. V.; Lilley, K. S. Addressing Accuracy and Precision Issues in iTRAQ Quantitation. *Mol. Cell. Proteomics* **2010**, *9* (9), 1885−1897. (b) Mahoney, D. W.; Therneau, T. M.; Heppelmann, C. J.; Higgins, L.; Benson, L. M.; Zenka, R. M.; Jagtap, P.; Nelsestuen, G. L.; Bergen, H. R.; Oberg, A. L. Relative Quantification: Characterization of Bias, Variability and Fold Changes in Mass Spectrometry Data from iTRAQ-Labeled Peptides. *J. Proteome Res.* **2011**, *10* (9), 4325−4333. (c) Ow, S. Y.; Salim, M.; Noirel, J.; Evans, C.; Rehman, I.; Wright, P. C. iTRAQ Underestimation in Simple and Complex Mixtures: "The Good, the Bad and the Ugly". *J. Proteome Res.* **2009**, *8* (11), 5347−5355.

(9) Burkhart, J. M.; Vaudel, M.; Zahedi, R. P.; Martens, L.; Sickmann, A. iTRAQ protein quantification: A quality-controlled workflow. *Proteomics* **2011**, *11* (6), 1125−1134.

(10) Ow, S. Y.; Salim, M.; Noirel, J.; Evans, C.; Wright, P. C. Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation. *Proteomics* **2011**, *11* (11), 2341−2346.

(11) Savitski, M. M.; Mathieson, T.; Zinn, N.; Sweetman, G.; Doce, C.; Becher, I.; Pachl, F.; Kuster, B.; Bantscheff, M. Measuring and Managing Ratio Compression for Accurate iTRAQ/TMT Quantification. *J. Proteome Res.* **2013**, *12* (8), 3586−3598.

(12) Ow, S. Y.; Noirel, J.; Salim, M.; Evans, C.; Watson, R.; Wright, P. C. Balancing robust quantification and identification for iTRAQ: Application of UHR-ToF MS. *Proteomics* **2010**, *10* (11), 2205−2213.

(13) (a) Oberg, A. L.; Mahoney, D. W.; Eckel-Passow, J. E.; Malone, C. J.; Wolfinger, R. D.; Hill, E. G.; Cooper, L. T.; Onuma, O. K.; Spiro, C.; Therneau, T. A.; Bergen, H. R., III Statistical analysis of relative labeled mass spectrometry data from complex samples using ANOVA. *J. Proteome Res.* **2008**, *7* (1), 225−233. (b) Hill, E. G.; Schwacke, J. H.; Comte-Walters, S.; Slate, E. H.; Oberg, A. L.; Eckel-Passow, J. E.; Therneau, T. M.; Schey, K. L. A statistical model for iTRAQ data analysis. *J. Proteome Res.* **2008**, *7* (8), 3091−3101. (c) Oberg, A. L.; Mahoney, D. W. Statistical methods for quantitative mass spectrometry proteomic experiments with labeling. *BMC Bioinformatics* **2012**, *13*, S7 DOI: 10.1186/1471-2105-13-S16-S7. (d) Vaudel, M.; Burkhart, J. M.; Radau, S.; Zahedi, R. P.; Martens, L.; Sickmann, A. Integral Quantification Accuracy Estimation for Reporter Ion-based Quantitative Proteomics (iQuARI). *J. Proteome Res.* **2012**, *11* (10), 5072−5080. (e) Boehm, A. M.; Putz, S.; Altenhofer, D.; Sickmann, A.; Falk, M. Precise protein quantification based on peptide quantification using iTRAQ (TM). *BMC Bioinf.* **2007**, *8*, 214. (f) Kim, P. D.; Patel, B. B.;

DOI: 10.1021/acs.jproteome.5b00254
J. Proteome Res. XXXX, XXX, XXX−XXX

Yeung, A. T. Isobaric labeling and data normalization without requiring protein quantitation. *Journal of biomolecular techniques: JBT* **2012**, *23* (1), 11−23.
(14) Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20* (18), 3551−67.
(15) Searle, B. C. Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics* **2010**, *10* (6), 1265−9.
(16) Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **2002**, *74* (20), 5383−92.
(17) Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **2003**, *75* (17), 4646−58.
(18) (a) Hu, J.; Qian, J.; Borisov, O.; Pan, S.; Li, Y.; Liu, T.; Deng, L.; Wannemacher, K.; Kurnellas, M.; Patterson, C.; Elkabes, S.; Li, H. Optimized proteomic analysis of a mouse model of cerebellar dysfunction using amine-specific isobaric tags. *Proteomics* **2006**, *6* (15), 4321−34. (b) Gan, C. S.; Chong, P. K.; Pham, T. K.; Wright, P. C. Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *J. Proteome Res.* **2007**, *6* (2), 821−827. (c) Song, X. M.; Bandow, J.; Sherman, J.; Baker, J. D.; Brown, P. W.; McDowell, M. T.; Molloy, M. P. iTRAQ experimental design for plasma biomarker discovery. *J. Proteome Res.* **2008**, *7* (7), 2952−2958.
(19) Searle, B. C.; Turner, M. In *Advancing Statistical Analysis of Multiplexed MS/MS Quantitative Data with Scaffold Q+*, 60th Annual ASMS Conference on Mass Spectrometry and Allied Topics, Vancouver, B.C., June, 2012.
(20) Shadforth, I. P.; Dunkley, T. P.; Lilley, K. S.; Bessant, C. i-Tracker: for quantitative proteomics using iTRAQ. *BMC Genomics* **2005**, *6*, 145.
(21) Cleries, R.; Galvez, J.; Espino, M.; Ribes, J.; Nunes, V.; de Heredia, M. L. BootstRatio: A web-based statistical analysis of fold-change in qPCR and RT-qPCR data using resampling methods. *Comput. Biol. Med.* **2012**, *42* (4), 438−45.
(22) Serang, O.; Cansizoglu, A. E.; Kall, L.; Steen, H.; Steen, J. A. Nonparametric Bayesian evaluation of differential protein quantification. *J. Proteome Res.* **2013**, *12* (10), 4556−65.
(23) Wilson, J., Pappin, D. J. In *FDR of "Different" Empirical, Probability Based Determination of Up- and Down-Regulation in Quantitative Proteomics*, 61st Conference of the American Society of Mass Spectrometry, 2013.

DOI: 10.1021/acs.jproteome.5b00254
J. Proteome Res. XXXX, XXX, XXX−XXX