Stable isotope dimethyl labeling coupled to selected reaction monitoring enhances throughput by multiplexing relative quantitation of targeted proteins

Thin Thin Aye[1*], Teck Yew Low[2,3], Yngvild Bjørlykke[1],
Harald Barsnes[1], Albert J.R. Heck[2,3], Frode S. Berven[1,4*]

[1] Proteomics Unit (PROBE), Department of Biomedicine, University of Bergen, Bergen, Norway.

[2] Biomolecular Mass Spectrometry and Proteomics Group, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584CH, Utrecht, The Netherlands.

[3] Netherlands Proteomics Centre, Padualaan 8, 3584 CH, Utrecht, The Netherlands.

[4] The Norwegian Multiple Sclerosis Competence Centre, Haukeland University Hospital, Bergen, Norway.

* Author to whom correspondence should be addressed.
Address reprint request to
Drs. Frode S. Berven and Thin Thin Aye
PROBE, Department of Biomedicine
University of Bergen
Jonas Liesvei 91
N-5009 Bergen, Norway
Phone: (+47) 55 58 67 43
Fax: (+47) 55 58 63 60
Email: Frode.Berven@biomed.uib.no
Email: Thin.aye@biomed.uib.no

Abstract

In this study we present a proof-of-concept for targeted relative protein quantitation workflow using chemical labeling in the form of dimethylation, coupled with selected reaction monitoring (Dimethyl-SRM). We first demonstrate close to complete isotope incorporation for all peptides tested. The accuracy, reproducibility, and linear dynamic range of quantitation are further assessed based on known ratios of non-human standard proteins spiked into human cerebrospinal fluid (CSF) as a model complex matrix. Quantitation reproducibility below 20% (CV<20%) was obtained for analyte concentrations present at a dynamic range of 4 orders of magnitude lower than that of the background proteins. An error of less than 15% was observed when measuring the abundance of 45 major human plasma proteins. Dimethyl-SRM was further examined by comparing the relative quantitation of eight proteins in human CSF with the relative quantitation obtained using synthetic heavy peptides coupled to stable isotope dilution–SRM (SID-SRM). Comparison between the two methods reveals that the correlation between dimethyl-SRM and SID-SRM is within 0.3-39% variation, demonstrating the accuracy of relative quantitation using dimethyl-SRM. Dimethyl labeling coupled with SRM provides a fast, convenient and cost-effective alternative for relative quantitation of a large number of candidate proteins/peptides.

INTRODUCTION

Quantitative proteomics is increasingly used for the discovery of diagnostic and prognostic biomarkers, therapeutic targets and for unraveling novel biological mechanisms.[1] These efforts are augmented by stable isotope labeling which improves the reproducibility of mass spectrometry (MS)-based quantitation[2-7]. Currently, stable isotope-based measurements are predominantly performed in a "discovery" or "non-targeted" manner, where run-to-run reproducibility and comprehensiveness of quantitation are hampered by sample complexity and the sampling speed of MS.[8, 9] Moreover, to quantify a large portion of the proteome, MS analyses are usually performed after exhaustive pre-fractionation[10], which limits the sample size (i.e. n <10) being analyzed.

Reliable and routine MS-based proteomics necessitate consistency, speed, robustness and sensitivity. To achieve these, data can be acquired in a "targeted" manner by employing selected reaction monitoring (SRM)-MS.[11] SRM is a tandem MS mode unique to triple-quadrupole mass spectrometers, being capable of rapid, sensitive, and specific quantitation of selected analytes in highly complex samples.[12] With SRM, *a priori* selected proteins can be quantified with less pre-fractionation, thereby generating desirable and consistent data[13, 14] in reduced analysis times. LC-SRM enables assays for any protein targets, so as to serve as alternative to antibody-based detection and quantitation, where antibody availability and specificity are often limiting. Besides, LC-SRM is also very sensitive as it was demonstrated that proteins can be quantified at extremely low copy numbers in cells[13] and in plasma at levels approaching ng/ml.[15, 16]

In quantitative SRM assays, proteotypic peptides are first selected from discovery experiments, or from databases such as PeptideAtlas[17]. Then, corresponding stable isotopic synthetic (SIS) peptides are spiked in as internal standards to estimate the concentrations of selected proteins by stable isotope dilution–SRM (SID-SRM). Despite having high accuracy and precision, SID-SRM is in practice limited by the high cumulative costs of SIS peptides. As an alternative, SRM can be combined with peptides chemically-derivatized with stable isotopes. Although quantitative methods based on chemical labeling are mostly adopted in non-targeted MS[2, 3, 5-7], they are increasingly applied in targeted MS. Desouza *et. al.*[18] combined SRM with mTRAQ reagent to achieve absolute quantification of peptides and proteins via isotope-dilution mass spectrometry and later it was successfully applied to SRM-based biomarker validation.[19] In addition, by introducing ^{18}O-labeled proteome reference as a Global Internal Standard (GIS), Kim *et al.* [20] showed that relative quantitation of large numbers of targeted proteins in plasma can be achieved with SRM.

In this study, we chose to couple dimethyl labeling with SRM, as it is inexpensive and straightforward.[2, 21] Besides, Munoz *et al.*[22] had demonstrated that over 7,000 proteins can be confidently quantified by triplex dimethyl labeling in a non-targeted study of human embryonic stem (hES) cells. Here, we developed, optimized and characterized a targeted high-throughput approach combining dimethyl labeling with SRM (dimethyl-SRM). We first examined the reproducibility, accuracy, dynamic range, assay linearity, labeling efficiency and isotopic effects of this newly developed method, followed by its application to human plasma and cerebrospinal fluid (CSF); both being widely used for biomarker studies and are generally considered among the most challenging samples due to the wide dynamic range of their proteomes.[23-25] In addition, we benchmarked the performance of dimethyl-SRM against SID-SRM using CSF.

MATERIALS AND METHODS

Materials: Dithioreithol, ammonium bicarbonate, and sodium cyanoborohydride were purchased from Fluka (Buchs, Switzerland). Iodoacetamide, triethyl ammonium bicarbonate, and formaldehyde (37%) were purchased from Sigma (St. Louis, MO). CD_2O (98% D, 20 wt%), $^{13}CD_2O$ (99% ^{13}C, 98% D, 20 wt%), and sodium cyanoborodeuteride (96% D) were from Isotec (Miamisburg, OH). Five non-human standard proteins (β-galatosidase, thioredoxin, hexokinase, aldehyde dehydrogenase and α-amylase) were purchased from Sigma-Aldrich (Germany). Unless stated otherwise, all additional chemicals were purchased from commercial sources and were of analytical grade.

Protein digestion and Stable isotope dimethyl labeling: All protein samples were denatured and digested as described in supplemental information (SI) and the tryptic peptides were desalted using the *Oasis*® HLB µElution 96 well plate (Oasis, Millipore), vacuum-dried prior to labeling with dimethyl reagents. The labeling procedure is described in SI. The three differentially labeled samples were pooled and again desalted for SRM analysis without fractionation.

Plasma and CSF sample preparation: Plasma sample was obtained from the Department of Neurology, Haukeland University Hospital. Immediately after collection in EDTA, the blood sample was centrifuged at 1000g for 15min at 22˚C to pellet cells. The supernatant were immediately frozen at -80˚C. CSF samples were collected from six patients undergoing lumbar puncture as part of routine diagnostic evaluations at the Haukeland University Hospital. CSF samples were centrifuged at 450*g* for 5 minutes to remove cells and cell debris, and the

supernatant were frozen at -80 °C. The collection of human samples was approved by the regional ethics committee, and written informed consent was obtained from all patients.

NanoLC-SRM analysis: An Ultimate 3000 system (Dionex) LC packing NanoLC (Ultimate 3000 system, Dionex) coupled with a QTRAP4000 (AB Sciex) with a nano-electrospray ionization source controlled by Analyst 1.5 software (AB Sciex) were used for all LC-SRM analyses. The sample (1µg) was loaded onto a pre-column (Dionex, Acclaim PepMap Nano Trap column, C18, 75 µm i.d. x 2 cm, 3 µm) followed by separation on the analytical column (Dionex, Acclaim PepMap100 RS LCnano column, 75 µm x 15 cm, C18, 2 µm) with flow rate of 250 nl/min. The LC gradient was 70 min with the following mobile phases: Mobile phase A (0.1% FA) and mobile phase B (0.1% FA / 90% ACN). The gradient used was as follows: 5-10% B from 0-3 minutes, 10-45% B from 3-45 minutes and 45-90% B from 45-45.5 minutes, and hold at 90% B from 45.5-51.5 minutes. From 51.5-55 minutes ramp from 90-5% B and re-equilibrate column at 5% B from 55-70 minutes. All acquisition methods used the following parameters: 2200-V ion spray voltage, a curtain gas setting of 25 psi, a 200 °C interface heater temperature, and Q1 and Q3 set to unit resolution (0.6–0.8 Da full width at half-height). For the study on 45 selected plasma proteins, a Q-Trap 5500 (AB Sciex) coupled to a Dionex Ultimate NCS-3500RS LC system (Sunnyvale, CA) was used for LC SRM-MS analysis. The MS run time was 68 min. Scheduled SRM option was used for data acquisition with a target scan time of 1s and a 4 min SRM detection window.

Peptide selection and SRM transition selection in MRMPilot: In all experiments, pre-selection criteria of proteotyptic peptides were based on Lange *et al.*[11] We also developed an in-house script for MRMPilot, to generate predicted transitions for dimethyl peptides. The detailed settings include: peptide lengths between 7-17 amino acids, trypsin chosen as the enzyme, no-miscleavages, carbamidomethyl (C) as fixed modification, light-, intermediate- and heavy-dimethyl (both K and N-term) as variable modifications. As each peptide was monitored in light-, intermediate- and heavy-labeled channels, this resulted in nine transitions per peptide in most of the experiments presented in this manuscript. Detailed information can be accessed in SI.

SIS peptide optimization: SIS peptides were custom-synthesized by Thermo Fisher Scientific GmBH (Germany) in crude quality. Each peptide was dissolved in 0.1% formic acid at a nominal concentration of 25 pmol as stock solution (Note: the concentration was estimated based on the crude concentration provided). Seven out of eight selected peptides were used for optimization in order to obtain the best combination of Q/Q3 ion pairs. ELPEHTVK (vitamin

D-binding protein), was excluded due to poor solubility in the designated buffer. Optimization procedure is described in SI.

SRM data analysis: The SRM data were processed using MultiQuant 2.2 (AB Sciex). The detailed setting of MultiQuant can be accessed in SI. Besides the ratios reported by MultiQuant, the integrated peak areas were validated manually and evaluated based on correct peak detection, accurate integration, matrix interference and retention time and S/N level. All subsequent data analyses were performed in MS-Excel to generate standard deviations and coefficients of variation. The most abundant transitions for each heavy/intermediate/light peptide pair were used to quantify the peptide ratios. The protein ratio was calculated by averaging the peptide ratios.

RESULTS AND DISCISSIONS

Analytical strategy

We aim to establish a platform that couples dimethyl labeling to SRM to enable rapid, cost-effective and reproducible relative quantitation of targeted peptides. The workflow consists of (i) generation of a global internal standard (GIS); (ii) tryptic digestion of the GIS and individual samples to be compared; (iii) isotopic dimethyl labeling at the peptide level; (iv) proteotypic peptide selection from the proteins of interest; (v) predicting the optimal SRM transitions and CE values; (vi) making the transition list with predicted instrument parameters; and finally (vii) LC-SRM and data analysis (Figure 1). In all experiments, three forms of dimethyl labels, i.e., light, intermediate and heavy are introduced into individual peptide mixtures to be quantified after trypsinization. Each label carries a mass difference of 4 Da from one another. Recently, Cappadona *et al.* [26] reported that as the mass and charges of dimethylated peptides increase, isotopic envelopes from differentially labeled peptides start to overlap, thus affecting quantitative accuracy. Hence, we implemented additional selection criteria such as peptide length and charge to minimize the overlap between the dimethylated peptides.

Characterization of dimethyl-SRM/MS

Labeling efficiency: Incomplete labeling of peptides leads to inaccuracy in the quantitation. To determine the degree of incomplete labeling, we labeled separately, equal amounts of BSA tryptic digest with light-, intermediate- and heavy-dimethyl reagents and mixed in a ratio of 1:1:1. Nine peptides were selected with two transitions per peptide for SRM assays. Transitions were monitored for peptides harboring three different dimethyl groups as well as for the unlabeled form to reveal potential incomplete labeling. For five processed replicates (Table S-

1), LC-SRM results consistently showed that non-labeled peptides were undetectable, except for two potentially unlabeled peptides (AEFVEVTK and YLYEIAR) which had S/N values below 10 while the S/N levels of their respective labeled versions were approximately 1000-folds higher. Since three labeling reactions contribute to the monitored unlabeled portion, our results are comparable to other reports of dimethyl labeling efficiency of 97% or better.[21, 27-29] This suggests that incomplete dimethyl labeling of peptides is not a critical issue impeding SRM assays.

Quantitation accuracy: Equal amounts of BSA tryptic digest were separately labeled with light-, intermediate- and heavy-dimethyl reagents and mixed in ratios of 1:1:1, 1:1:2 and 4:2:1. Then, same nine peptides as in the experiment described above were selected with three transitions per peptide for SRM. The triplex labeled peaks for the DDSPDLPK peptide is shown in Figure S-1. All nine targeted peptides are identified with S/N >3, however only eight of them are successfully quantified with S/N >10. The % error was calculated by deriving the % deviation of observed ratios from the expected ratio. An error of <20% indicates the good accuracy of the dimethyl-SRM method, and detailed results are described in Table S-2.

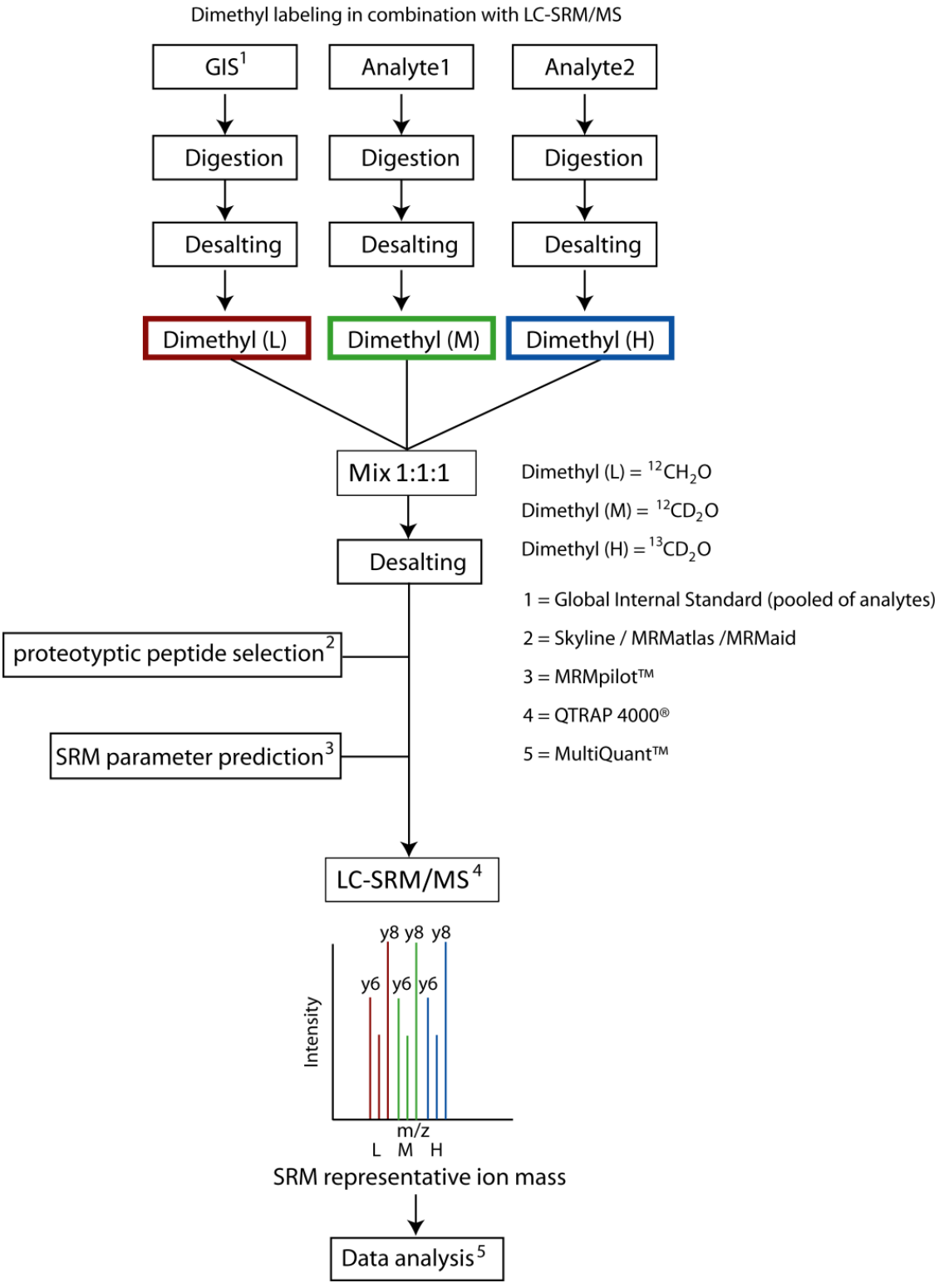

Figure 1. Dimethyl-SRM strategy. Workflow of the high throughput relative quantitation using Dimethyl-SRM.

Inter-assays reproducibility: To evaluate the accuracy and precision of quantitation for dimethyl-SRM in a biologically complex proteome, 30µg of trypsinized CSF sample was split into three equal fractions,

followed by labeling with light-, intermediate- and heavy-dimethyl reagents and then mixed in a 1:1:1 ratio. Ten highly abundant proteins in CSF were selected for LC-SRM. Three peptides per protein and two transitions per peptide were monitored in 68 min LC-SRM runs (Figure S-2). For peptides with no interference, the most abundant transitions (i.e., highest S/N) for each labeled species were used for quantitative analysis. As shown in Table S-3, 27 out of 30 targeted peptides were successfully quantified with S/N >10. The reproducibility (% CV for four process replicates) varied from 4-19% (intermediate version, M/H) and from 2-16% (light version, L/H), respectively. Theoretically, the ratio for heavy, intermediate and light is 1:1:1 for each peptide. Hence, the % error of the observed ratio from the expected ratio was calculated, showing an error <20%. It indicates that dimethyl-SRM is sufficiently robust for CSF, a complex matrix with high complexity and dynamic range.

Dynamic range and limit of quantitation: Both dynamic range and limit of quantitation were assessed by spiking different but known amounts of five non-human proteins into a complex human background matrix. Prior to dimethyl labeling, a mixture of five non-human proteins were spiked into human CSF at seven different concentrations (0.1fmol-100fmol) per µg CSF (corresponding to 1pg-10ng of protein amounts in 1µg CSF) generating seven different hybrid mixtures which were digested individually. Each hybrid mixture was labeled with light-, intermediate-, and heavy-dimethyl groups and mixed as shown in Table S-4A. As a process replica, an inverted experiment was performed (i.e., label swapping). For the subsequent 68 min SRM analysis, we selected two peptides per protein which were observed with no background interference. In addition, one peptide (DLGEENFK) from human serum albumin was also monitored as an internal standard in the quantitation. All five standard proteins were detected at a spike-in level of 10pg based on at least one peptide per protein, suggesting that the limit of detection is at low-picogram level in un-depleted CSF. Six out of ten peptides exhibit a linear response (R^2 >0.92) in relative concentration of 1-100fmol (Figure S-3). We observe non-linearity at a mixing ratio of 1:5:500 due to the matrix interference at low concentration (0.1fmol). Human serum albumin was detected and quantified consistently across all the experiments. In 1µg CSF, there is presumably at least 700ng of serum albumin (HSA). Our results demonstrate that four proteins were quantified in the presence of 10000-fold excess of HSA, suggesting that the dynamic range of our method is at least four orders of magnitude when compared to HSA, the most abundant protein in CSF. The area ratios from all three replicates are consistent (CV = 0.4-20.9%), showing that the quantitative accuracy was minimally affected by the label swapping (Table S-4B).

Chromatographic elution and isotopic effect: In the isotopic dimethyl labels, hydrogen atoms are replaced with deuterium (D) atoms which are slightly more hydrophilic than hydrogen. For dimethyl labeled peptides in reversed phase chromatography, the retention times (RT) for labeled peptides (4D per intermediate-label and 6D per heavy-label) can be affected. Hence, we investigated isotopic effects for 72 triplex dimethylated peptides originating from 16 proteins in all previous datasets mentioned

above for evaluation of accuracy, linearity and sensitivity. For each labeled peptide, the RT of the most intense transition was extracted from the MultiQuant output file. Then, the RT differences (ΔRT) between the three dimethylated species were calculated by subtracting the early eluting peptide form from the late eluting peptide form (i.e., $RT_{light\ dimethylated}$ - $RT_{heavy\text{-}dimethylated}$) (Table S-5). ΔRTs (in seconds) for all peptides are illustrated in Figure 2. For 67 out of 72 dimethylated peptides (93%), all three isotopic forms elutes within 6s, showing minimal changes in RT. One possibility is that we selected peptides without any tryptic miscleavages for SRM analysis, hence the maximum number of dimethyl groups per target peptide is two, limiting the isotopic effect. Five peptides showed retention time difference (ΔRT) greater than 6s. Following manual inspection, we found that these peptides originated from BSA, HSA and plasminogen which are highly abundant. For these overloaded peptides, their XIC peak shapes tend to broaden and thus their corresponding RT cannot be accurately determined.

Relative quantification of human plasma protein markers

In order to evaluate the accuracy and reproducibility of dimethyl-SRM in high throughput, we targeted peptides from 45 plasma proteins derived from the SID-SRM study[16] and the ^{18}O-SRM study[20]. Based on our peptide selection criteria, 57 peptides were selected, where 43 candidates were from Kuzyk *et al.*[16] excluding peptides containing miscleavages or methionine, and 14 peptides were from Kim *et al.*[20] We first trypsinized a plasma sample, and equally divided the digest into 3 aliquots which were individually labeled with light-, intermediate- and heavy-dimethyl reagents. Labeled peptides were mixed in five different ratios (L:M:H) of 1:3:10, 3:1:10, 1:1:1, 3:10:1 and 10:3:1. All 513 transitions were then scheduled and analyzed in 3 technical replicates. Among the 57 peptides, 55 peptides were detected and quantified in all five mixing ratios. One peptide, AEIEYLEK (L-selectin), was found in four mixing ratios except in 10:3:1. This peptide could also not be quantified by Kim *et al.* [20], while Kuzyk *et al.* [16] reported that L-selectin is far less concentrated than other serum proteins. This implies that transitions derived from unmodified peptides or databases such as PeptideAtlas can be directly adapted and transplanted for dimethyl-SRM studies. Moreover, modification by dimethyl reagents does not complicate transition prediction, or preclude transplanted target peptides from detection and quantitation by LC-SRM, avoiding the need for creating new transition databases for dimethylated peptides. Reliable detection and quantitation of L-selectin indicates that the method covers a high dynamic range of the plasma proteome. THLPEVFLSK (Vitamin D-binding protein), was not observed in any of the five mixing ratios. For 12 proteins that were analyzed with two target peptides, we reported

only the peptide with the highest signal intensity. As a result, we consolidated the data to include only 44 peptides so that only one peptide per protein is reported and tabulated in Table S-6A and B. Good assay linearity (R^2 >0.98) was observed for the 44 quantifiable peptides, including L-selectin. Regarding accuracy the error was less than 15%, and the CV was less than 10% for the reproducibility measurements.

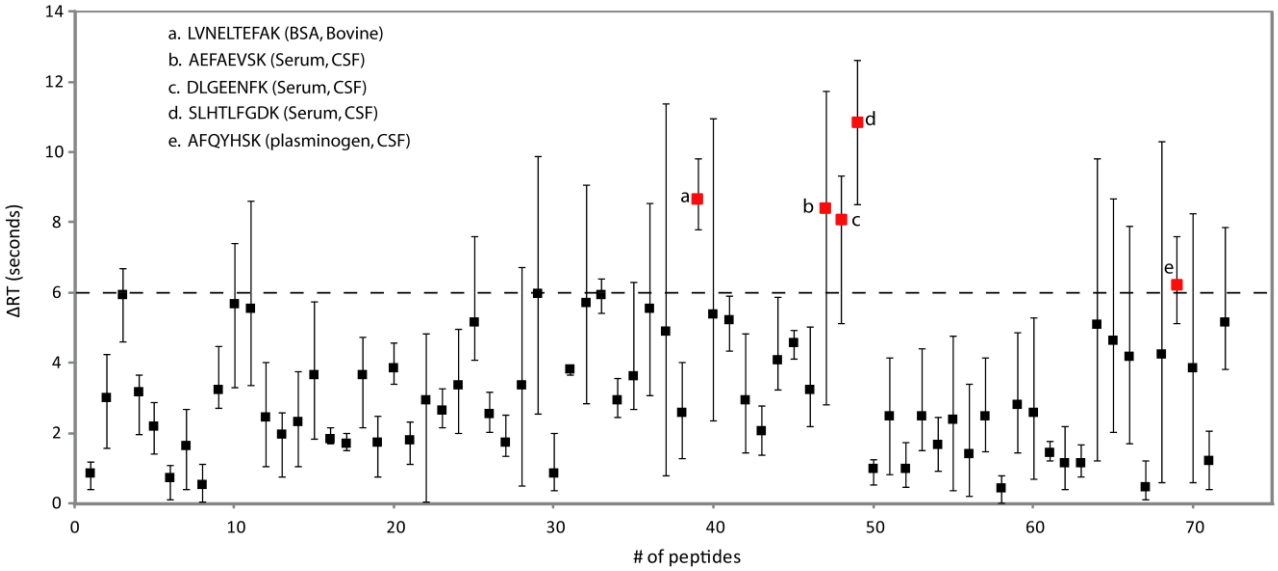

Figure 2. Evaluation of isotopic effects. Retention time differences (ΔRT) of 72 triplex dimethylated peptides were investigated. Five peptides with ΔRT >6s are colored in red. The standard deviations of each peptide for four process replicates are depicted as error bar.

Relative quantitation of CSF proteins using SID-SRM and dimethyl-SRM

Since stable isotope dilution–SRM (SID-SRM) is a standardized and widely-used technique, we decided to benchmark dimethyl-SRM further with SID-SRM performed in-house by comparing the relative quantification of eight proteins across six CSF samples from six individual patients. These proteins (α-2 macroglobulin, vitamin-D binding protein, α-1 acid glycoprotein 1, prostaglandin-H2 D-isomerase, secretogranin-1 and, cystatin-C, contactin-1, and chromogranin-A) where chosen as they previously have been suggested as biomarker candidates for neurological disease such as Multiple Sclerosis (MS), Alzheimer disease (AD) and Parkinson's disease (PD) [24]. The six patients (NN1, NN10, NN20, NN29, NN41 and NN48) did not have any neurological symptoms and thus their CSFs can be regarded as neurologically normal. The experimental setup is described in Figure 3. In short, 40µg proteins from each of the six patients were digested with lysC/trypsin and then split into two equal fractions of 20µg for each patient. One aliquot was then used for SID-SRM and the other for dimethyl-SRM.

SID-SRM strategy for CSF sample: We used eight SIS peptides, each proteotypically representing the eight target proteins in CSF. Prior to the SID-SRM experiment, we generated linear response curves in order to assess LOQ (S/N >10) of all the targeted endogenous peptides in CSF samples (Figure S-4). The peptides GLSAEPGWQAK (chromogranin-A), IVESYQIR (contactin-1) and ALDFAVGEYNK (Cystatin-C) were excluded due to low S/N level of endogenous peptides; while ELPEHTVK (vitamin D-binding protein) was excluded due to poor solubility of synthetic peptides in designated buffer. This left us with four SIS peptides (Table S-7) that were added to each of the six samples in identical amounts, and close to the endogenous peptide level in CSF. The MS run time for each analysis was 68 min with three technical replicates. The relative quantitative ratios were derived from the peak areas representing endogenous peptides and normalized against the amount of synthetic peptide spiked in CSF. The final peptide ratios were calculated by averaging the three technical replicas (run-to-run variation of 0.42-9.98%). Overall, data from the four selected peptides representing four proteins indicated a biological variation of 23-53% across six patients (Table S-7).

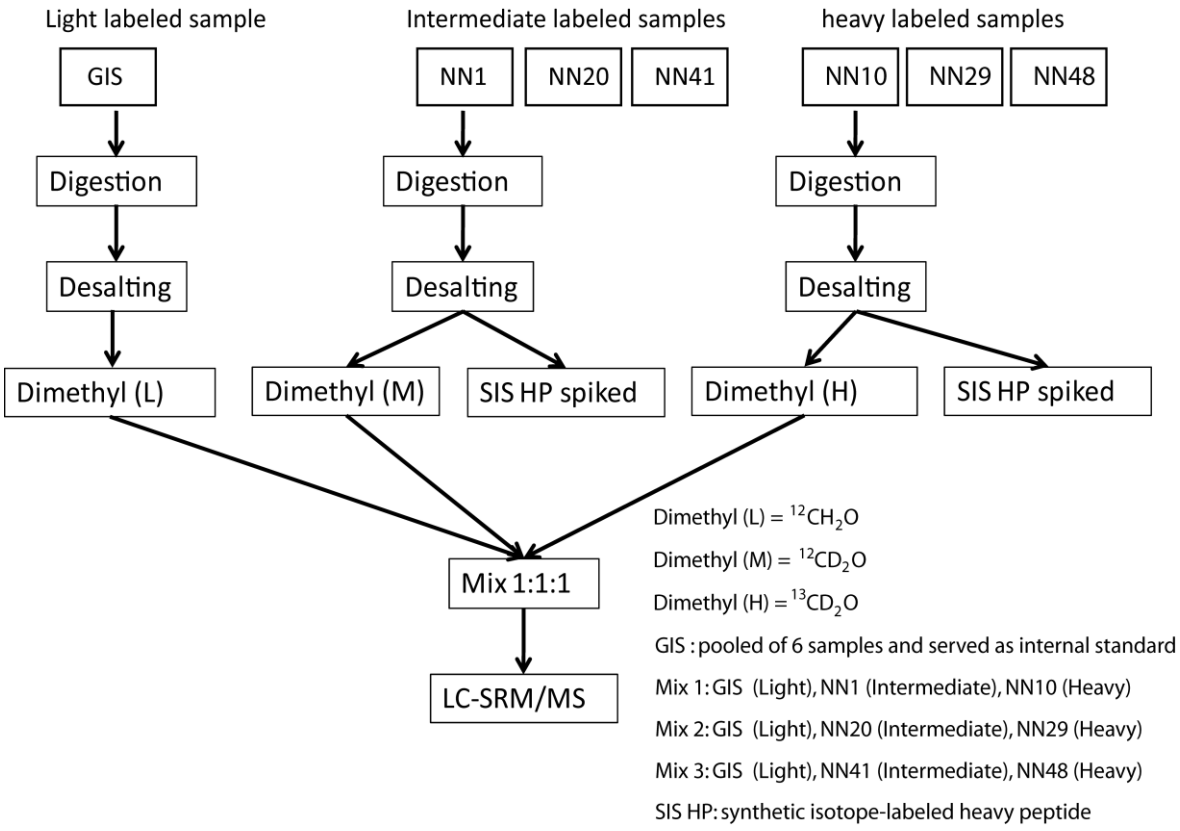

Figure 3. Workflow for CSF biomarker study. 40µg of proteins from each of the six patients were used. The proteins were digested with lysC/trypsin and then split into two equal fractions after desalting. One aliquot was used for SID-SRM and the other for dimethyl-SRM. Eight proteins were selected for analysis. In SID-SRM, the estimated SIS HP peptide amount added to each sample was close to the endogenous peptide level in the sample. In dimethyl-SRM, 10µg CSF of each patient was pooled to generate a global internal standard (GIS).

Dimethyl-SRM strategy for CSF sample: We pooled CSFs from all six patients to generate a Global Internal Standard (GIS) and labeled with light-dimethyl. Three samples (NN1, NN20 and NN41) were independently labeled with intermediate-dimethyl, whereas each of the other three samples (NN10, NN29 and NN48) was labeled with the heavy-dimethyl. Equal amounts of the GIS, one intermediate- and one heavy-dimethylated sample were mixed, generating three set of mixtures for LC-SRM (Table S-8). We targeted the same eight proteins from the SID-SRM experiment. We included two proteotypic peptides for each of the eight selected proteins, resulting in 16 targeted peptides to obtain confirmatory data for estimating protein abundance without incurring extra costs or efforts. The peptide ratios between the GIS and the individual samples were averaged across three technical replicates (CV= 1-17%). Overall, most targeted proteins show similar protein expression profile across six different samples with CVs (%) of 23-47% (Figure S-5). We then selected only quantitative data of one peptide per protein (the same peptides as for the SID-SRM approach) for evaluating both SRM methods at the peptide level.

SID-SRM vs. Dimethyl-SRM: The performance of SID-SRM and dimethyl-SRM were evaluated using the same targeted proteins and peptides. As shown in Figure 4, the error between SID-SRM and dimethyl-SRM extracted ratios was in general below 20%, except for four of the 24 measurements. Maximum error was 39% for AIGYLNTGYQR (α-2-macroglobulin) in patient NN41. We managed to detect and quantify eight out of eight selected peptides from CSF samples using dimethyl-SRM (100% success rate), while only four peptides were successfully detected and quantified using SID-SRM (Table S-9). From our experience, based on the peptides that we have tested, the success rate for enabling relative quantification using the dimethyl-SRM method is consistently high. In our hands, dimethyl-SRM gave a technical variation of 1-17.29 % and biological variation of 23-47% while for SID-SRM, the corresponding values are 0.42-9.9% and 22.5-52.8%, respectively. Hence, SID-SRM achieved better reproducibility with CV <10% compared to dimethyl-SRM (CV <20%) indicating that SID-SRM is the most precise and accurate quantitative method of the two. Nevertheless, dimethyl-SRM exhibits several advantages over SID-SRM. First, due to the availability of three labels, three samples, including the GIS, can be multiplexed in a single LC-SRM analysis, doubling the sample throughput. Secondly, with GIS, labeled counterparts of virtually every peptide are present in the samples to be measured, including post-translationally modified peptides such as phosphorylation[30, 31], and peptides that are not stable or easily soluble as synthetic pure peptides in solution, or difficult to synthesize for different reasons. Also the optimization of the spike-in amount of the labeled peptides is not necessary when a pool of samples with representative levels of all proteins is used as the GIS. The increased sample complexity that occur when doing the dimtehyl-SRM approach can be remedied by the SRM-mode which allows complex samples to be analyzed with high selectivity due to two levels of mass selection. Finally, dimethyl-SRM could be attractive to use to triage (qualify or discard) a large number of biomarker candidates prior to committing to the subsequent more time-, cost- and resource- demanding steps of the more accurate SID-SRM approach.

Dimethyl labeling introduces a mass increment of 28 Da (light), 32 Da (intermediate) and 36 Da (heavy) to each susceptible amino acid residue. Therefore the minimum mass difference between two differentially labeled peptides is only 4 Da compared to 8 Da difference in SID-SRM. Due to low mass difference between different isotopic labeles, the chances of overlapping for the isotopic envelopes from their differentially labeled versions increases.[26] In dimethyl-SRM, this overlap may result in the co-isolation of 2 differentially labeled species by the Q1 filter (with a resolution of 0.7Da FWHM) and thus may reduce the assay accuracy. This potential effect would be eliminated if choosing to use only light and heavy labels, which would give a 8 Da difference, and could also be avoided by limiting the targeted peptide length (7-17AA), charge state and amino acid composition to minimize such overlaps.

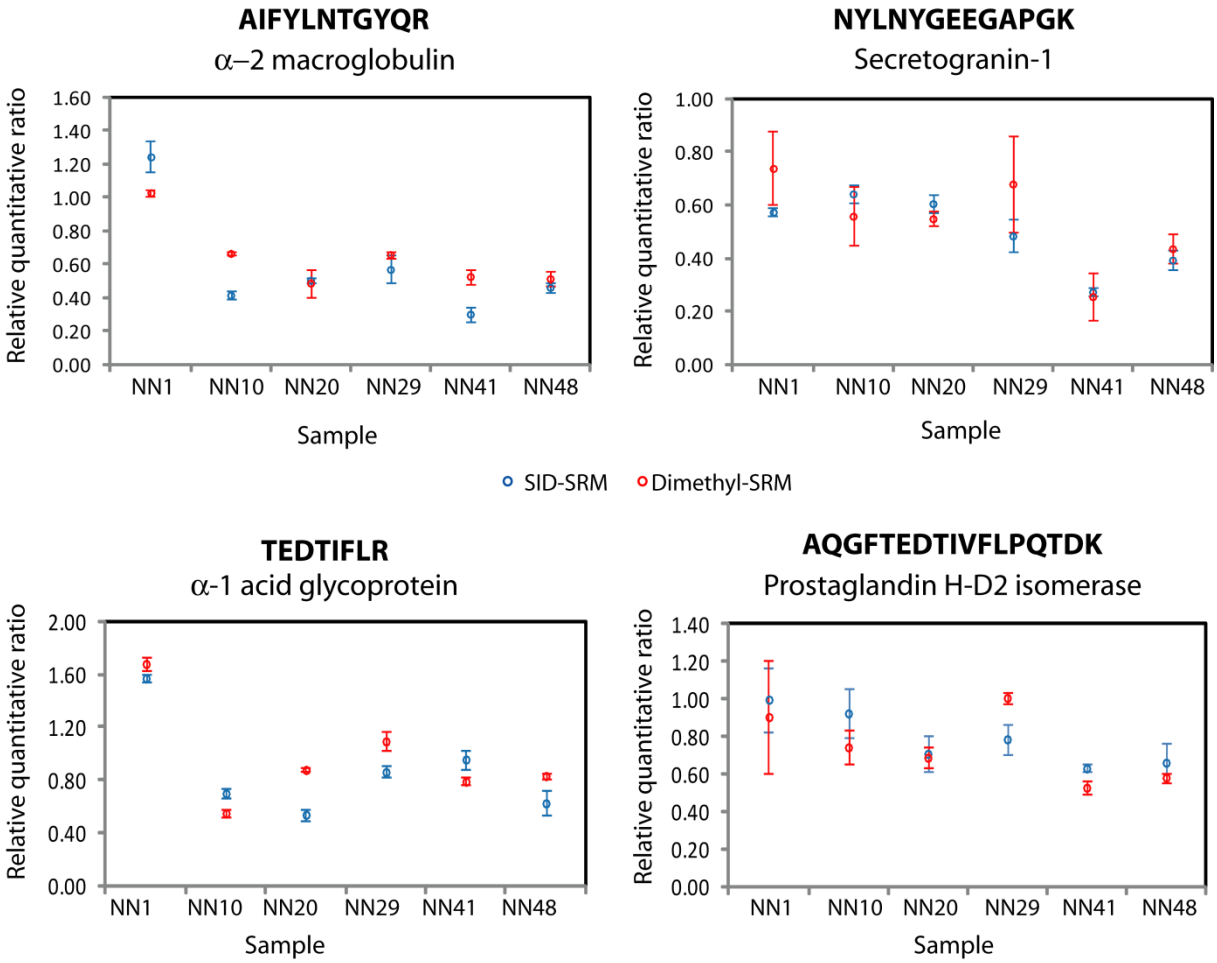

Figure 4. Correlation between dimethyl-SRM and SID-SRM. The performance of SID-SRM and dimethyl-SRM were evaluated based on four peptides quantified. The standard deviations of each peptide for three process replicates were depicted as error bar.

CONCLUSION

In summary, by combining a simple whole proteome in-solution digest with dimethyl labeling and an SRM, we present a flexible workflow to monitor relative changes in protein expression in biological sample. This approach is fast, straightforward and inexpensive. Besides, by multiplexing three different labels, including the GIS, in one MS analysis, sample throughput is increased. Dimethyl-SRM is able to quantify proteins over a wide dynamic range, with acceptable reproducibility (CV<20%) and accuracy (error <20%). Although our results show that SID-SRM (typically CV <10% [15, 16, 33]) outperforms dimethyl-SRM in quantitative accuracy and precision, dimethyl-SRM can still be a complementary method with advantages over using SID-SRM in certain areas such as analyses of post translationally modified peptides, and peptides that are difficult to synthesize. It is also a fast and cost efficient alternative for preliminary or confirmatory targeted screening of proteins before building actual SID-SRM assays.

ACKNOWLEDGEMENTS

The study was supported by the National Program for Research in Functional Genomics (FUGE) funded by the Research Council of Norway, the Western Norway Regional Health Authority, the Meltzer foundation, and the Kristian Gerhard Jebsen Foundation. We thank other members of PROBE for fruitful discussions, especially Olav Mjaavatten and Hilde Kristin Garberg for excellent technical assistance.

Supporting information available

Supplementary data set containing Supplementary Figure S1-4 and Supplementary Table S1-8 is available free of charge at http://pubs.acs.org.

The raw data associated with this manuscript can be downloaded from Tranche using the following hash:

ogLqfOB0B6sV5721y3FBaBn7GV5Uby2lq54HiYJkLDxNN48BZm5mGVaEDZdWUiC0KUXuSSL
dmXzpm6xSsWUb5HMN834AAAAAAAANTA==

REFERENCES

(1) Yan, W.; Chen, S. S. *Brief Funct Genomic Proteomic* **2005**, *4*, 27-38.

(2) Boersema, P. J.; Raijmakers, R.; Lemeer, S.; Mohammed, S.; Heck, A. J. *Nat Protoc* **2009**, *4*, 484-494.

(3) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nature Biotechnology* **1999**, *17*, 994-999.

(4) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. *Mol Cell Proteomics* **2002**, *1*, 376-386.

(5) Qian, W. J.; Monroe, M. E.; Liu, T.; Jacobs, J. M.; Anderson, G. A.; Shen, Y. F.; Moore, R. J.; Anderson, D. J.; Zhang, R.; Calvano, S. E.; Lowry, S. F.; Xiao, W. Z.; Moldawer, L. L.; Davis, R. W.; Tompkins, R. G.; Camp, D. G.; Smith, R. D. *Molecular & Cellular Proteomics* **2005**, *4*, 700-709.

(6) 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. *Mol Cell Proteomics* **2004**, *3*, 1154-1169.

(7) Thompson, A.; Schaefer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R. A. W.; Mohammed, A. K. A.; Hamon, C. *Analytical Chemistry* **2006**, *78*, 4235-4235.

(8) Cox, J.; Mann, M. *Nat Biotechnol* **2008**, *26*, 1367-1372.

(9) Elias, J. E.; Haas, W.; Faherty, B. K.; Gygi, S. P. *Nature Methods* **2005**, *2*, 667-675.

(10) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198-207.

(11) Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. *Mol Syst Biol* **2008**, *4*, 222.

(12) Gallien, S.; Duriez, E.; Domon, B. *J Mass Spectrom* **2011**, *46*, 298-312.

(13) Picotti, P.; Bodenmiller, B.; Mueller, L. N.; Domon, B.; Aebersold, R. *Cell* **2009**, *138*, 795-806.

(14) Wolf-Yadlin, A.; Hautaniemi, S.; Lauffenburger, D. A.; White, F. M. *Proc Natl Acad Sci U S A* **2007**, *104*, 5860-5865.

(15) Keshishian, H.; Addona, T.; Burgess, M.; Kuhn, E.; Carr, S. A. *Mol Cell Proteomics* **2007**, *6*, 2212-2229.

(16) Kuzyk, M. A.; Borchers, C. H.; Smith, D.; Yang, J. C.; Cross, T. J.; Jackson, A. M.; Hardie, D. B.; Anderson, N. L. *Molecular & Cellular Proteomics* **2009**, *8*, 1860-1877.

(17) Deutsch, E. W.; Lam, H.; Aebersold, R. *EMBO Rep* **2008**, *9*, 429-434.

(18) DeSouza, L. V.; Romaschin, A. D.; Colgan, T. J.; Siu, K. W. *Anal Chem* **2009**, *81*, 3462-3470.

(19) DeSouza, L. V.; Krakovska, O.; Darfler, M. M.; Krizman, D. B.; Romaschin, A. D.; Colgan, T. J.; Siu, K. W. *Proteomics* **2010**, *10*, 3108-3116.

(20) Kim, J. S.; Fillmore, T. L.; Liu, T.; Robinson, E.; Hossain, M.; Champion, B. L.; Moore, R. J.; Camp, D. G., 2nd; Smith, R. D.; Qian, W. J. *Mol Cell Proteomics* **2011**, *10*, M110 007302.

(21) Hsu, J. L.; Huang, S. Y.; Chow, N. H.; Chen, S. H. *Anal Chem* **2003**, *75*, 6843-6852.

(22) Munoz, J.; Low, T. Y.; Kok, Y. J.; Chin, A.; Frese, C. K.; Ding, V.; Choo, A.; Heck, A. J. *Mol Syst Biol* **2011**, *7*, 550.

(23) Anderson, N. L.; Anderson, N. G. *Mol Cell Proteomics* **2002**, *1*, 845-867.

(24) Kroksveen, A. C.; Opsahl, J. A.; Aye, T. T.; Ulvik, R. J.; Berven, F. S. *J Proteomics* **2011**, *74*, 371-388.

(25) Lathrop, J. T.; Anderson, N. L.; Anderson, N. G.; Hammond, D. J. *Curr Opin Mol Ther* **2003**, *5*, 250-257.

(26) Cappadona, S.; Munoz, J.; Spee, W. P.; Low, T. Y.; Mohammed, S.; van Breukelen, B.; Heck, A. J. *J Proteomics* **2011**, *74*, 2204-2209.

(27) Boersema, P. J.; Aye, T. T.; van Veen, T. A.; Heck, A. J.; Mohammed, S. *Proteomics* **2008**, *8*, 4624-4632.

(28) Hsu, J. L.; Huang, S. Y.; Chen, S. H. *Electrophoresis* **2006**, *27*, 3652-3660.

(29) Raijmakers, R.; Berkers, C. R.; de Jong, A.; Ovaa, H.; Heck, A. J.; Mohammed, S. *Mol Cell Proteomics* **2008**, *7*, 1755-1762.

(30) Song, C.; Wang, F.; Ye, M.; Cheng, K.; Chen, R.; Zhu, J.; Tan, Y.; Wang, H.; Figeys, D.; Zou, H. *Anal Chem* **2011**, *83*, 7755-7762.

(31) Wu, C. J.; Chen, Y. W.; Tai, J. H.; Chen, S. H. *J Proteome Res* **2011**, *10*, 1088-1097.

(32) Picotti, P.; Rinner, O.; Stallmach, R.; Dautel, F.; Farrah, T.; Domon, B.; Wenschuh, H.; Aebersold, R. *Nat Methods* **2010**, *7*, 43-46.

(33) Anderson, L.; Hunter, C. L. *Mol Cell Proteomics* **2006**, *5*, 573-588.