Combination of Gas-Phase Fractionation and MS3 Acquisition Modes for Relative Protein Quantification with Isobaric Tagging

ARTICLE *in* JOURNAL OF PROTEOME RESEARCH · SEPTEMBER 2012

Impact Factor: 4.25 · DOI: 10.1021/pr300519c · Source: PubMed

3 AUTHORS, INCLUDING:

Loïc Dayon
Nestlé Institute of Health Sciences S.A.
41 PUBLICATIONS **1,113** CITATIONS
SEE PROFILE

Martin Kussmann
Nestlé Institute of Health Sciences S.A.
104 PUBLICATIONS **2,189** CITATIONS
SEE PROFILE

Journal of proteome research

Technical Note

pubs.acs.org/jpr

Combination of Gas-Phase Fractionation and MS^3 Acquisition Modes for Relative Protein Quantification with Isobaric Tagging

Loïc Dayon,*[,†] Bernhard Sonderegger,[†] and Martin Kussmann[†,‡,§]

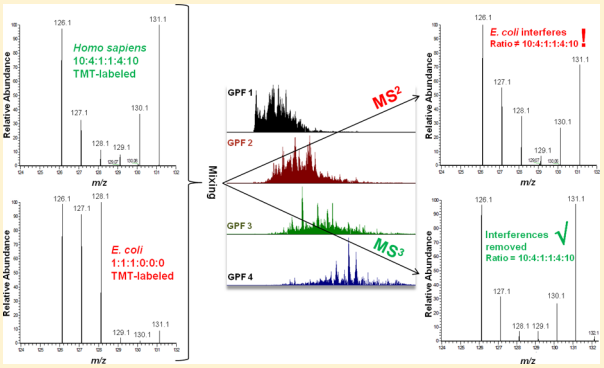
[†]Proteomics and Metabonomics Core, Nestlé Institute of Health Sciences, Lausanne, Switzerland

[‡]Faculty of Life Sciences, Ecole Polytechnique Fédérale Lausanne (EPFL), Lausanne, Switzerland

[§]Faculty of Science, Aarhus University (AU), Aarhus, Denmark

Supporting Information

ABSTRACT: Relative quantification of peptides and proteins with isobaric tags such as iTRAQ or TMT is commonly used in comparative quantitative proteomics based on tandem mass spectrometry (MS/MS). Nonetheless, isobaric tagging inherently suffers from the cofragmentation/interference phenomenon that may compromise the quality of the quantitative data. An MS^3 acquisition mode has been recently proposed to address this issue. Because of the additional ion isolation and fragmentation step, the MS^3 acquisition mode significantly alleviates this interference effect. However, MS^3 acquisition exhibits a lower sensitivity and a higher duty cycle, both of which reduce the number of identified and quantified proteins. In the present study, we evaluated the combination of gas-phase fractionation (GPF) and MS^3 acquisition modes to optimize both identification and quantification of tryptic peptides labeled with TMT using a hybrid ion trap-orbitrap (LTQ-OT) instrument. An interference model was used where TMT-labeled human plasma proteolytic digests were spiked with TMT-labeled *E. coli* proteolytic digests. When combined with GPF, the MS^3 acquisition mode was compared with MS^2 modes such as high-energy collision dissociation (HCD) and combined collision-induced dissociation (CID)/HCD. We demonstrated the benefit of using both GPF and MS^3 to analyze tryptic peptides labeled with TMT in terms of quantification precision and accuracy as well as proteome coverage. We further explored parameters such as the influence of automatic gain control and additional MS^3 scans. The TMT-GPF-MS^3 workflow was shown to be a powerful alternative for quantitative proteomic studies that offers improved identification/quantification accuracy and enhanced proteome coverage without the need for extensive sample fractionation before MS analysis.



KEYWORDS: *isobaric tag, quantitative proteomics, tandem mass tag, gas phase fractionation, third-stage MS*

INTRODUCTION

Relative protein quantification with mass spectrometry (MS) that uses stable isotope mass labels is widely employed in proteomics. Such an approach consists of the differential labeling of two or more peptide or protein samples with chemical tags that can be discriminated with MS. This methodology enables pooling of the samples after labeling that thereby offers multiplexing and reduces bias induced by parallel sample handling. Isobaric labeling such as tandem mass tags (TMTs) and isobaric tags for relative and absolute quantification (iTRAQ) enables multiplexed quantification of peptides and proteins with tandem MS (MS/MS or MS^2);[1−3] the different labels incorporated into the free amino-terminus and epsilon-amino groups of lysine residues of proteolytic peptides are revealed after MS^2 through release of their specific reporter ions. Relative peptide (and protein from inference) quantification between samples and conditions is achieved by comparing the peak intensities of the reporter ions. Quantification performance might nonetheless be compromised by cofragmentation of more than one peptide species at a time that results in inaccurate measurement and ratio compression.[4] In particular, quantitative ratio compression results from the fact that, in complex mixtures, a majority of proteins displays a 1:1 ratio between experimental conditions as a consequence of the generally stable concentrations of many proteins within the proteome of investigation. Therefore, when using isobaric tagging-based approaches, these "constant" proteins act as a background that indeed reduces the ratios of truly differentially expressed proteins when cofragmentation/interference occurs.

Gas-phase fractionation (GPF) across the m/z axis[5] and data-independent acquisition strategies[6] such as precursor acquisition independent from ion count (PAcIFIC)[7] and SWATH−MS[8] use the mass spectrometer to fractionate complex peptide mixtures and require reduced degree of liquid- or gel-based

Received: June 11, 2012
Published: September 4, 2012

fractionation before reversed-phase liquid-chromatography (RP-LC) MS analysis. GPF across the m/z axis basically consists of splitting the MS analysis into several smaller m/z windows and reinjection of the sample several times to cover the full m/z range of the precursor ions (e.g., typically m/z = 300−1500). GPF combined with data-dependent acquisition was previously shown to enhance proteome coverage and reduce under-sampling, compared to classical data-dependent methods without GPF.[5] However, due to the absence of any upstream sample fractionation, GPF-like methods may nevertheless suffer from increased probability of cofragmentation events during RP-LC−MS/MS analysis that results in poorer accuracy and precision when performing quantification with isobaric tags.[9]

Several groups have proposed solutions to cope with this cofragmentation/interference issue. For instance, Bantscheff and co-workers used delayed fragmentation and optimized isolation width for high-energy collision dissociation (HCD) using a hybrid linear ion trap-orbitrap (LTQ-OT) instrument.[10] Coon and colleagues performed gas-phase purification by proton-transfer ion−ion reactions to reduce the charge state of precursor ions and the probability of concurrent isolation of multiple peptide ions.[11] Gygi's group has recently introduced a new quantification pipeline based on MS^3 acquisition that eliminates ratio distortion in isobaric tagging experiments.[12] In this approach, first-generation fragments obtained from collision-induced dissociation (CID) in the LTQ (i.e., MS^2) are further fragmented by HCD that reveals the reporter ions in the OT analyzer (i.e., MS^3). Because of the additional isolation and fragmentation, the MS^3 acquisition mode significantly reduces the interference effect due to the lower chance of isobaric peptides that also present isobaric fragments. The MS^3-based quantification method reported by Gygi and co-workers uses the protease Lys-C for protein digestion, which ensures that isobaric tags are incorporated at both peptide termini such that all generated fragments during MS^2 are quantifiable with MS^3.

In view of these recent developments, we evaluated the combination of GPF and MS^3 for relative protein quantification of complex mixtures with isobaric tagging. We thereby aimed to (i) reduce the need for sample preparation/fractionation before RP-LC−MS/MS while preserving accuracy and precision in a multiplexed quantitative proteomic experiment; and (ii) compensate the lower acquisition speed of the MS^3 mode and enhance proteome coverage by fine-tuned sample reinjections using GPF. As previously described,[12] an interference model was used to evaluate the quantification performance obtained with TMTs: the to-be-compared TMT-labeled human plasma proteolytic digests were spiked with TMT-labeled *E. coli* proteolytic digests, which acted as the interfering background. Furthermore, we evaluated the use of both trypsin and endoproteinase Lys-C for protein digestion. When combined with GPF, the MS^3 acquisition mode was compared with MS^2 acquisition modes such as HCD and combined CID/HCD[13−15] and assessed in terms of quantification precision and accuracy as well as proteome coverage. We also explored the influence of automatic gain control, normalized collision energy (NCE) and additional scans to optimize the TMT-GPF-MS^3 workflow.

EXPERIMENTAL SECTION

Materials

Iodoacetamide (IAA), tris(2-carboxyethyl) phosphine hydrochloride (TCEP), triethylammonium hydrogen carbonate buffer (TEAB) 1 M pH = 8.5 and sodium dodecyl sulfate (SDS) were purchased from Sigma (St. Louis, MO). Formic acid (FA, 99%) was from BDH (VWR International Ltd., Poole, U.K.). Hydroxylamine solution 50 wt % in H_2O (99.999%) was from Aldrich (Milwaukee, WI). Water (18.2 MΩ.cm at 25 °C) was obtained from a Milli-Q apparatus (Millipore, Billerica, MA) and acetonitrile was from BDH. The sixplex tandem mass tags (TMTs) were purchased from Thermo Scientific (Rockford, IL). Protein assay dye reagent concentrate was from Bio-Rad (Hercules, CA). Sequencing grade modified trypsin was from Promega (Madison, WI). *E. coli* sample (Bio-Rad) was dissolved in H_2O and precipitated with prechilled acetone (−20 °C) overnight. After centrifugation at 8000× *g* for 10 min (4 °C), acetone was decanted and the pellets were dried for 10 min. Lyophilized Human plasma (Sigma) was resuspended in 1 mL H_2O. Aliquots were kept at −20 °C.

Sample Preparation

The protein concentration in the studied samples was determined by the Bradford method.[16] Each human plasma and *E. coli* protein sample (400 μg) was dissolved in 380 μL of TEAB 100 mM and 20 μL of SDS 2%. A volume of 21.2 μL TCEP 20 mM was added and incubation was performed for 1 h at 55 °C. A volume of 22 μL IAA 150 mM was added (incubation for 1 h in the dark). Enzymatic digestion was performed by addition of 40 μL trypsin or Lys-C (both at 0.25 μg/μL in TEAB 100 mM) and incubation overnight at 37 °C. Human plasma protein tryptic and Lys-C digests were each aliquoted according to the following protein amounts: 100 μg (i.e., 120.8 μL), 40 μg (i.e., 48.32 μL), 10 μg (i.e., 12.08 μL), 10 μg (i.e., 12.08 μL), 40 μg (i.e., 48.32 μL), 100 μg (i.e., 120.8 μL) and completed with addition of 0, 72.48, 108.72, 108.72, 72.48, 0 μL of buffer to equalize the sample volumes before labeling with sixplex TMT. *E. coli* protein tryptic and Lys-C digests were each aliquoted in 3 × 100 μg of proteins (i.e., 120.8 μL). TMT labeling was performed by addition of 0.8 mg sixplex TMT reagent in 41 μL CH_3CN (incubation for 1 h at room temperature). Human plasma protein tryptic and Lys-C digests containing 100, 40, 10, 10, 40, and 100 μg were respectively labeled with reagents with reporter ions at m/z = 126.1, 127.1, 128.1, 129.1, 130.1, and 131.1 Th. *E. coli* protein tryptic digests were labeled with reagents with reporter ions at m/z = 126.1, 127.1, and 128.1 Th. *E. coli* protein Lys-C digests were labeled with reagents with reporter ions at m/z = 129.1, 130.1, and 131.1 Th. After reaction, a volume of 8 μL of hydroxylamine 5% in H_2O was added to each tube to react for 15 min. All trypsin-digested samples were pooled together as were all Lys-C-digested samples. The 2 pooled samples were further purified with Oasis HLB cartridges (1 cm^3, 30 mg) from Waters (Milford, MA) followed by strong cation-exchange solid-phase extraction using home-prepared columns packed with SP Sepharose Fast Flow (Sigma). Samples were then evaporated to dryness before storage at −20 °C.

RP-LC−MS/MS

The samples were dissolved in H_2O/CH_3CN/FA 96.9/3/0.1 for RP-LC−MS/MS analysis. LC−MS/MS was performed on a LTQ orbitrap Elite from Thermo Scientific (San Jose, CA)

dx.doi.org/10.1021/pr300519c | J. Proteome Res. 2012, 11, 5081−5089

Table 1. Number of MS^n Spectra, Identifications in the *Homo sapiens* and *E. coli* Databases, and Missing Quantitative Values for Human Peptides in TMT-GPF Experiments[a]

acquisition mode	enzyme used	number of ions used for quantification (i.e., AGC)	NCE 1/%	NCE 2/%	average number of MS^2 spectra per analysis	average number of MS^3 spectra per analysis	average number of Mascot queries	average number of identified unique peptides[b]	quantification success of the most intense reporter ions (i.e., 126/131)	total missing reporter-ion values
CID/HCD	Trypsin	100000	30	60	48206	0	23779	3602	99.8%	2.4%
HCD	Trypsin	100000	35		29100	0	28484	3469	99.6%	2.7%
MS^3	Trypsin	100000	30	60	25363	20989	24931	3525	55.6%	59.9%
MS^3	Trypsin	200000	30	60	22979	17081	22611	3077	56.2%	55.6%
MS^3	Trypsin	300000	30	60	21416	15151	21089	2811	59.2%	50.6%
$2MS^3$	Trypsin	300000	30	60	15137	19510	14880	2393	64.6%	44.6%
MS^3	Lys-C	100000	30	60	23606	18384	22857	1889	60.6%	57.1%
MS^3	Lys-C	200000	30	60	21983	16939	21230	1864	58.9%	62.0%
MS^3	Lys-C	300000	30	60	20745	15474	19944	1808	64.3%	54.8%
HCD[c]	Lys-C	250000	45		n.a.	0	n.a.	43656	97.7%	4.5%
MS^{3c}	Lys-C	200000	35	60	n.a.	n.a.	n.a.	38122	88.7%	22.3%

[a]Data based on 3 replicates. [b]In both *Homo sapiens* and *E. coli* databases. [c]Data retrieved from Ting et al., 2011; in that study performed with an LTQ-OT Velos (Thermo Scientific), the data was obtained from 20 strong cation-exchange fractions and is provided here for *E. coli* peptides since the human proteome from HeLa cells served as the interference matrix.

equipped with an Easy-nLC 1000 (Thermo Scientific). Proteolytic peptides (1 μg) were trapped on a Acclaim Pepmap 75 μm × 2 cm (C18, 3 μm, 100 Å) precolumn (Dionex, Olten, Switzerland). Following washing, they were separated on a Acclaim PepMap RSLC 75 μm × 15 cm (C18, 2 μm, 100 Å) column (Dionex) coupled to a stainless steel nanobore emitter (40 mm, OD 1/32") (Thermo Scientific). The analytical separation was performed for 80 min with a gradient of H_2O/FA 99.9/0.1 (solvent A) and CH_3CN/FA 99.9/0.1 (solvent B). The gradient was performed as follows: 0−2 min 95% A and 5% B, then to 70% A and 30% B at 62 min, and 20% A and 80% B at 72 min at a flow rate of 220 nL min^{-1}, followed by re-equilibration of the column. For MS survey scans, the OT resolution was 60000 and the ion population was 1×10^6 with an m/z window from 300 to 1500 Th or the specified GPF m/z window (see Results and Discussion part). For MS^2 in the LTQ with CID activation, the ion population was 1×10^4 (isolation width of 2 Th) with a maximum injection time of 150 ms. For MS^2 detection in the OT with HCD activation, ion population was set to 1×10^5 (isolation width of 2 Th), with a resolution of 15000, first mass at m/z = 100 Th, and a maximum injection time of 750 ms. For the MS^3 acquisition mode, MS^2 scans were performed in the LTQ with CID as described above followed directly after by MS^3 scans in the OT after HCD activation of the most intense MS^2 fragment ion comprised in a 110−160% m/z window of the precursor-ion m/z. For MS^3, ion population was set to 1×10^5, 2×10^5, or 3×10^5 (isolation width of 4 Th), with a resolution of 15000, first mass at m/z = 100 Th, and a maximum injection time of 750 ms (charge state screening was disabled). A maximum of 10 precursor ions (most intense) were selected for activation and subsequent MS^n analyses. CID was performed at 30% of the normalized collision energy (NCE) in all cases. Different collision energies were evaluated for HCD, that is, 25, 30, 35, 40, 45, 50, and 60% NCE (see Table 1 and Supporting Information). The different operating modes evaluated for identification and quantification of peptide/protein are detailed in the Results and Discussion part. Briefly, one acquisition mode used HCD alone as fragmentation method. Another performed parallel fragmentation of a given precursor with CID and HCD (CID/HCD mode).[13] The MS^3 acquisition mode consisted in the fragmentation of a given precursor with CID followed by HCD of a selected fragment ion.[12] Control experiments were carried out with CID alone for peptide/protein identification considerations and GPF m/z window determination. In this case, the 20 most intense precursor ions were selected. All analyses were performed at least in triplicate (e.g., 12 repeats for single MS analyses and 3 repeats for each GPF fraction).

Data Analysis

Proteome Discoverer (version 1.3.0.339, Thermo Scientific) was used as data analysis interface with the possibility to select scan events for peptide/protein identification and/or quantification. Identification was performed in the Swiss-Prot database (57.15) with only *Homo sapiens* and *E. coli* entries (42912 entries). Mascot (version 2.03.02, Matrix Sciences, London, U.K.) was used. Variable amino acid modifications were oxidized methionine. TMT-labeled peptide amino terminus and TMT-labeled lysine (+ 229.163 Da) were set as fixed modifications as well as carbamidomethylation of cysteines. Trypsin or Lys-C was selected as the enzyme, with one potential missed cleavage. Peptide and fragment ion tolerance was respectively 20 ppm and either 0.6 or 0.02 Da when fragments were recorded in the LTQ or in the OT. The percolator[17] module was used and Mascot "ion score" cutoff value was derived to 13. The average false discovery rates were about 1%.[18] The Proteome Discoverer software was used to extract the reporter-ion abundance values for the human peptides. No individual protein quantification was carried out. The Log2 or Log10 of the reporter-ion ratios were used to evaluate the quantification performances.

RESULTS AND DISCUSSION

Increased Number of Identified TMT-labeled Peptides and Proteins When Combining GPF and MS^3

A 1:1 mixture of the sixplex-TMT-labeled proteolytic peptides obtained from a human plasma sample and an *E. coli* cell lysate was analyzed on an LTQ-OT. This mixture consisted of TMT-labeled human proteolytic peptides in relative amounts 10:4:1:1:4:10 and interfering TMT-labeled proteolytic peptides from *E. coli* in relative amounts 1:1:1:0:0:0 or 0:0:0:1:1:1 depending on the enzyme used (see Experimental Section for

dx.doi.org/10.1021/pr300519c | J. Proteome Res. 2012, 11, 5081−5089

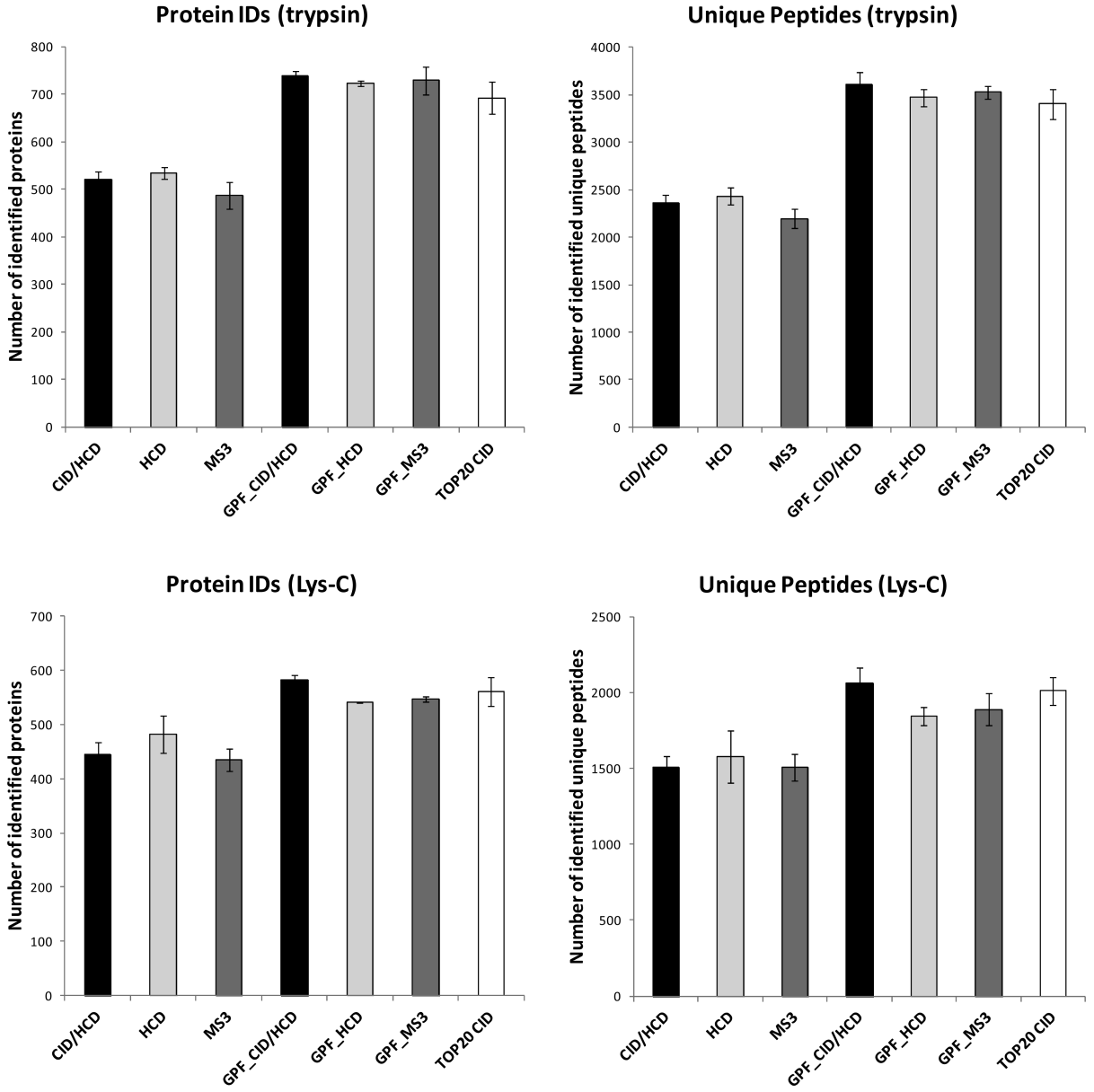


Figure 1. Average number of proteins and unique peptides identified in the *Homo sapiens* and *E. coli* databases as a function of the different acquisition modes (i.e., CID/HCD, HCD, and MS^3 (automatic gain control (AGC) = 1×10^5) with and without GPF) and of the proteolytic enzyme used for digestion. In all cases, a data-dependent analysis of the 10 most intense precursor ions in the MS survey scans was used. For single MS analyses, 12 injections were made ($n = 12$) and for GPF analyses 3 injections of each of the 4 m/z fractions were performed ($n = 3$). Reference number of identifications provided by CID alone with data-dependent analysis of the 20 most intense ions in the MS survey scans is also given ($n = 3$) (see Experimental Section).

details). GPF mass windows were first determined experimentally by analyzing the 1:1 mix with MS/MS. A data-dependent acquisition method that performs MS^2 with CID on the 20 most intense ions detected in the MS survey scans was used. Four GPF mass windows were defined. Each of the GPF windows was designed to contain the same number of MS^2 events. For tryptic peptides, the GPF windows were m/z = 300−536, 531−679.5, 674.5−840.5, and 835.5−1500 Th. For peptides that result from the digestion with Lys-C, the GPF windows were m/z = 300−542.5, 537.5−695, 690−876, and 871−1500 Th. Because the endoproteinase Lys-C digests the proteins only at the C-terminus of lysine residues, the GPF windows were slightly shifted to higher m/z values with respect to those determined with tryptic peptides.

Several MS^n acquisition methods were evaluated for relative quantification of TMT-labeled peptides, CID/HCD,[13] HCD, and MS^3.[12] For all these methods, normalized collision energies (NCEs) were optimized for higher numbers of identifications of peptides/proteins with our LTQ-OT. For the CID/HCD acquisition mode, we used NCEs of 30 and 60% for CID and HCD, respectively. When HCD was used alone for fragmentation, we evaluated an NCE of 35% to maximize the peptide/protein identifications (see Supporting Information). For the MS^3 mode, we used NCEs of 30 and 60% for CID (MS^2) and HCD (MS^3), respectively. All these methods that actually used HCD scans targeted a maximum of 10 precursor ions selected in the MS survey scans. The combination of GPF with all these acquisition modes was evaluated.

dx.doi.org/10.1021/pr300519c | *J. Proteome Res.* 2012, 11, 5081−5089

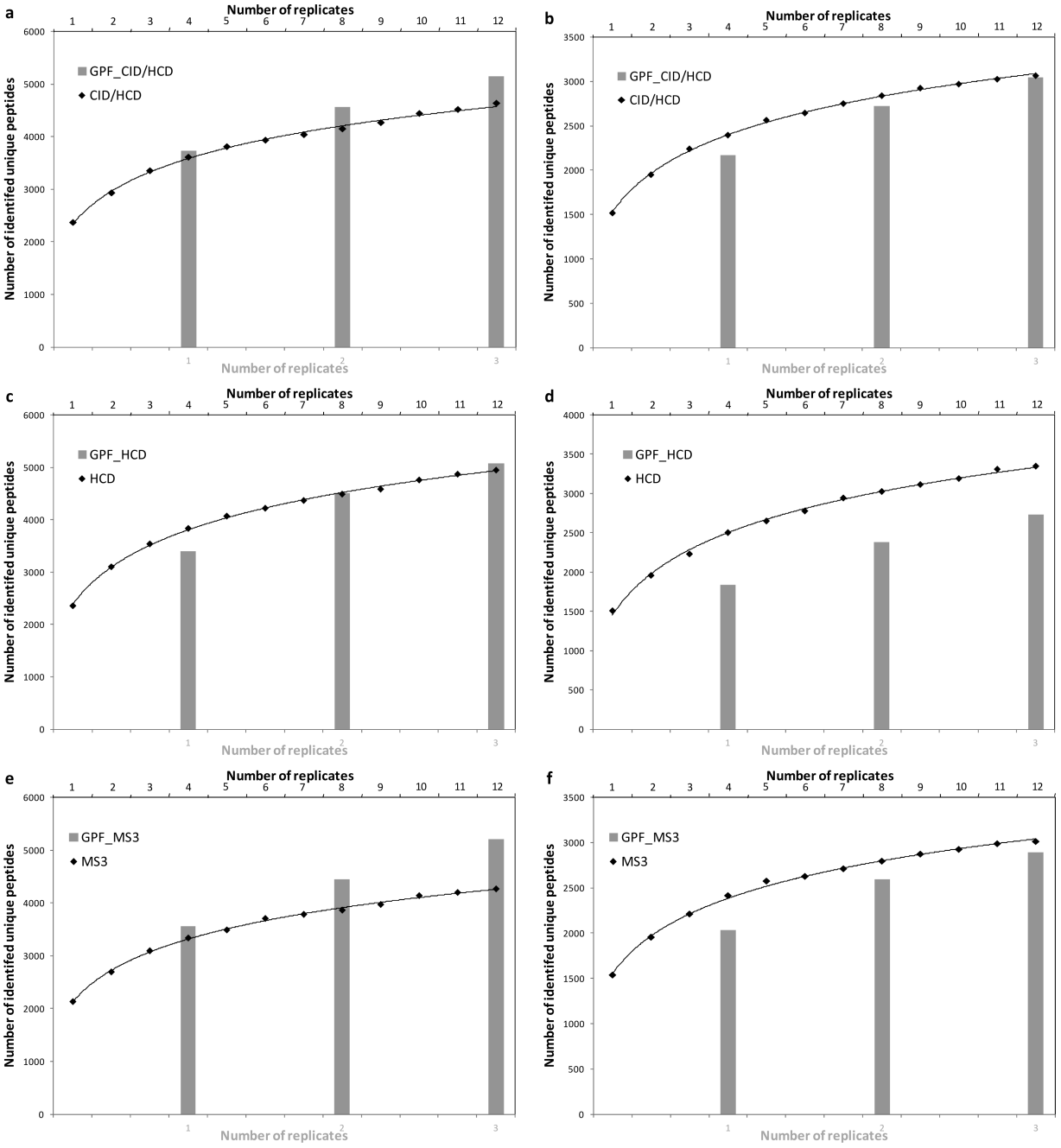


Figure 2. Incremental number of unique peptides identified in the *Homo sapiens* and *E. coli* databases as a function of the number of LC−MS/MS analyses (i.e., injections) and dependent on the acquisition mode following protein digestion with trypsin or Lys-C. CID/HCD (a and b), HCD (c and d), and MS^3 (AGC = 1×10^5) (e and f) acquisition modes with and without GPF were used. One microgram of proteolytic peptides of TMT-labeled human plasma and *E. coli* digests was loaded on the RP-LC column.

GPF provided an increased coverage of the proteomes in the studied samples (i.e., mixture of human plasma and lysate of *E. coli* cell lysate) compared to a single MS analysis over the full m/z range. This observation was true for any of the proteolytic enzymes used for protein digestion as well as for any of the acquisition modes that were tested (Figure 1). On average, an increase of 42% and 23% in the number of identified proteins was achieved respectively for samples digested with trypsin and Lys-C when using GPF. It actually reached the proteome coverage obtained with CID only with a maximum of 20 precursor ions selected in the MS survey scans. GPF triplicate results were also compared with the results of 4 combined single MS analyses in triplicate. For the CID/HCD, HCD, and MS^3 acquisition modes, an increase of 17, 6, and 21% in the number of identified proteins was obtained, respectively, for the sample digested with trypsin when using GPF (+4, −8, and +10% respectively at the unique peptide level). For the sample digested with Lys-C, the number of identified proteins increased by 5% for the MS^3 mode (−8% at the unique peptide level) and decreased by 3 and 12% for the CID/HCD and HCD modes, respectively (−11 and −23% at the unique peptide level) when using GPF. Further, after normalizing the identification output to the invested analysis time (Figure 2), it clearly appeared that GPF was less efficient than multiple

dx.doi.org/10.1021/pr300519c | *J. Proteome Res.* 2012, 11, 5081−5089

injections when digesting proteins with Lys-C. This result was particularly apparent for the HCD acquisition mode, which is the faster of the acquisition modes used herein. For single MS analyses, the MS^3 acquisition mode was shown to provide less identified proteins and unique tryptic peptides than the CID/HCD and HCD acquisition modes (Figures 1 and 2). Application of GPF corrected with discrepancy (Figure 1). In accordance with a previous report,[5] the benefit of using GPF with tryptic peptides was clearly demonstrated. This result was still true even though the LTQ-OT instrument used in this study was much faster than the one used in that report. We therefore decided to focus on the quantification of tryptic peptides combined with GPF.

The use of GPF intends to minimize both sample handling before MS analysis and the loss in analytical depth intrinsic to the MS^3 method (e.g., reduced acquisition speed due to the sensitivity loss). The comparison of the GPF (4 m/z fractions) with a 2-dimensional LC workflow[19] (4 retention-time fractions; see Supporting Information) showed that both approaches provided similar proteome coverage at the level of tryptic digest analysis. Only 2% more protein identifications were obtained with GPF for that level of fractionation (Supporting Information). In terms of peptide/protein identification, we found that the use of trypsin resulted in significantly higher proteome coverage compared to that of Lys-C (Figures 1 and 2, and Supporting Information). In addition to the higher number of peptides produced with trypsin with respect to Lys-C, the higher charge distribution of Lys-C peptides may reduce their identification efficiency (see Supporting Information).[20] Gygi and co-workers[12] presented their data on peptides obtained from proteolytic digestion with Lys-C because these peptides are both TMT-labeled at their N- and C-termini, ensuring the release of TMT reporter ions from both b- and y-type ions during MS^3. Some y-type fragment ions from tryptic peptides (i.e., those with a terminal arginine residue) do not release TMT reporter ions and are not quantifiable. Nonetheless, we showed here that the significantly higher number of tryptic peptides may actually compensate this limitation.

Quantification of Tryptic Peptides when Combining Isobaric Labeling with GPF

As a first evaluation of the quantification performance obtained with isobaric tagging after GPF, we calculated for the human tryptic peptides the ratios between reporter ions at m/z = 126.1 and 131.1 Th, which were either subject or not subject to interferences from the *E. coli* tryptic peptides. Figure 3 shows that MS^3-based quantification definitively provided better performance. The curve maximum for the distribution of $\mathrm{Log2}(i_{126.1}/i_{131.1})$ was indeed closer to an x-axis value of 0 for the MS^3 mode, thereby showing better accuracy. The smaller full-width half-maximum and enhanced curve symmetry attested to better precision. This fact was further illustrated by a $\mathrm{Log2}(i_{126.1}/i_{131.1})$ median value of 0.25, 0.22, and −0.04 and a $\mathrm{Log2}(i_{126.1}/i_{131.1})$ average value of 0.46, 0.40, and −0.03 for respectively the CID/HCD, HCD, and MS^3 modes. As expected and previously reported,[12] the MS^3 mode clearly enhanced the quality of the quantitative data.

We then evaluated in more detail the effects of the interference caused by the *E. coli* peptides on the quantification of human peptides. In the mass-tag channels with reporter ions at m/z = 129.1, 130.1, and 131.1 Th, a clear separation appeared between $i_{129.1}/i_{131.1}$ and $i_{130.1}/i_{131.1}$ (theoretical ratio of

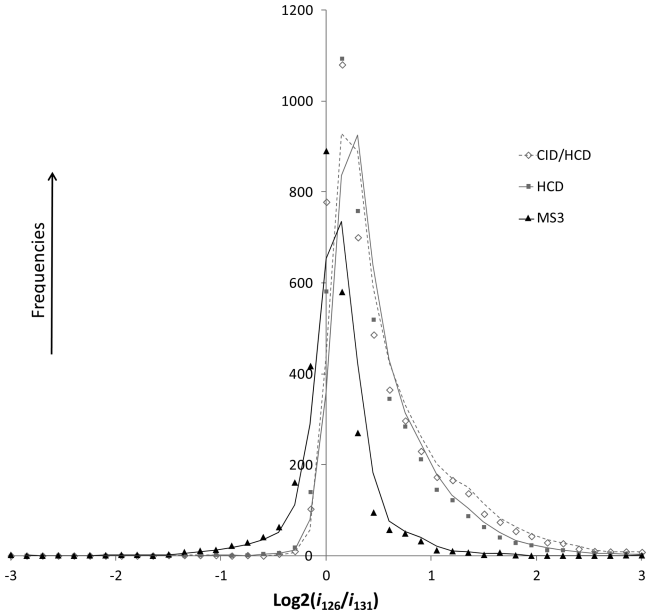


Figure 3. Distribution representation of $\mathrm{Log2}(i_{126.1}/i_{131.1})$ for the identified human tryptic peptides as a function of the acquisition mode (CID/HCD, HCD, and MS^3 (AGC = 3×10^5)) for tryptic peptides labeled with TMT. Data was obtained from the triplicate MS analyses with GPF (m/z = 300−536, 531−679.5, 674.5−840.5, and 835.5−1500 Th).

1:10 and 1:4) because these channels were free of interference (Figure 4b). This separation occurred whatever the acquisition mode used. On the contrary, for channels with reporter ions at m/z = 126.1, 127.1, and 128.1 Th, a clear effect of the interference was seen (Figure 4a). Under these circumstances, MS^3-based quantification helped by substantially removing the interference induced by the *E. coli* tryptic peptides. A clear separation was shown between ratios 1:10 and 4:10 with the MS^3 acquisition mode.

Nevertheless, the number of quantified peptides was less with the MS^3 acquisition mode with respect to the CID/HCD and HCD acquisition modes. This lower number was due on the one hand to the use of tryptic peptides and, on the other hand, to the decreased sensitivity of the MS^3 acquisition compared to MS^2 (Table 1). Actually, missing values (i.e., absence of reporter-ion signal) were a more common feature with the MS^3 acquisition mode. The use of either MS^3 or MS^2 to record reporter ions should be therefore chosen carefully, to optimize the trade-off between the number of quantifiable peptides/proteins and the performance required for the quantification.

In the context of limited upstream sample fractionation and the concomitant use of GPF, the MS^3 acquisition mode restored the quantification performances of isobaric tags, through alleviation of cofragmentation/interference, which is a major issue with MS^2 acquisition modes (i.e., CID/HCD and HCD).

Optimized Combination of GPF and MS^3 Acquisition Modes for Relative Quantification of Tryptic Peptides with Isobaric Tagging

The quantification performance of MS^3 combined with GPF was evaluated for several automatic gain controls (AGCs). It was shown that increasing the value of AGC increased the quantification performance significantly (Figure 5) but compromised the number of identified peptides (Figure 6

dx.doi.org/10.1021/pr300519c | J. Proteome Res. 2012, 11, 5081−5089

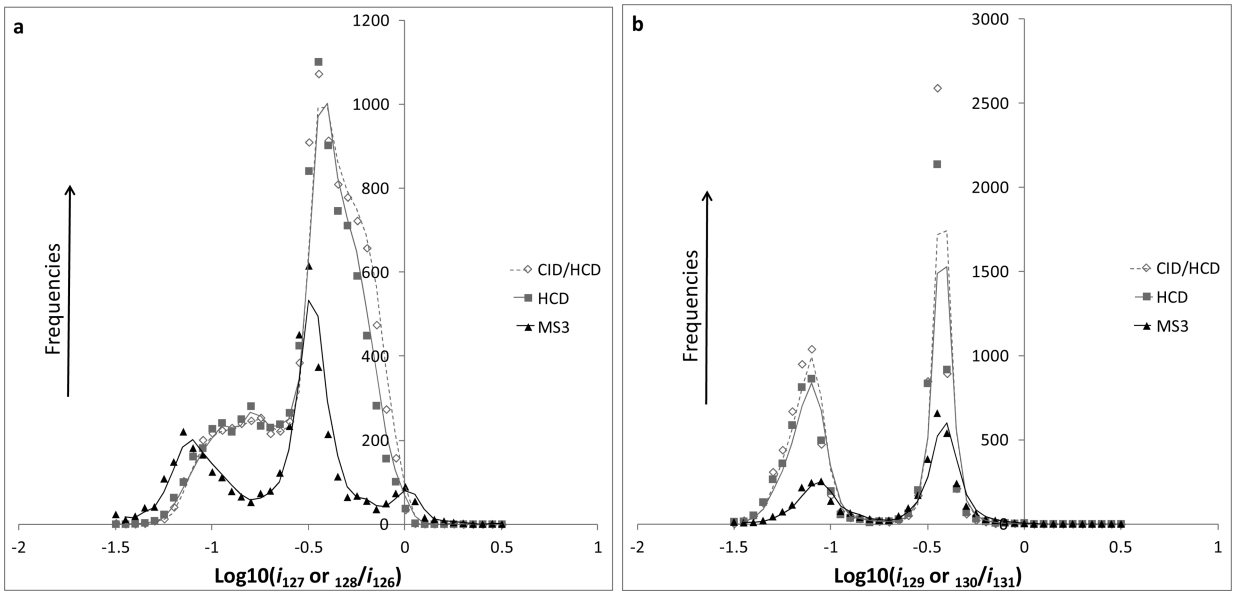


Figure 4. Distribution representation of (a) Log10($i_{127.1\text{ or }128.1}/i_{126.1}$) and (b) Log10($i_{129\ 0.1\text{or }130.1}/i_{131.1}$) for the identified human tryptic peptides as a function of the acquisition mode (CID/HCD, HCD, and MS^3 (AGC = 3×10^5)). Data was obtained from the triplicate MS analyses with GPF (m/z = 300−536, 531−679.5, 674.5−840.5, and 835.5−1500 Th).

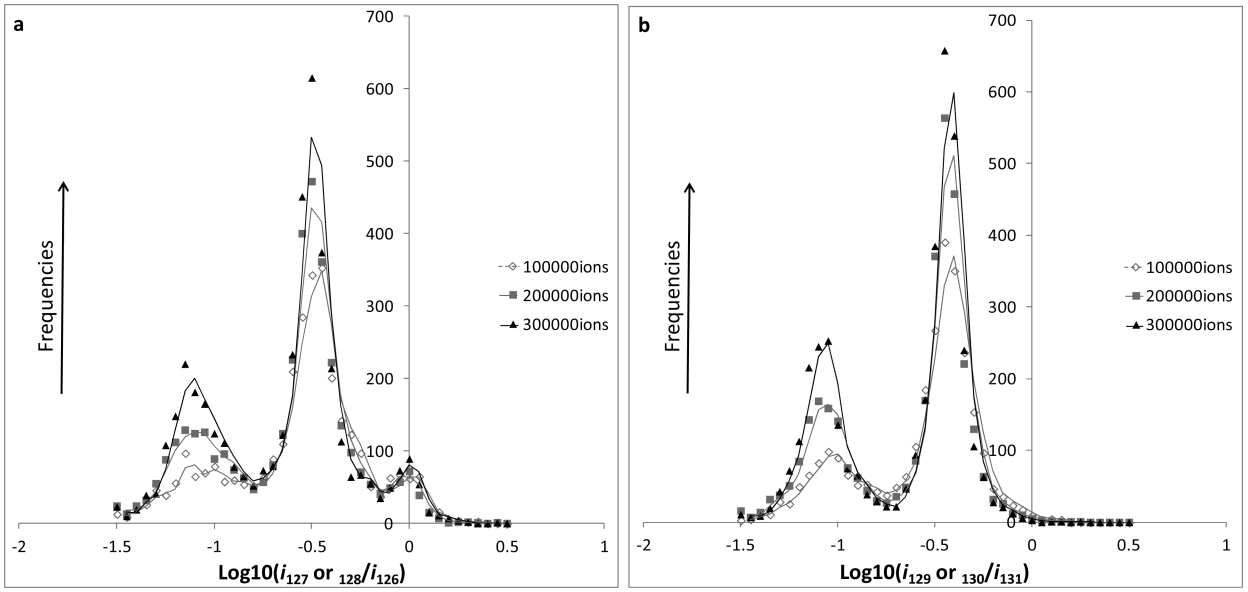


Figure 5. Distribution representation of (a) Log10($i_{127.1\text{ or }128.1}/i_{126.1}$) and (b) Log10($i_{129.1\text{ or }130.1}/i_{131.1}$) for the identified human tryptic peptides as a function of the AGC used for the MS^3 acquisition mode. Data was obtained from the triplicate MS analyses with GPF (m/z = 300−536, 531−679.5, 674.5−840.5, and 835.5−1500 Th).

and Table 1) with a 20% decrease in the number of identified unique peptides when increasing the AGC from 1×10^5 to 3×10^5. The NCE used for the MS^3 scans, on the contrary, had no significant influence. The percentage of missing reporter ions was indeed of 50.6, 49.0, and 53.1% for NCE of 60, 50, and 40% (data not shown).

The results from Table 1 confirmed that tryptic peptides can be quantified with MS^3 as an alternative to the quantification of peptides obtained from the digestion with endoproteinase Lys-C. The absolute number of quantifiable peptides with isobaric tags was indeed higher after digestion with trypsin (Table 1). The drawback was that some of the proteins were not quantifiable with tryptic peptides. In our conditions, this was the case for 13.1% of the human and *E. coli* proteins, and 21.8% of the human proteins, where no quantitative data were available even for the most concentrated sample (i.e., labeled with TMT with reporter ion at m/z = 126.1 Th). We calculated that 63% of the identified peptides contained a TMT-labeled C-terminal lysine residue while 37% had a sequence terminating by an arginine. For MS^3 with AGC = 3×10^5, the percentage of missing reporter-ion values decreased to 41.5% when considering only the peptides with a C-terminal lysine, and the quantification success of the most intense reporter ions (i.e., at m/z = 126 and 131 Th) reached 71.6% (see also Table 1). As a comparison, we determined that Ting et al. had 22.3% of missing reporter-ion values in their analyses of 20 strong cation-exchange LC fractions (Table 1).[12] Direct comparison was nevertheless difficult due to different samples analyzed by Ting

dx.doi.org/10.1021/pr300519c | *J. Proteome Res.* 2012, 11, 5081−5089

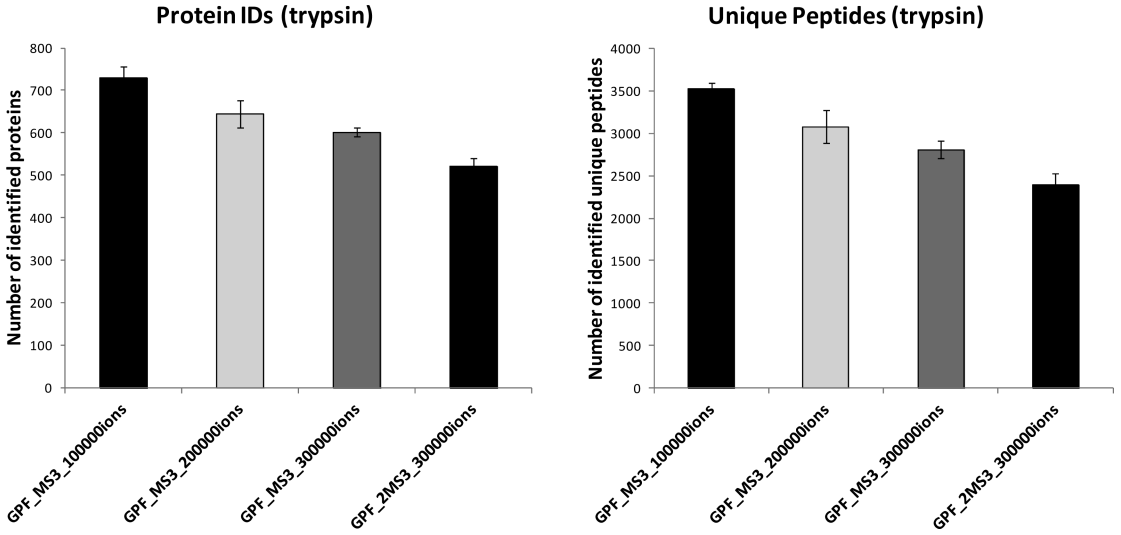


Figure 6. Average number of proteins and unique peptides identified in the *Homo sapiens* and *E. coli* databases as a function of the AGC used; an MS^3 acquisition mode on the 2 most intense fragment ions in the MS^2 spectra was also investigated. One microgram of proteolytic peptides of TMT-labeled human plasma and *E. coli* digests was loaded on the RP-LC column.

et al. and us, and also different loading amounts between the two studies.

In a further optimization effort, we set up a double MS^3 acquisition mode ($2MS^3$) which triggered the two most intense fragment ions in the MS^2 spectra for further MS^3. This methodology aimed to reduce the number of nonquantified tryptic peptides by increasing the probability to indeed select a b-type ion for MS^3 (Table 1 and Supporting Information). Although the acquisition speed was reduced, resulting again in less identified unique peptides (−14.9%; Figure 6), the proportion of quantifiable peptides increased significantly and the number of missing reporter-ion values decreased (Table 1 and Supporting Information). For the $2MS^3$ acquisition mode, the percentage of missing reporter-ion values decreased to 35.5% and the quantification success of the most intense reporter ions reached 76.9% for tryptic peptides with a C-terminal lysine.

CONCLUSION

GPF aims to limit sample prefractionation in a shotgun proteomic approach and uses the mass spectrometer as a fractionation tool in the m/z dimension, instead of for instance LC columns in the retention time dimension. We showed cofragmentation/interference to be an issue for isobaric TMT quantification to analyze very complex peptide mixtures with GPF and MS^2 acquisition modes (i.e., CID/HCD and HCD) on a LTQ-OT. Therefore, the MS^3 acquisition mode was used to restore the quantification performances of isobaric tags when using GPF. Combination of TMT, MS^3 and data-independent acquisition methods[6−8] will be the natural extension of the strategy described herein.

We confirmed that MS^3 provides better quantitative data with isobaric tags compared to MS^2 level quantification[12] and extended it to the quantification of tryptic peptides. Some TMT-labeled tryptic peptides were not quantifiable because some of their fragment ions did not reveal tag moieties. The absolute number of quantified peptides was higher with trypsin than with endoproteinase Lys-C because trypsin yielded significantly more peptide/protein identifications.

Furthermore, GPF was valuable to increase the number of identified peptides when using the MS^3 acquisition mode. The MS^3 mode required both MS^2−CID scans and MS^3−HCD scans for peptide identification and quantification. The MS^3 mode was thus intrinsically slower than the HCD acquisition mode. Moreover, due to the reduced MS^3 sensitivity and need of more ions to be recorded, the duty cycle of the MS^3 acquisition mode was increased compared to the CID/HCD mode, making the MS^3 mode the least efficient of the three evaluated acquisition modes in terms of acquisition speed. However, and importantly, we managed to compensate the decreased acquisition speed of the MS^3 mode by combining MS^3 with GPF, thereby diminishing the under-sampling effect and increasing the proteome coverage.

In conclusion, combining the benefits of GPF (e.g., higher number of identification) and MS^3 (e.g., better quantitative data) was demonstrated to provide an efficient solution for relative quantification of peptides/proteins with isobaric tags such as TMT. This said, increased MS^3 sensitivity is still needed to fully leverage GPF-TMT and more sensitive instruments will help generalizing such an approach. TMT-GPF-MS^3 promises utility for accurate inclusion mass screening[21] deploying, for instance, isobaric channels to generate a calibration curve, as has been demonstrated for proteomics[22,23] or, more recently, for amine-specific metabolomics.[24] We consider our TMT-GPF-MS^3 method as a promising future alternative or complement to second stage MS modes, especially in view of MS instruments and methods becoming more sensitive at the third stage MS level.

ASSOCIATED CONTENT

Supporting Information

Number of MS^2 spectra, identified proteins, and identified unique peptides dependent on NCE used for HCD. Number of unique peptides identified as a function of the number of LC−MS/MS analyses (i.e., injections) and dependent on the acquisition mode. Charge distribution of precursor ions. First-dimensional RP liquid chromatograms. Typical MS base peak chromatograms obtained with GPF and 2-dimensional LC approaches. Evaluation of the quantification performance of the

dx.doi.org/10.1021/pr300519c | J. Proteome Res. 2012, 11, 5081−5089

$2MS^3$ acquisition mode. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Proteomics/Metabonomics Core, Nestlé Institute of Health Sciences SA, Campus EPFL, Quartier de l'innovation, bâtiment G, 1015 Lausanne, Switzerland. E-mail: loic.dayon@rd.nestle.com.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Alexandre Panchaud and John Corthésy are thanked for fruitful discussion and technical support.

REFERENCES

(1) Dayon, L.; Hainard, A.; Licker, V.; Turck, N.; Kuhn, K.; Hochstrasser, D. F.; Burkhard, P. R.; Sanchez, J.-C. Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Anal. Chem.* **2008**, *80* (8), 2921−2931.

(2) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3* (12), 1154−69.

(3) 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.

(4) 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.

(5) Scherl, A.; Shaffer, S. A.; Taylor, G. K.; Kulasekara, H. D.; Miller, S. I.; Goodlett, D. R. Genome-specific gas-phase fractionation strategy for improved shotgun proteomic profiling of proteotypic peptides. *Anal. Chem.* **2008**, *80* (4), 1182−1191.

(6) Venable, J. D.; Dong, M. Q.; Wohlschlegel, J.; Dillin, A.; Yates, J. R. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat. Methods* **2004**, *1* (1), 39−45.

(7) Panchaud, A.; Scherl, A.; Shaffer, S. A.; von Haller, P. D.; Kulasekara, H. D.; Miller, S. I.; Goodlett, D. R. Precursor acquisition independent from ion count: how to dive deeper into the proteomics ocean. *Anal. Chem.* **2009**, *81* (15), 6481−6488.

(8) Gillet, L. C.; Navarro, P.; Tate, S.; Roest, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R. Targeted data extraction of the MS/MS spectra generated by data independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* **2012**, DOI: 10.1074/mcp.O111.016717.

(9) Panchaud, A.; Jung, S.; Shaffer, S. A.; Aitchison, J. D.; Goodlett, D. R. Faster, quantitative, and accurate precursor acquisition independent from ion count. *Anal. Chem.* **2011**, *83* (6), 2250−2257.

(10) Savitski, M. M.; Sweetman, G.; Askenazi, M.; Marto, J. A.; Lang, M.; Zinn, N.; Bantscheff, M. Delayed fragmentation and optimized isolation width settings for improvement of protein identification and accuracy of isobaric mass tag quantification on Orbitrap-type mass spectrometers. *Anal. Chem.* **2011**, *83* (23), 8959−8967.

(11) Wenger, C. D.; Lee, M. V.; Hebert, A. S.; McAlister, G. C.; Phanstiel, D. H.; Westphall, M. S.; Coon, J. J. Gas-phase purification enables accurate, multiplexed proteome quantification with isobaric tagging. *Nat. Methods* **2011**, *8* (11), 933−935.

(12) Ting, L.; Rad, R.; Gygi, S. P.; Haas, W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat. Methods* **2011**, *8* (11), 937−940.

(13) Dayon, L.; Pasquarello, C.; Hoogland, C.; Sanchez, J.-C.; Scherl, A. Combining low- and high-energy tandem mass spectra for optimized peptide quantification with isobaric tags. *J. Proteomics* **2010**, *73* (4), 769−777.

(14) Koecher, T.; Pichler, P.; Schutzbier, M.; Stingl, C.; Kaul, A.; Teucher, N.; Hasenfuss, G.; Penninger, J. M.; Mechtler, K. High precision quantitative proteomics using iTRAQ on an LTQ Orbitrap: a new mass spectrometric method combining the benefits of all. *J. Proteome Res.* **2009**, *8* (10), 4743−4752.

(15) Zhang, Y.; Ficarro, S. B.; Li, S.; Marto, J. A. Optimized Orbitrap HCD for quantitative analysis of phosphopeptides. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (8), 1425−1434.

(16) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248−54.

(17) Kall, L.; Canterbury, J. D.; Weston, J.; Noble, W. S.; MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **2007**, *4* (11), 923−925.

(18) Elias, J. E.; Haas, W.; Faherty, B. K.; Gygi, S. P. Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nat. Methods* **2005**, *2* (9), 667−675.

(19) Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *J. Sep. Sci.* **2005**, *28* (14), 1694−703.

(20) Thingholm, T. E.; Palmisano, G.; Kjeldsen, F.; Larsen, M. R. Undesirable charge-enhancement of isobaric tagged phosphopeptides leads to reduced identification efficiency. *J. Proteome Res.* **2010**, *9* (8), 4045−52.

(21) Whiteaker, J. R.; Lin, C.; Kennedy, J.; Hou, L.; Trute, M.; Sokal, I.; Yan, P.; Schoenherr, R. M.; Zhao, L.; Voytovich, U. J.; Kelly-Spratt, K. S.; Krasnoselsky, A.; Gafken, P. R.; Hogan, J. M.; Jones, L. A.; Wang, P.; Amon, L.; Chodosh, L. A.; Nelson, P. S.; McIntosh, M. W.; Kemp, C. J.; Paulovich, A. G. A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nat. Biotechnol.* **2011**, *29* (7), 625−NIL_108.

(22) Dayon, L.; Turck, N.; Kienle, S.; Schulz-Knappe, P.; Hochstrasser, D. F.; Scherl, A.; Sanchez, J.-C. Isobaric tagging-based selection and quantitation of cerebrospinal fluid tryptic peptides with reporter calibration curves. *Anal. Chem.* **2010**, *82* (3), 848−858.

(23) Dayon, L.; Turck, N.; Scherl, A.; Hochstrasser, D. F.; Burkhard, P. R.; Sanchez, J.-C. From relative to absolute quantification of tryptic peptides with tandem mass tags: application to cerebrospinal fluid. *Chimia* **2010**, *64* (3), 132−135.

(24) Yuan, W.; Anderson, K. W.; Li, S.; Edwards, J. L. Subsecond absolute quantitation of amine metabolites using isobaric tags for discovery of pathway activation in mammalian cells. *Anal. Chem.* **2012**, *84* (6), 2892−2899.