ON THE STATISTICAL SIGNIFICANCE OF COMPRESSED RATIOS IN ISOBARIC LABELLING: A CROSS-PLATFORM COMPARISON

Ana Martinez-Val¹, Fernando Garcia¹, Pilar Ximénez-Embún¹, Nuria Ibarz¹, Eduardo Zarzuela¹,
Isabel Ruppen¹, Shabaz Mohammed^{2,3}, Javier Munoz^{1*}.

¹ Proteomics Unit, Spanish National Cancer Research Centre (CNIO), 28029 Madrid, Spain. ProteoRed-ISCI.

² Department of Biochemistry, New Biochemistry Building, South Parks Road, OX1 3QU Oxford, UK

³ Departments of Chemistry, University of Oxford, Physical & Theoretical Chemistry Laboratory, South Parks Road, OX1 3QZ Oxford, UK

Correspondence to: jmunozpe@cnio.es, (34) 917 328 000 3110

Abstract

Isobaric labelling is gaining popularity in proteomics due to its multiplexing capacity. However, co-peptide fragmentation introduces a bias that undermines its accuracy. Several strategies have been shown to partially and, in some cases, completely solve this issue. However, it is still not clear how ratio compression affects the ability to identify a protein's change of abundance as statistically significant. Here, by using the “two proteomes” approach (*E.coli* lysates with fixed 2.5 ratios in the presence or absence of human lysates acting as the background interference) and manipulating isolation width values, we were able to model isobaric data with different levels of accuracy and precision in three types of mass spectrometers: LTQ Orbitrap Velos, Impact and Q Exactive. We determined the influence of these variables on the statistical significance of the distorted ratios and compared them to the ratios measured without impurities. Our results confirm previous findings¹⁻⁴ regarding the importance of optimizing acquisition parameters in each instrument in order to minimize interference without compromising precision and identification. We also show that, under these experimental conditions, the inclusion of a second replicate increases statistical sensitivity 2-3 fold and counterbalances to a large extent the issue of ratio compression.

Keywords: Isobaric labelling, iTRAQ, accuracy, precision, statistical significance, ratio compression, Isobar.

Introduction

The quantification, by means of mass spectrometry, of proteins and post-translational modifications (PTMs) is a powerful approach to get insights into the molecular mechanisms of a disease or a biological process ⁵. When designing a comparative study several strategies can be adopted, wherein isobaric labelling is highly competitive mainly due to its multiplexing capacity. Isobaric reagents label peptides from distinct experimental conditions with tags that have identical overall mass but vary in the distribution of its heavy isotopes. Relative quantification is then performed upon MS/MS fragmentation of the selected peptides, on the basis of the so-called reporter ions. Currently, two commercially available reagents exist: isobaric tags for relative and absolute quantification (iTRAQ) ⁶ and tandem mass tag (TMT) ^{7,8}. The multiplexing capacity of these reagents varies: four experimental conditions in the case of iTRAQ 4-plex, six in the TMT 6-plex and eight in the iTRAQ 8-plex. Moreover, recent advances in the resolution of mass spectrometers allow them to distinguish the miniscule (6 mDa) difference in $^{12}\text{C}/^{13}\text{C}$ and $^{14}\text{N}/^{15}\text{N}$ isotopic pairs ^{9,10}. Such fidelity made possible the synthesis of 10-plex reagents using the same TMT employed for 6-plex, although some precautions must be taken to avoid ion coalescence ¹¹. Most recently, a 28-plex labelling architecture that has the ability to generate multiple series of reporter ions was demonstrated¹². Such a degree of multiplexing allows one to combine several biological samples into a single experiment ¹³ and to increase time-course series ¹⁴ which before had to be analysed in separate experiments ¹⁵.

Despite all these benefits, isobaric labelling is affected by an issue that undermines its accuracy. Precursor ions selected for fragmentation are frequently co-isolated with other peptides. As a result, the reporter ion intensities are biased by the fragmentation of both peptides ^{16,17}. Since the vast majority of proteins remain unaltered, the ratios of differentially expressed proteins are often determined in the presence of this constant background, shifting the measured ratios towards unity; this issue is popularly known as “ratio compression” ¹⁶. The consequences of ratio compression are overwhelming as actual fold changes of ten can be experimentally measured as 3.5 ¹⁸. Understanding and minimizing this effect is crucial when trying to extract the biological significance in a quantitative analysis. For instance, Ow

*et al.*¹⁶ showed that reducing sample complexity by pre-fractionation had a positive effect on ratio compression¹⁹. However, it was also shown that pre-fractionation allows the detection of novel peptide species²⁰ and Ting *et al.*¹⁸ provided further evidence that this only achieves a modest improvement in reducing interference¹⁸. Optimizing instrument settings by delaying fragmentation to the apex of the chromatographic peak² and narrowing the isolation width values can also result in a significant improvement in hybrid linear ion trap-Orbitrap devices^{2,18}. Recently, two groups proposed novel MS acquisition methods that solve this issue. On the one hand, Coon and co-workers carried out proton transfer reactions to reduce the charge state of co-isolated ions and reduce interference²¹. On the other hand, Gygi and colleagues, used MS3 fragmentation events to produce quantitative mass spectra free of background reporter ions¹⁸. In addition to these experimental strategies, Bantscheff and co-workers devised a post-acquisition algorithm that deconvolutes the degree of interference on the basis of the preceding survey spectra²². Albeit all these strategies are highly valuable, they only partially solve the problem or are restricted to modern and sophisticated instruments that have not been yet implemented in routine proteomics labs.

The main concern in isobaric labelling is that the underestimation of the real ratio might affect the statistical sensitivity to detect such changes as significant. To what extent this underestimation in accuracy hampers the statistical significance has been only considered in a few reports²³. In fact, to select differentially expressed proteins, it is a common practice to apply *ad hoc* thresholds to expression changes which are often as low as 1.2-fold. Also, these arbitrary criteria neglect the quality of the protein ratio, which often is calculated on the basis of several peptides ratios with a given variance. As a consequence, novel statistical models and computational approaches have been developed^{24–27}. The simplest method is the use of a one sample Student's t test taking as a protein ratio the average or median of all its peptides ratios, and assuming a normal distribution. Also, it has been described that the moderated t-test can be applied to account for sample variation²⁸. On the other hand, previous research has shown that not all peptides are evenly susceptible to identification and quantification and that the precision of the measurements decreases with low intensity signals²⁹. To correct for this variability, Hundermarkt *et al* developed a noise model²⁹ which is used by Isobar²⁷, an R package specifically

designed for isobaric labelling. Isobar also accounts for the biological and sample (*i.e.* technical) variability using a Cauchy distribution. Other methods, such as the one described in Navarro *et al.* ²⁶ propose a WSPP (weighted spectrum peptide protein) model in which the variability is estimated at each level (*e.g.* spectrum, peptide and protein) and afterwards integrated following the error propagation theory.

In addition to these new algorithms, the increasing multiplexing of new isobaric reagents discussed above enables the introduction of replicates, hence increasing the statistical probability of measuring significant changes despite the ratio compression. Here, we sought to test such hypothesis using a controlled experimental set-up that suffers from a strong isobaric interference. Our data describe the dependencies between precision and accuracy and the statistical significance of compressed ratios for three types of mass spectrometers and demonstrate the great benefit of including replicates in depicting truly differentially expressed proteins.

Experimental Procedures

Sample preparation

E.coli were lysed using 10 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.1 mM EDTA, 1:1000 (v/v) of benzonase and 1:100 (v/v) of Halt™ phosphatase and protease inhibitor cocktail. U2OS cells were lysed using 7 M urea, 2 M thiourea, 100 mM Hepes pH 7.5, 1:1000 (v/v) of benzonase and 1:100 (v/v) of Halt™ phosphatase and protease inhibitor cocktail. Both U2OS and *E.coli* were homogenized by sonication and cleared by centrifugation (20,000 g, 30 min, 4°C). 800 µg of *E.coli* lysate were divided into two aliquots of 400 µg to perform the digestion in parallel - referred as digestion A and digestion B - using the filter aided sample preparation (FASP) method³⁰. Samples were dissolved in 8M urea and 0.1M TEAB (UTEAB). Proteins were reduced in 10 mM DTT during 1h, at 40°C, and alkylated in 50 mM of IAA for 20 min in dark. First digestion using Lys-C (1:50 w/w, Wako) was performed during 4 hours at room temperature, followed by an 8-fold dilution in 50 mM TEAB. Second digestion using trypsin (1:100 w/w, Promega) was carried out overnight at 37°C. U2OS extract was in-solution digested using similar conditions to *E. coli*. All digestions were quenched with 1% TFA. For labelling, 96 µg of *E.coli* digested peptides were used for each channel using the iTRAQ® Reagent 8plex kit (AB Sciex). Digestion A was used for channels 113, 115, 117 and 119, whilst digestion B was used for 114, 116, 118 and 121. On the other hand, 100 µg of U2OS were labelled with the first four tags. Then, *E.coli* peptides were combined in 2.5:1 ratios, whereas U2OS peptides were mixed in 1:1 ratios as shown in Figure 1. Finally, 120 µg of *E.coli* were mixed with 150 µg of U2OS in order to obtain a ratio between species of 0.8 to 1. The clean-up was performed with Supelclean™ LC-SCX SPE Tubes 1 ml (SUPELCO®). Sample was loaded in 1% TFA, washed with 0.2% TFA and finally eluted with 450 µl of 5% (v/v) NH₃OH and 80% (v/v) of acetonitrile. Eluate was dried in vacuum and resuspended in 0.1% formic acid for subsequent analysis by LC-MS/MS.

LC-MS/MS

All samples were analyzed by LC-MS/MS on three different instruments. The Impact (Bruker Daltonics) was coupled online to a nanoLC Ultra system (Eksigent), equipped with a CaptiveSpray nanoelectrospray ion source supplemented with a CaptiveSpray nanoBooster operated at 0.2 bar/minute with isopropanol as dopant. Samples were loaded onto a reversed-phase C18, 5 μ m, 0.1 x 20 mm trapping column (NanoSeparations) and washed for 15 min at 2.5 μ l/min with 0.1% FA. The peptides were eluted at a flow rate of 300 nl/min onto a home-made analytical column packed with ReproSil-Pur C18-AQ beads 3 μ m, 500 x 0.075 mm, heated to 45 °C. Solvent A was 4% ACN in 0.1% FA and Solvent B acetonitrile in 0.1% FA. The following gradient was used: 0–2 min 2% B, 2–119 min 2–20% B, 119–129 min 20–34% B, 129–140 min 98% B, 140–145 min 2% B. The MS acquisition time used for each sample was 145 min. The Q-q-TOF Impact was operated in a data dependent mode. The spray voltage was set to 1.35 kV and the temperature of the source was set to 180°C. The MS survey scan was performed at a spectra rate of 2.5 Hz scanning between 80 and 1600 m/z. The minimum MS signal for triggering MS/MS was set to 500 counts. The 30 most abundant isotope patterns with charge ≥ 2 and m/z > 350 from the survey scan were sequentially isolated and fragmented by collision induced dissociation (CID) using a collision energy of 23 – 56 eV as function of the m/z value. The m/z values triggering MS/MS with a repeat count of 1 were put on an exclusion list for 30 s using the *rethinking* option: the precursor intensities were re-evaluated in the scan (n) regarding their values in the previous scan (n-1). Any m/z with intensity exceeding 5 times the measured value in the preceding survey scan was reconsidered for MS/MS. Data acquired were transformed to MGF format using the Compass DataAnalysis program. The LTQ Orbitrap Velos mass spectrometer (Thermo Scientific) was coupled to an Eksigent nano LC system (Eksigent) through a nanoelectrospray ion source (Proxeon Biosystems). Peptides were loaded from a cooled nanoLC AS-2 autosampler. 5 μ L from each sample were loaded onto a ReproSil Pur C18-Aq 5 μ m 0.3 x 10 mm trapping cartridge (SGE Analytical), and washed for 10 min at 2.5 μ L/min with loading buffer (0.1% FA). The peptides were eluted from a RP ReproSil Pur C18-AQ 1.9 μ m, 400 x 0.075 mm home-made column by a binary gradient consisting of 4% ACN in 0.1% FA (buffer A) and 100% ACN in 0.1%FA (buffer B), with a flow rate of 250 nL/min. Peptides

were separated using the following gradient: 0-2 min 4% B, 2-133 min 30% B and 133-143 min 98% B. The column was operated at a constant temperature of 50°C. Full scan MS spectra were acquired in the Orbitrap (scan range 350-1500 m/z, resolution 30,000 at m/z 400, AGC target 1e6, maximum injection time 500 ms). After the MS scans, the 15 most intense peaks were selected for HCD fragmentation at 38 % of normalised collision energy. HCD spectra were also acquired in the Orbitrap (resolution 7500, AGC target 3e4, maximum injection time 250 ms). The m/z values triggering MS/MS with a repeat count of 1 were put on an exclusion list for 70 s. The minimum MS signal for triggering MS/MS was set to 5000 counts. In all cases, one microscan was recorded, the Q value was set to 0.25 and an activation time to 0.10 ms was used. Waveform filter was also activated. The Q Exactive (Thermo) was connected to an EASY-nLC 1000 UHPLC system (Proxeon) through an EASY-Spray nano-electrospray ion source (Thermo). The peptides were trapped on an in-house packed trap column (ReproSil C18, 3µm, 20 x 0.075 mm) using solvent A (0.1% Formic Acid, 5 % DMSO in water) at a pressure of 500 bar. The peptides were separated on an EASY-spray Acclaim PepMap® analytical column (RSLC C18 2µm, 500 x 0.075 mm) using a linear gradient (length: 120 minutes, 8 % to 28 % solvent B (0.1% formic acid, 5 % DMSO in acetonitrile), flow rate: 200 nL/min). The column was operated at a constant temperature of 45°C. Peptides were electrosprayed directly into the mass spectrometer operating in a data-dependent mode. Full scan MS spectra were acquired in the Orbitrap (scan range 350-1500 m/z, resolution 70,000 at m/z 200, AGC target 3e6, maximum injection time 100 ms). After the MS scans, the 20 most intense peaks were selected for HCD fragmentation at 32 % of normalised collision energy. HCD spectra were also acquired in the Orbitrap (resolution 17500, AGC target 5e4, maximum injection time 120 ms). Different isolation windows were used. For the LTQ Orbitrap Velos, 0.5, 1, 1.5 and 2 Da; and for the Impact and Q Exactive, 0.7, 1, 1.5, 2 and 3 Da. For each analysis, 1 µg of sample was injected and two technical replicates (runs) were performed.

Data Processing

All files were analysed using Proteome Discoverer 1.4 (Thermo Scientific) with Sequest HT as the search engine against a concatenated Uniprot database of *Homo sapiens* (20,187 sequences), *Escherichia coli* (4,305 sequences) and supplemented with frequently observed contaminants (397

sequences). iTRAQ 8plex tag in lysine and N-terminal were included as fixed modifications, together with carbamidomethylation of cysteine. Oxidation of methionine and iTRAQ 8plex labelling of tyrosine were included as variable modifications. Precursor mass tolerance was 20 ppm for all instruments and fragment mass tolerance was 0.05 Da for Impact and 0.025 for LTQ Orbitrap Velos and Q Exactive. The integration of reporter ions was performed using the most confident centroid with a tolerance of 50 ppm. Reagents impurities were corrected as indicated by the manufacturer. PSMs were filtered using Percolator (v2.04) ³¹ with a FDR of 1%. Quantification results at the PSM level were exported for further analysis.

Statistical analysis

Quantification and statistical analysis was performed using Isobar (version 1.10) ²⁷ in R (version 3.1.2 “Pumkin Helmet” ³²). To this end, we first used a noise model ²⁹ that accounts for the technical variation due to the instrument, using the ratios obtained from channels 113 and 114. These channels were selected because both correspond to ratios 1 to 1 in both human and *E.coli* peptides. Moreover, the intensities measured for these channels were the highest allowing to model the variance for all the dynamic range. Second, a null protein distribution was used to model sample variability (*i.e.* 113 and 114 channels obtained from independent digestions). Afterwards, protein ratios were calculated for either channels with and without co-isolation. Ratios were obtained as a value for each “one vs. one” comparison between conditions (*e.g.* 113/115). Moreover, since the experimental design includes two technical replicates of each condition, ratios can also be calculated as “two vs. two” comparisons. This is calculated by Isobar as the weighted average of the four available combinations between replicates (*e.g.* 113/115, 114/116, 113/116 and 114/115). Besides we calculated again the same ratios but discarding those proteins quantified by only one PSM. Finally, we calculated the statistical sensitivity of all the data sets as the true positive rate (TPR). Here, TPR is defined as the number of *E. coli* proteins detected as significant divided by the total number of *E. coli* proteins quantified.

Results

The mass analyser used for precursor isolation influences the degree of ratio compression

To evaluate the effect of ratio compression on the statistical analysis we used the “two proteomes approach” previously described for TMT 6-plex¹⁸ and adapted it to iTRAQ 8-plex (see Materials and Methods and Figure 1A). This model imitates the complexity of a typical biological comparison, in which the human U2OS cell line mimics the large set of non-regulated proteins, whilst the *E. coli* proteome represents the set of regulated proteins, which in this case were experimentally defined with a fold-change of 2.5. This value was chosen because it was the smallest ratio examined by Ting *et al*¹⁸ and, under identical experimental conditions (*e.g.* the number of quantified peptides per protein and the quality of such measurements), it is closer to the null distribution and hence is less likely to be detected as significant as if we had chosen a larger ratio. Also, the results from our current work might be easily comparable to those obtained using the MS3 method¹⁸.

The degree of ratio compression is determined by the contribution of human peptide species co-isolated during selection of *E. coli* precursor ions before MS/MS fragmentation¹⁹. The sensitivity and the selectivity of precursor isolation might be different between mass analysers of different nature³. Here, we tested this hypothesis in the LTQ Orbitrap Velos (LIT-OT-q), the Impact (Q-q-TOF) and the Q Exactive (Q-OT-q) mass spectrometers (Figure 1B). While the first one uses a linear ion trap for ion isolation, the other two instruments are equipped with resolving quadrupoles. All the three platforms were operated at an isolation width value of 2 Da which is commonly used in data dependent acquisition of complex proteomic samples.

The extent of ratio compression in our data sets was evaluated at the “PSMs level”, using each quantification event as a unique and independent measurement. To this end, we calculated the $\log_2$ ratios for all the *E.coli* PSMs that were affected by co-isolation with the human background (*i.e.* 113/115, hereafter referred as “compressed”) and for those without isobaric interference (*i.e.* 117/119, hereafter referred as “uncompressed”) (Figure 1A). In the absence of any sort of interference the theoretical $\log_2$

ratio for the *E. coli* peptides should be 1.32. However, one should also consider uncontrollable deviations from this value due to small pipetting errors when samples were mixed. For the LTQ Orbitrap Velos, the median of the compressed ratios was 0.671 while for the uncompressed ratios was 1.392 (Figure 2, Supplementary Figure 2A and Supplementary Table S1). Similar medians were observed for the other pairwise comparisons in compressed and uncompressed channels (data not shown). Thus, the degree of ratio compression, calculated as the percentage of compressed/uncompressed ratios, was found to be 51.8% for this instrument (Figure 2). In the case of the Q-q-TOF Impact, the compressed ratios (median = 0.563) and uncompressed (median = 1.296) showed a degree of compression of 56.5% (Figure 2, Supplementary Figure 1B and Supplementary Table S2) while the compression for the Q Exactive was the highest and ratios were compressed by 65.4% (Figure 2, Supplementary Figure 1C and Supplementary Table S3). The differences observed among these instruments prompted us to investigate in more detail the relationship between selectivity and ratio compression.

Our comparison among platforms using the same isolation width value might be inaccurate, as the response factor might be different depending on the linearity and calibration curves of this function by each manufacturer. Previously, Savitski *et al.* demonstrated that narrowing the isolation width values on the LTQ Orbitrap Velos improves accuracy in isobaric labelling², and this result was later confirmed by Ting *et al.*¹⁸. Therefore, we sought to investigate the extent of this finding also on the Impact and Q Exactive. To this end, we repeated our measurements at a number of isolation width values for the three mass spectrometers. For the LTQ Orbitrap Velos, we examined ratio compression at isolation width values of 2.0, 1.5, 1 and 0.5 Da, whilst for the Impact and Q Exactive we used 3.0, 2.0, 1.5, 1.0 and 0.7 Da (Figure 3). The response of the LTQ Orbitrap Velos to this parameter was confirmed to be remarkable as the ratios were compressed only by 29.6% at the narrowest isolation width value tested (0.5 Da) (Figure 3A). In the case of the Impact, the difference between the widest (3 Da) and narrowest (0.7 Da) window resulted in a moderate improvement in accuracy, as the compression changed from 60.5% to 52.5% (Figure 3B). In the Q Exactive, the compression of ratios slightly improved from 69.8%-56.2% (Figure 3C). Interestingly, the compression observed in the Velos at the widest window (51.8%) outperformed all the values examined for the other two platforms (Figure 3). Also, the 0.7 Da

window set in the Q Exactive achieved the same degree of compression than the Impact working at 2.0 Da (Figure 3). Taken together, these results suggest that the degree of isobaric interference is in part determined by the mass analyser used, or more specifically the quality of the isolation applied for precursor isolation and may indicate that ion trap device from the LTQ Orbitrap Velos enable higher accuracy in isobaric labelling when compared to two quadrupole devices tested here.

The relationship between precision and accuracy is platform-dependent

The improvement in selectivity of using narrow isolation width values comes at the expense of a decrease in ion transmission, *i.e.* sensitivity. It has been shown by others that low reporter ion intensities result in a loss of precision²⁹. Hence, we analysed the relationships between isolation window, precision and accuracy in our data sets. To this end, the data was normalized in each platform as the percentage of the highest intensity measured. As expected, the widest window resulted in the highest reporter intensities and the signal decreased as the window was narrowed for the three instruments (Figure 4A and Supplementary Figure 1). The Impact exhibited a solid linear response, but the intensity dropped more than 80% between 0.7 and 3.0 Da (Figure 4A). The LTQ Orbitrap Velos also showed a good linearity and a 40% loss in signal was observed. (Figure 4A). The intensity in the Q Exactive decreased approximately 50%, but the linear response was only observed between 2-1 Da (Figure 4A). In this instrument, the intensities did not vary significantly between 3-2 Da and 1-0.7 Da (Figure 4A). Then we calculated the precision (*i.e.* standard deviation of all the PSMs ratios) for all the conditions evaluated. Note that precision is dimensionless and thus the comparison between data sets is straightforward here. The most precise measurements (*i.e.* lowest STDEV) were observed at the highest reporter intensities while the lowest intensities resulted in the most dispersed ratios (*i.e.* highest STDEV) (Figure 4B and Supplementary Figure 1). The Impact achieved the best precision (0.24) closely followed by the Q Exactive (0.26) (Figure 4B). In both instruments, the precision decreased almost linearly in all the range examined with the exception of the Impact at isolation windows smaller than 1.5 Da, in which precision was significantly worse (Figure 4B). Contrary, the precision in the LTQ Orbitrap Velos did not decrease as much as in the other two platforms. The best precision observed for this instrument at 2.0 Da (0.31) was achieved by the other instruments at approximately 1-1.5 Da (Figure

4B). Next, we investigated the dependencies between precision and accuracy in our data sets. In all the instruments, the increase in selectivity resulted in a gain of accuracy (Figure 4C and Supplementary Figure 1C). Inversely, this variable had a negative effect in the sensitivity, and consequently the precision of the measurements decreased (Figure 4C). Taken together, these observations reveal important differences in the relationships between precision and accuracy of isobaric quantification amongst these instruments in the experimental conditions tested here (Figure 4C). On the one hand, the LTQ Orbitrap Velos achieves the best accuracy levels with a moderate precision that seems to be robust to changes in signal intensity (Figure 4C). On the other hand, the precision levels of both Q Exactive and Impact are similar and surpass those seen for the Velos but the Impact achieves higher accuracy than the Q Exactive albeit with a significant loss of precision at narrow isolation windows (Figure 4C).

The impact of precision and accuracy on the statistical significance of differential proteins

A primary goal in a quantitative proteomic analysis is to detect all the differentially expressed proteins present in a biological assay. The statistical significance mainly depends on: (i) how different the means of the two samples are and (ii) the reproducibility of the measurements. While the first one refers to the accuracy, the second is for the precision. Above, we have just described that the improvement in accuracy results in a detrimental loss of precision, wherein noticeable differences are apparent between instruments. Next, we investigated the influence of each of these variables in the statistical analysis of our data sets using Isobar²⁷ (Figure 1C). Given the differences in the absolute number of identified proteins in each of the instruments, the True Positive Rate (TPR) was calculated (*i.e.* percentage of significant *E. coli* proteins from all quantified *E. coli* proteins). First, we analysed the data using a “one vs. one” sample comparison (*i.e.* 113 vs. 115 and 117 vs 119 for the compressed and uncompressed channels respectively). Surprisingly, in the Q Exactive and Impact, and despite the lack of isobaric interference, it was not possible to identify all the “uncompressed” *E. coli* proteins as statistically significant (Figure 5A). In the LTQ Orbitrap Velos, 96-99% of the proteins were detected as significant (Figure 5A). With regard to the compressed ratios, despite the loss in identifications observed in the

Velos, the absolute number of statistically significant proteins that suffer from isobaric interference remained very similar (123-136 proteins) (Figure 5A). Interestingly, the other two instruments showed opposite behaviours. On the one hand, in the Q-q-TOF Impact the percentage dropped from 40% to 20% (Figure 5A). On the other hand, in the Q Exactive this percentage increased from 21% to 37% with the highest sensitivity at 1.5 Da (44%). Together, these results show important differences across these platforms in the ability to detect proteins as significant when ratios of 2.5 are measured in the presence of isobaric interference.

The inclusion of replicates improves statistical significance and counteracts ratio compression

In a quantitative proteomics experiment, the inclusion of replicates increases the probability of discriminating biologically significant changes from the stochastic variability of the measurements³³. Here, we sought to investigate the benefit of including a second replicate in the statistical analysis. All “two vs. two” comparisons were performed by Isobar for compressed (113 and 114 vs. 115 and 116) and uncompressed channels (117 and 118 vs. 119 and 121) and corresponding p-values were estimated. Remarkably, this “summarization” analysis led to an overall increase in the percentage of significant proteins, although this effect was more prominent in the compressed channels (Figure 5B). For the Velos, the inclusion of a second replicate allowed the identification of nearly the same number of significant proteins in the compressed channels as in the uncompressed (Figure 5B), despite their large difference in accuracy (Figure 3A). This improvement was also evident in the Impact and Q Exactive (Figure 5B), in which the percentage of significant *E. coli* proteins that suffer from isobaric interference double and even triple when compared to the “one vs. one” strategy (Figure 5A). In the Impact, 86% of *E. coli* proteins were found significant despite their low accuracy (2.0 Da), while the Q Exactive enabled the identification of 79% (1.0 Da).

A common practice in quantitative proteomics is to filter out proteins which are quantified on the basis of just one peptide, since no variance can be estimated for those protein ratios. To evaluate the effect of discarding these “*a priori*” low quality quantified proteins, we removed from our data sets the “single-

hit” proteins and calculated again the significance levels. In all the conditions the effect of isobaric interference was nearly abolished. The number of *E. coli* proteins that were found significant was nearly the same between proteins ratios that suffered from compression and those that did not (Figure 5C). For instance, in the absence of any interference, 279 significant proteins (100% of all quantified) were found in the Impact (2 Da), almost the same number as in the compressed ratios (268 proteins) (Figure 5C). Similar results were found in the Q Exactive and in the LTQ Orbitrap Velos (Figure 5C). Noticeably, by applying this filter, the total number of quantified proteins decreased in all the instruments by 20-30%. However, the number of *E. coli* proteins that were found statistically significant over the total number of quantified *E. coli* were maximized (*i.e.* true positive rate, sensitivity or recall). This might be important for some cases in which the false negative rate (*i.e.* percentage of differential proteins that are not detected as significant) needs to be minimized.

In the case of PTMs, the vast majority of peptides are frequently quantified on the basis of just one measurement³⁴. Thus, if no variance can be estimated, the statistical outcome might be penalised. In this exercise, the Isobar statistical analysis was performed at the peptide level using the “two vs. two” comparison. The percentages of significant peptides from the compressed channels (Figure 5D) were barely lower than those obtained at the protein level (Figure 5B). In the LTQ Orbitrap Velos, 77-94% of the compressed peptides were detected as significant (Figure 5D) whilst we were able to detect 88-99% of the proteins (Figure 5B). Similar results were obtained in the other two instruments. Interestingly, the distribution of peptide ratios showed higher accuracy than those calculated at the protein level (Supplementary Figure S2) in agreement with other reports³⁵. A possible explanation could be that a great proportion of the protein ratios is calculated on the basis of just a few peptides. Often these peptides are low abundant and consequently the reporter ions are measured with poor S/N levels which is known to correlate with isobaric interference²².

Discussion

Ratio compression remains a concern in isobaric tagging. Although some experimental strategies can solve this problem^{18,36}, so far they have only been implemented in hybrid ion trap mass spectrometers. Therefore, it is crucial to understand this issue^{1,37}, minimize the degree of compression during acquisition¹⁹ and optimize the data analysis²². All these studies showed a partial or even complete improvement in accuracy but overlooked the impact that isobaric interference might have on the statistical significance of the data. This is especially relevant considering the recent development of statistical approaches specifically designed for isobaric labelling^{26,27,35}. Here, by manipulating isolation width values, we were able to model data sets with different levels of accuracy and precision. In the experimental set up used here, we determined the influence of these variables on the statistical significance of ratios distorted by the interference and compared them to ratios measured without such impurities (Figure 6). Most importantly, our analyses were systematically conducted on three instruments and showed each had distinct profiles.

In the Impact, the improvement of statistical sensitivity might be explained by a gain in precision (Figure 6A) given that accuracy remained unchanged (Figure 6B). On the other side, accuracy prevails over precision in the Velos and Q Exactive since the latter shows anti-correlation with the statistical significance (Figure 6A). The modest improvement in sensitivity in the Q Exactive is likely due to an increase in accuracy (Figure 6B). Although the Velos achieved the best sensitivity and the most accurate ratios (Figure 6B), the absolute number of differential proteins decreased remarkably (Figure 5A). It would be of interest to compensate the loss of ion transmission resulting of small isolation windows by increasing the injections time³⁸. Certainly, this will have a negative impact in the duty cycle but the quality of the quantitative data might likely improve, and hence, the statistical sensitivity. Our data also showed differences between the two quadrupole-based instruments. Although they enable similar levels of accuracy, the Impact suffers from a significant loss in precision at small isolation windows which the Q Exactive seems to better tolerate (Figure 4C). Both proton-transfer reactions²¹ and MS3¹⁸ approaches mentioned above have a significant loss in reporter ion intensities (also reported in³⁷).

Recently, the implementation of isolation waveforms with multiple frequency notches (*i.e.* multi-notch MS3)³⁹ showed a 10-fold increase in reporter signal compared to the original method. Nevertheless, the compensation in ion statistics will definitely affect the precision of the measurements and consequently the statistical significance, which should be investigated in detail⁴.

The “two proteomes model”¹⁸ used here mimics, to a certain extent, a real biological scenario in which a set of differential proteins co-occur with a much larger proportion of non-changing proteins. Nevertheless, this design presents some caveats that need consideration. In this set up, it is impossible to imitate the biological variability that exists in a real study. Our two replicates correspond to independent *E. coli* digests and hence they only reflect the error that is introduced during sample preparation. However, many other factors, such as genetic background or tissue heterogeneity impose a higher order of variability that cannot be implemented in our model. Therefore, the sensitivity levels found in our data sets will likely be an underestimation of the real levels in a case study. Here, we have evaluated the statistical significance of isobaric data using a defined 2.5 fold-change. At this ratio, we have shown that the inclusion of two replicates enables the identification of most of the differential proteins as significantly regulated (Figure 5B). It would be interesting to determine whether the inclusion of a third replicate could further increase the sensitivity of the analysis and reach the same levels found in the absence of ratio distortion (Figure 5B). Likewise, the statistical sensitivity might increase for differential proteins with ratios larger than 2.5, but the expected loss in precision might have an impact on the outcome. On the other hand, for ratios smaller than 2.5, a prior analysis of the statistical power³³ could be done. This will be useful to determine the appropriate number of replicates.

When multiple conditions need to be tested, the researcher might replicate the assay in a separate experiment⁴⁰. In such design, the p-values of each experiment can be combined through Stouffer’s Z-transformation²⁴. We have analysed our data sets using this approach and the results were very similar to those obtained by the summarization of two channels in Isobar (data not shown). However, this strategy is limited to those proteins quantified in the two runs. In principle, the controlled experimental conditions used here should allow the estimation of not only the statistical sensitivity (number of significant *E. coli* proteins out of all *E. coli* proteins) but also the specificity (number of significant

human proteins from all human proteins) (*i.e.* true negatives). However, there is an experimental bias in human peptides when they are measured in the presence of *E. coli* peptides. This causes a shift in their ratios towards positive values, therefore increasing the likelihood of being identified as significant (data not shown). Consequently, the number of significant human proteins is rather inflated and does not reflect the false positives of the experiment and should not be used to determine specificity levels.

Conclusions

This study provides a systematic evaluation of the performance of three popular MS platforms – LTQ Orbitrap Velos, Q Exactive and Impact – regarding the quantification of proteins differentially expressed by 2.5 fold which suffer from isobaric interference. Our results point out the importance of understanding how ratio compression is affected in each instrument, and the necessity to optimize acquisition parameters to minimize as much as possible isobaric interference without compromising the identification rate and the precision of the measurements. We also show that these parameters have a different impact on the statistical outcome and that the inclusion of replicates compensates to a large extent the loss of accuracy. The combination of a design that resembles the complexity of a real biological assay, and its systematic analysis in three instruments at different levels of accuracy and precision represents a useful resource for the proteomics community. We expect that these data sets might serve as a platform for future studies of isobaric labelling and the development of new statistical tools. The proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE ⁴¹ partner repository with the identifier PXD003640.

Acknowledgments

The CNIO Proteomics Unit belongs to ProteoRed, PRB2-ISCI, supported by grant PT13/0001. Part of this work was funded by SAF2013-45504-R (MINECO). J.M. is also supported by Ramon y Cajal

Programme (MINECO) RYC-2012-10651, A.M.V. is supported by BES-2014-070098 (MINECO) and E.Z. by PEJ-2014-P-01082 (MINECO).

Associated Content

Supplementary Table S1. Quantification results for *E.coli* PSMs at different isolation widths and for the two technical replicates analysed in the LTQ Orbitrap Velos.

Supplementary Table S2. Quantification results for *E.coli* PSMs at different isolation widths and for the two technical replicates analysed in the Impact.

Supplementary Table S3. Quantification results for *E.coli* PSMs at different isolation widths and for the two technical replicates analysed in the QExactive.

Supplementary Figure S1. Distribution of ratios of *E. coli* peptides versus reporters' intensities.

Supplementary Figure S2. Distribution of ratios measured at PSM and protein level.

References

- (1) 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.
- (2) 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.
- (3) 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.
- (4) Williamson, J. C.; Edwards, A. V. G.; Verano-Braga, T.; Schwämmle, V.; Kjeldsen, F.; Jensen, O. N.; Larsen, M. R. High-performance hybrid Orbitrap mass spectrometers for quantitative proteome analysis: observations and implications. *Proteomics* **2016**, *16* (6), 907–914.
- (5) Altelaar, A. F. M.; Munoz, J.; Heck, A. J. R. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nat. Rev. Genet.* **2013**, *14* (1), 35–48.
- (6) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3* (12), 1154–1169.
- (7) 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.

- (8) Thompson, A.; Schäfer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K. A.; 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.
- (9) 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.
- (10) McAlister, G. C.; Huttlin, E. L.; Haas, W.; Ting, L.; Jedrychowski, M. P.; Rogers, J. C.; Kuhn, K.; Pike, I.; Grothe, R. a.; Blethrow, J. D.; et al. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* **2012**, *84* (17), 7469–7478.
- (11) Werner, T.; Sweetman, G.; Savitski, M. F.; Mathieson, T.; Bantscheff, M.; Savitski, M. M. Ion coalescence of neutron encoded TMT 10-plex reporter ions. *Anal. Chem.* **2014**, *86* (7), 3594–3601.
- (12) Braun, C. R.; Bird, G. H.; Wühr, M.; Erickson, B. K.; Rad, R.; Walensky, L. D.; Gygi, S. P.; Haas, W. Generation of multiple reporter ions from a single isobaric reagent increases multiplexing capacity for quantitative proteomics. *Anal. Chem.* **2015**, *87* (19), 9855–9863.
- (13) Isasa, M.; Rose, C. M.; Elsasser, S.; Navarrete-Perea, J.; Paulo, J. A.; Finley, D. J.; Gygi, S. P. Multiplexed, Proteome-wide Protein Expression Profiling: Yeast Deubiquitylating Enzyme Knockout Strains. *J. Proteome Res.* **2015**, *14* (12), 5306–5317.
- (14) Murphy, J. P.; Stepanova, E.; Everley, R. A.; Paulo, J. A.; Gygi, S. P. Comprehensive Temporal Protein Dynamics during the Diauxic Shift in *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* **2015**, *14* (9), 2454–2465.
- (15) Benevento, M.; Tonge, P. D.; Puri, M. C.; Hussein, S. M. I.; Cloonan, N.; Wood, D. L.; Grimmond, S. M.; Nagy, A.; Munoz, J.; Heck, A. J. R. Proteome adaptation in cell reprogramming proceeds via distinct transcriptional networks. *Nat. Commun.* **2014**, *5*,

- 5613.
- (16) 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.
  - (17) Bantscheff, M.; Boesche, M.; Eberhard, D.; Matthieson, T.; Sweetman, G.; Kuster, B. Robust and sensitive iTRAQ quantification on an LTQ Orbitrap mass spectrometer. *Mol. Cell. Proteomics* **2008**, *7* (9), 1702–1713.
  - (18) Ting, L.; Rad, R.; Gygi, S. P.; Haas, W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nature Methods*. 2011, pp 937–940.
  - (19) 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.
  - (20) Michalski, A.; Cox, J.; Mann, M. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *J. Proteome Res.* **2011**, *10*, 1785–1793.
  - (21) 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.
  - (22) Savitski, M. M.; Mathieson, T.; Zinn, N.; Sweetman, G.; Doce, C.; Becher, I.; Pahl, F.; Kuster, B.; Bantscheff, M. Measuring and managing ratio compression for accurate iTRAQ/TMT quantification. *J. Proteome Res.* **2013**, *12* (8), 3586–3598.
  - (23) Mertins, P.; Udeshi, N. D.; Clauser, K. R.; Mani, D.; Patel, J.; Ong, S. -e.; Jaffe, J. D.; Carr, S. a. iTRAQ Labeling is Superior to mTRAQ for Quantitative Global Proteomics and Phosphoproteomics. *Mol. Cell. Proteomics* **2012**, *11* (6), M111.014423–M111.014423.
  - (24) Pascovici, D.; Song, X.; Solomon, P. S.; Winterberg, B.; Mirzaei, M.; Goodchild, A.; Stanley,

- W. C.; Liu, J.; Molloy, M. P. Combining protein ratio p-values as a pragmatic approach to the analysis of multirun iTRAQ experiments. *J. Proteome Res.* **2015**, *14* (2), 738–746.
- (25) Zhou, C.; Walker, M. J.; Williamson, A. J. K.; Pierce, A.; Berzuini, C.; Dive, C.; Whetton, A. D. A hierarchical statistical modeling approach to analyze proteomic isobaric tag for relative and absolute quantitation data. *Bioinformatics* **2014**, *30* (4), 549–558.
- (26) Navarro, P.; Trevisan-Herraz, M.; Bonzon-Kulichenko, E.; Núñez, E.; Martínez-Acedo, P.; Pérez-Hernández, D.; Jorge, I.; Mesa, R.; Calvo, E.; Carrascal, M.; et al. General statistical framework for quantitative proteomics by stable isotope labeling. *J. Proteome Res.* **2014**, *13* (3), 1234–1247.
- (27) Breitwieser, F. P.; Müller, A.; Dayon, L.; Köcher, T.; Hainard, A.; Pichler, P.; Schmidt-Erfurth, U.; Superti-Furga, G.; Sanchez, J. C.; Mechtler, K.; et al. General statistical modeling of data from protein relative expression isobaric tags. *J. Proteome Res.* **2011**, *10*, 2758–2766.
- (28) Kammers, K.; Cole, R. N.; Tiengwe, C.; Ruczinski, I. Detecting Significant Changes in Protein Abundance. *EuPA open proteomics* **2015**, *7*, 11–19.
- (29) Hundertmark, C.; Fischer, R.; Reinl, T.; May, S.; Klawonn, F.; Jänsch, L. MS-specific noise model reveals the potential of iTRAQ in quantitative proteomics. *Bioinformatics* **2009**, *25* (8), 1004–1011.
- (30) Wiśniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, *6* (5), 359–362.
- (31) Käll, 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.
- (32) R Core Team. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing Vienna Austria*. 2014, pp {ISBN} 3-900051 – 07-0.

- (33) Levin, Y. The role of statistical power analysis in quantitative proteomics. *Proteomics* **2011**, *11* (12), 2565–2567.
- (34) Erickson, B. K.; Jedrychowski, M. P.; McAlister, G. C.; Everley, R. A.; Kunz, R.; Gygi, S. P. Evaluating multiplexed quantitative phosphopeptide analysis on a hybrid quadrupole mass filter/linear ion trap/orbitrap mass spectrometer. *Anal. Chem.* **2015**, *87* (2), 1241–1249.
- (35) Zhou, C.; Walker, M. J.; Williamson, A. J. K.; Pierce, A.; Berzuini, C.; Dive, C.; Whetton, A. D. A hierarchical statistical modeling approach to analyze proteomic isobaric tag for relative and absolute quantitation data. *Bioinformatics* **2014**, *30* (4), 549–558.
- (36) Vincent, C. E.; Rensvold, J. W.; Westphall, M. S.; Pagliarini, D. J.; Coon, J. J. Automated gas-phase purification for accurate, multiplexed quantification on a stand-alone ion-trap mass spectrometer. *Anal. Chem.* **2013**, *85* (4), 2079–2086.
- (37) Altelaar, a. F. M.; Frese, C. K.; Preisinger, C.; Hennrich, M. L.; Schram, A. W.; Timmers, H. T. M.; Heck, A. J. R.; Mohammed, S. Benchmarking stable isotope labeling based quantitative proteomics. *J. Proteomics* **2012**, 1–13.
- (38) Kelstrup, C. D.; Young, C.; Lavalley, R.; Nielsen, M. L.; Olsen, J. V. Optimized fast and sensitive acquisition methods for shotgun proteomics on a quadrupole orbitrap mass spectrometer. *J. Proteome Res.* **2012**, *11* (6), 3487–3497.
- (39) McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wühr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal. Chem.* **2014**, *86* (14), 7150–7158.
- (40) Isasa, M.; Rose, C. M.; Elsasser, S.; Navarrete-Perea, J.; Paulo, J. A.; Finley, D. J.; Gygi, S. P. Multiplexed, Proteome-wide Protein Expression Profiling: Yeast Deubiquitylating Enzyme Knockout Strains. *J. Proteome Res.* **2015**, *14* (12), 5306–5317.

- (41) Vizcaíno, J. A.; Csordas, A.; Del-Toro, N.; Dienes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **2016**, *44* (D1), D447–D456.

Figure Legends

Figure 1. Experimental design. (A) Two *E. coli* samples were labelled with iTRAQ 8plex and mixed in ratios 2.5 to 1 as shown in the figure. U2OS cell line was labelled with only the first four channels and mixed in ratios 1 to 1. Finally, both preparations were mixed together in a 0.8.1 to generate a sample with and without isobaric interference. (B) Samples were analysed in duplicate in the LTQ Orbitrap Velos (LIT-OT-q), Impact (Q-q-TOF) and Q Exactive (Q-OT-q). By using different isolation window, different sets of accuracy and precision were modelled. (C) Statistical significance of *E. coli* proteins and peptides was determined using Isobar R package.

Figure 2. Distribution of ratios of *E. coli* peptides in the presence and absence of isobaric interference. Boxplots represent the median and interquartile range of the log₂ ratios of *E.coli* PSMs for the channels with and without compression due to human background at an isolation width of 2.0 Da. Red dashed line represents the theoretical value in which the *E.coli* peptides were spiked in.

Figure 3. Differences in compression rate as a function of isolation width in three MS platforms. Bar plots show the median of the log₂ ratios of *E.coli* PSMs for the compressed (dark colour) and uncompressed (light colour) conditions at different isolation widths. Percentages represent the degree of compression in the distorted channels calculated using as a reference the channels free of interference. Error bars refer to the standard deviation between technical replicates. LTQ Orbitrap Velos data is represented in blue, Impact data in green and Q Exactive is shown in yellow.

Figure 4. Relationships between intensity, accuracy, precision and isolation width. (A) Median of reporter ions intensities of *E. coli* peptides affected by compression at different isolation width values. Shown are the percentages relative to the highest intensity detected. (B) Relation between precision and isolation width. Precision is calculated as the standard deviation (STDEV) of the log₂ ratios for the compressed conditions. Error bars show the standard deviation between technical replicates. (C) Relation between precision and accuracy of both compressed conditions (dark colours) and uncompressed conditions (light colours). LTQ Orbitrap Velos data is represented in blue, Impact data

in green and Q Exactive is shown in yellow. Red dashed line represents the theoretical value in which the *E.coli* peptides were spiked in.

Figure 5. Statistically significant *E. coli* proteins and peptides identified by Isobar. The darkest colour represents the number of significant proteins/peptides of the compressed ratios, intermediate colour refers to the uncompressed conditions, and the lightest shows the total number of quantified protein/peptides. LTQ Orbitrap Velos data is represented in blue, Impact data in green and Q Exactive is shown in yellow (A) Results obtained for one vs. one comparison at protein level (113/115 was used for compressed channels and 117/119 for uncompressed). (B) Results obtained from the two vs. two comparisons for the compressed and uncompressed conditions. (C) Same as in B but discarding those proteins quantified by only one PSM. (D) Results obtained at the peptide level using the two replicates for the compressed and uncompressed channels.

Figure 6. Correlation between the precision and accuracy of isobaric labelling and the statistical sensitivity. (A) Correlation between precision (measured as the standard deviation of the log2 ratios of *E. coli* PSMs) and statistical sensitivity (percentage of significant *E. coli* proteins out of all quantified *E. coli* proteins) in compressed channels (dark colours) and uncompressed (light colours). (B) Correlation between accuracy (median of log2 ratios) and statistical sensitivity. Points are connected from the widest to the narrowest isolation width value. LTQ Orbitrap Velos data is represented in blue, Impact data in green and Q Exactive is shown in yellow.