UC Riverside

Title
EVALUATING TMTC+ FOR QUANTITATIVE PROTEOMICS ON A LTQ ORBITRAP VELOS PRO

Permalink
https://escholarship.org/uc/item/6sh4w75d

Author
Deloney, Nicholas M

Publication Date
2022-05-01

Data Availability
The data associated with this publication are not available for this reason: N/A

EVALUATING TMTC+ FOR QUANTITATIVE PROTEOMICS ON A LTQ ORBITRAP

VELOS PRO

By

Nicholas Mikel Deloney

A capstone project submitted for Graduation with University Honors

May 1, 2022

University Honors
University of California, Riverside

Dr. Joseph Genereux
Department of Chemistry

Dr. Richard Cardullo, Howard H Hays Jr. Chair
University Honors

ABSTRACT

Mass spectrometry-based proteomics is an extremely useful tool for characterizing proteins. One method for quantifying proteins using mass spectrometry is tandem mass tagging (TMT), where peptides are chemically labeled so that a reporter ion is released during peptide fragmentation. TMT can be performed with a single round of fragmentation and identification (TMT-MS2), but suffers from limitations in quantitative accuracy due to interference between different peptides. Alternatively, TMTc+ quantifies TMT-labeled proteins using complement reporter ions (Sonnett, Yeung, and Wühr 2018), offering higher reported accuracy. However, TMTc+ has only been performed on higher resolution quadrupole-Orbitraps. Our goal is to determine whether TMTc+ is effective on the lower resolution hybrid Orbitrap mass spectrometers. If not, we would like to optimize how TMTc+ can be used on these instruments to benchmark the reliability of TMT-MS2 data for specific classes of sample. To evaluate the effectiveness of TMTc+, we ran TMTc+ on new samples. Samples consisted of human cell lysate in a 1 to 4 to 10 to 4 to 1 ratio across five TMT channels, added into labeled *E. coli* lysates to increase the density of potentially interfering peptides. When comparing the ratios of TMTc+ to that of the reporter ions, we see that at high resolution TMTc+ has similar quantitative accuracy to TMT-MS2, but at lower resolutions TMTc+ does not perform as well as TMT-MS2. Additionally, we evaluate the basis for poor TMTc+ performance on an LTQ Orbitrap Velos Pro.

ACKNOWLEDGEMENTS

INTRODUCTION

Quantitative proteomics is an important tool in biological, health, and cancer research. Quantitative proteomics is an analytical tool used in chemistry to determine the amount of proteins in a given sample. In the case of quantitative proteomics, a sample is typically some amount of substance from a biological organism. Not only does quantitative proteomics find the relative amount of proteins in a sample, it also identifies proteins and other physiological aspects of a sample. This is useful because it can be helpful in determining the differences between a given sample, such as cancer cells and human cells.

One of the main methods for quantitative proteomics is mass spectrometry (MS). MS is an analytical technique in chemistry which measures the mass to charge ratio of ions in the gas phase. The mass to charge ratio of ions specifically relates the mass and the electric charge of a given particle. In MS, the mass to charge ratio is represented as m/z, where m refers to mass and z refers to the charge of the ion. MS measures the m/z of a given ion in a sample and outputs the intensity of the ions versus m/z, also known as the mass spectrum. TMT-MS2 is a method of MS where molecules of a sample are ionized and separated by m/z, individually split into smaller fragment ions, and the fragments are separated and identified by m/z. A mass spectrum shows this fragmentation at each m/z for each ion.

In order to better understand physiological differences between samples, identifying as many peptides in a sample as possible is useful. In MS, resolving power (RP) is the ability for an instrument to distinguish between two different peaks of similar m/z at the full width half maximum. High RP is useful for peptide identification as it can provide more accurate and precise measurements in peptide identification. Moreover, high RP is useful for resolving ion distributions. However, as RP increases the number of peptides identifiable go down as the

instrument takes longer to analyze each ion. We can see this relationship represented in Figure 1. The x-axis represents the RP of the run and the y-axis represents the number of peptides identified.
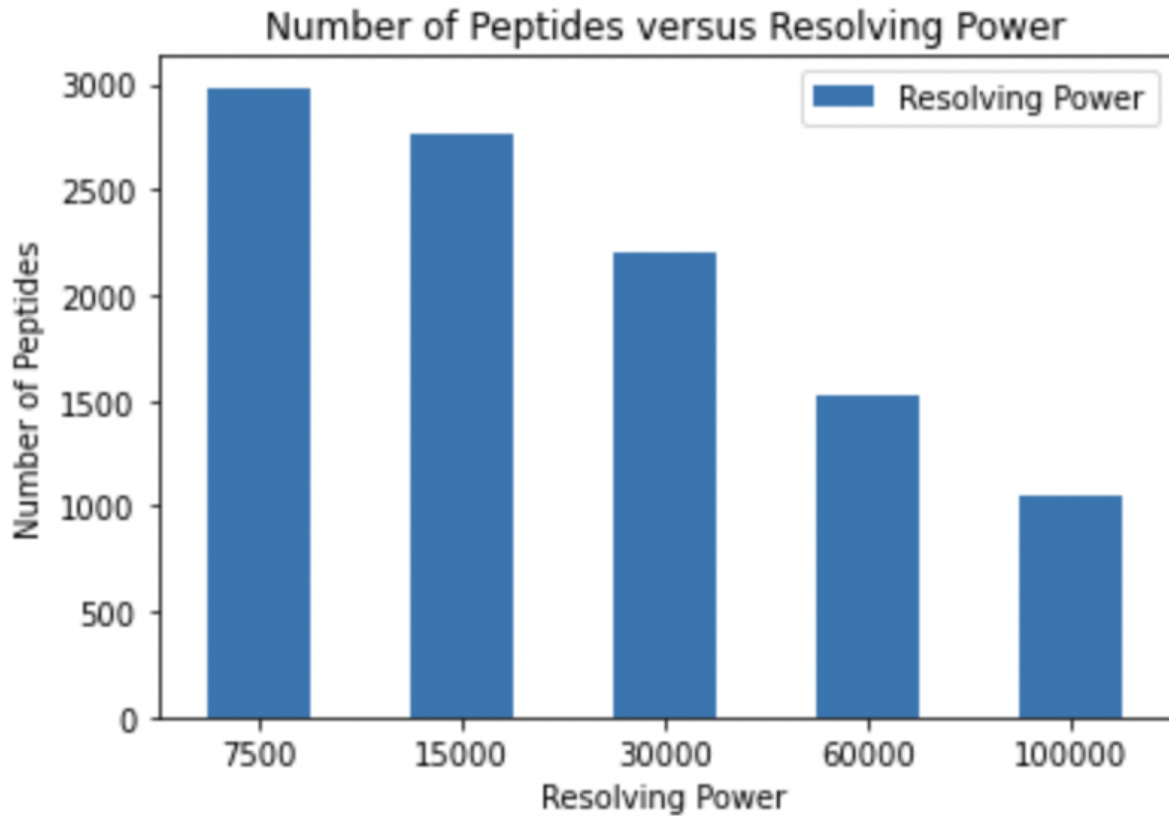

Figure 1: A bar graph depicting the number of peptides identified at each resolving power for a Human *E. coli* mixed sample.

As seen in Figure 1, the most number of peptides identified are at 7500 RP and at each RP less peptides are identified. In cases when resolving ion distributions are necessary, a higher RP is more useful but comes with the tradeoff of less peptide identification.

There are many different methods of quantification done using MS. In this experiment, we use isobaric labeling or tagging to quantify proteins with MS. The most popular isobaric label is known as a tandem mass tag (TMT). Isobaric tagging is a method used in quantitative

proteomics to label peptides or proteins with chemical groups which are identical in mass. Different samples are labeled with chemically distinct isobaric tags such that all the peptides in a given sample will fragment differently to release a reporter ion that is characteristic for that sample. In other words, we can measure the intensities of reporter ions released from the tags during MS. These results of these samples are then reported by the mass spectrum. With the mass spectrum, we are able to take the resulting data and identify the intensities reported for each sample, helping us to identify the differences in samples, which are the relative abundances between the samples as seen in Figure 2B.

Figure 2 shows a mass spectrum in the form four chromatograms of the relative abundance of ions at different m/z. The y-axis contains the relative abundance and the x-axis contains the m/z. The spectrum is a tandem spectrum where ions with m/z within a single window (695 ± 0.2) were fragmented and the products of the fragmentation analyzed.
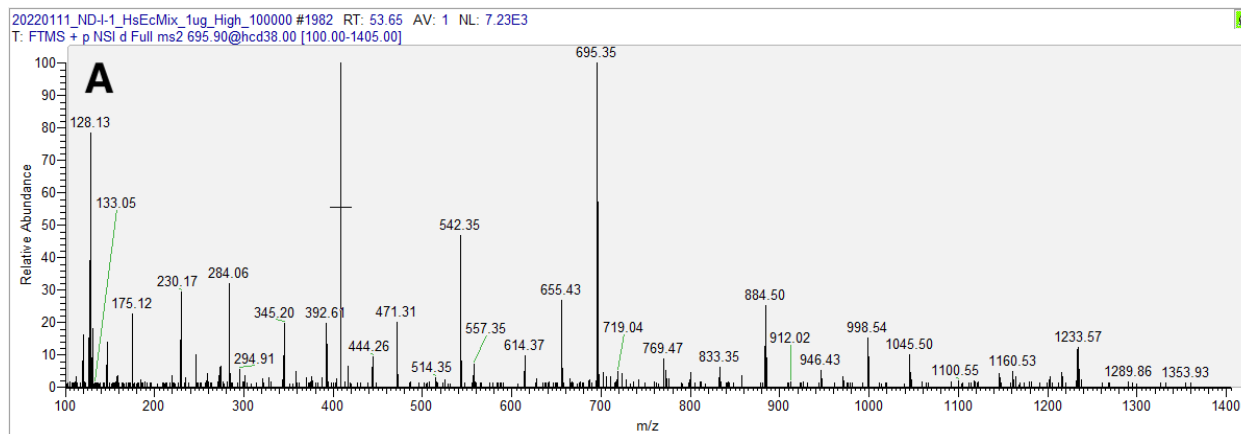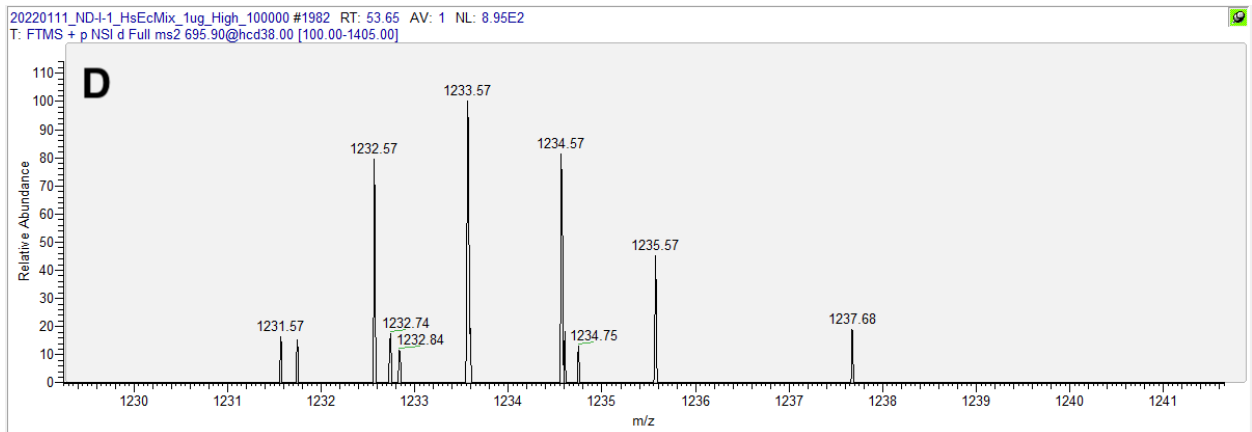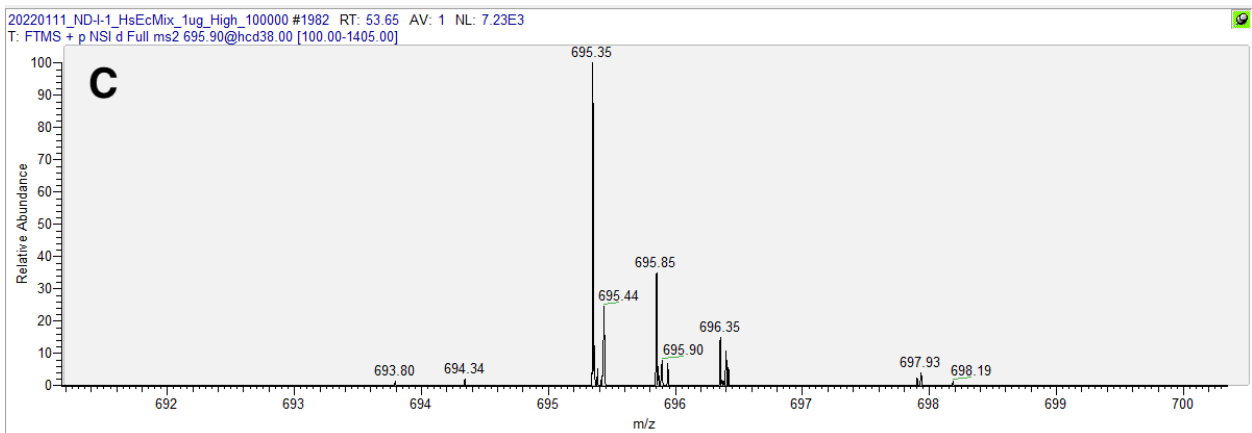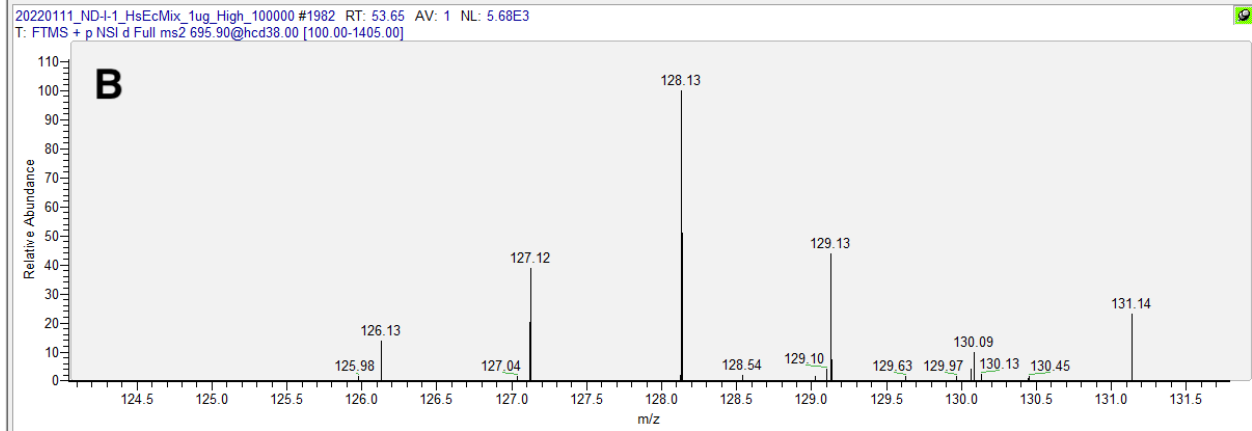
Figure 2: Chromatograms showing a mass spectrum output of a run containing thousands of spectra from the same scan. Each of the regions are all of the same scan.

In Figure 2A, an example of a tandem mass spectrum is shown. The precursor peptide that was isolated for fragmentation has a charge of 2+ and a mass of 1390.70, and hence has an

m/z of 695.35. The reporter ion is shown on the far left (m/z of 128.13), the ion of the precursor peptide in middle (695.35 m/z), and the complement ion on the far right (1233.57 m/z). Both the reporter ion and the complement ion are singly charged, so their m/z is nominally the same as their mass. Figure 2B shows a zoomed in view of the reporter ions, clearly showing very little isotopic impurities with the impurities added on top. Figure 2C shows the ion of the peptide labeled with TMT with the co-eluting and co-isolating ions having nearby m/z. Figure 2D shows the convolution between the different labels and the different isotopes of the complement peptide. The complement ions are very similar to the reporter ions except that as for the precursor, the impurities are separated.

Specifically, we use TMT in this experiment for quantification with MS. However, due to instrumentation being unable to isolate ions precisely if there are other ions nearby with similar m/z, the m/z reported by TMT-MS2 suffers from inaccuracy. TMTc is a method which almost eliminates this interference but suffers from low measurement precision. SPS-MS3-TMT is also another method which fixes the inaccuracy TMT-MS2 suffers from, but is only available to be used on the most expensive instrumentation. Another method, TMTc+, is reported to eliminate the inaccuracy of TMT-MS2 and poor precision of TMTc while still being viable on less expensive orbitraps.

TMTc+ adds modeling of the isolation step from TMT-MS2 into the deconvolution algorithm used by TMTc. Additionally, TMTc+ adds the shape of the isolation window into the deconvolution algorithm along with using the complement ions as seen in Figure 1D. Our goal is to see if TMTc+ can be used to benchmark with respect to how bad the ratio compression is to a given type of sample. Ratio compression occurs when significant changes in points of data go unnoticed as the averages of the data shows the overall changes in the data to be less significant.

If so, in cases when ratio compression is unknown, TMTc+ would be used at high resolution to compare with reporter ions. When the reporter ions agree with TMTc+, TMT-MS2 can then be used in place of TMTc+.

In Figure 3, ratio compression both with TMTc+ and without TMTc+ is displayed. In Figure 3A, the isolation window for reporter ions is shown. The contaminating ion is shown in an m/z similar to the target ion. Figure 3B and 3C represent the isolation window for TMTc+ and what the reporter and complement ions look like after isolation and fragmentation.
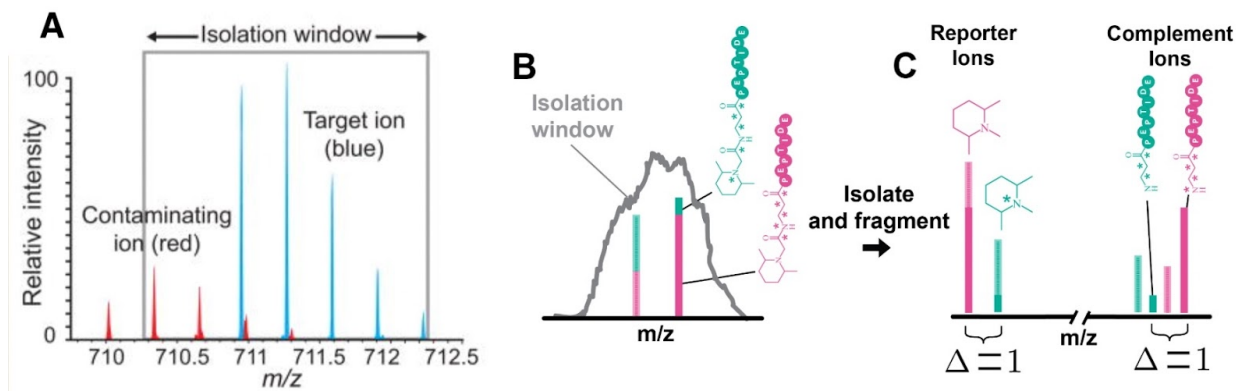

Figure 3A: Ting et. al., image screenshotted. Figure 3B, 3C: Sonnet et. al., image screenshotted. Schematic depicting isolation window and fragmentation for reporter ions and TMTc+.

In Figure 3A, there is clearly minimal ratio compression due to the contamination ion as seen by the relative intensity of the target ion and impurities of the contaminating ion near the target ion. Figure 3B shows indistinguishable peptide peaks where multiple peptides are co-isolated and shown in Figure 3C. Figure 3C shows how the reporter ions are distinguishable by mass while the complement ions are also distinguishable due to the peptides' unique masses. In other words, Figure 3 presents very little ratio compression for TMTc+ similar to that of the reporter ions.

Experimental Section

In order to determine if TMTc+ can be used to benchmark a given type of sample, we prepared multiple samples to test TMTc+ with. Samples consisted of human cell lysate in a 1 to 4 to 10 to 4 to 1 ratio across five TMT channels, added into labeled *E. coli* lysates to increase the density of potentially interfering peptides. The exact steps of sample preparation are listed below.

Clean-Up Protocol for MS

1) Use MS-grade solvents and nanopure water.

2) Begin with 300μg human cell lysate and 300μg *E. coli* of protein in a microcentrifuge tube.

3) Add 450μL MeOH into microcentrifuge tube and vortex.

4) Add 150μL $CHCl_3$ into microcentrifuge tube and vortex.

5) Add 450μL water and vortex well. Spin at greater than 16,000 rpm for 2 minutes.

6) Discard supernatant being careful not to disturb the pellet.

7) Add 500μL MeOH. Vortex, spin, and discard the supernatant.

8) Wash with MeOH. For SDS elution, wash twice.

9) Spin a second time to remove residual liquid from walls of tube.

10) Air dry pellet.

11) Only prepare buffers in disposable and filter tubes.

TMT Labeling - Digestion

1) Resuspend pellet in 5μL 1% aqueous rapigest.

2) Bring to 50μL 100mM HEPES.

3) Reduce with 1.28μL TCEP stock.

4) Alkylate with 1.44μL iodoacetamide stock.

5) Digest with 1 to 100 trypsin, 1 to 100 LysC overnight at 37°C, 600 rpm.

6) Bring to 180μL total volume with water so each sample is 50μg/30μL.

TMT Labeling

1) Aliquot both lysate samples to 5x30μL for 10 tubes.

2) Add to each tube 1 to 1 TMT with greater than 40% acetonitrile.

3) Each tube will receive 20μL (μg) of appropriate TMT label.

4) Bring to 40μL acetonitrile if samples are low.

5) Mix well and allow samples to react for 1 hour.

6) Quench by bringing to 10% ammonium bicarbonate.

7) Mix well and allow samples to sit for 1 hour.

8) Combine TMT labeled samples to 10μg total protein and evaporate to 10μL.

9) Resuspend in 200μL Buffer A (5% acetonitrile, 0.1% formic acid in water)

10) If the pH is greater than 2, acidify with formic acid.

11) Heat at 37°C for 1 hour, hard spin 30` to ensure rapigest precipitation.

12) Remove supernatant.

13) Store labeled peptides at -80°C.

14) Hard spin 30` after thawing samples.

LC-MS Settings

1) 2% DMSO added to both Buffer A and Buffer B.

2) Standard 1D run.

3) Scan range 500 to 1400 m/z.

4) Orbitrap: Auto m/z High.

5) Only charge states of 2+.

Each sample was labeled in the following format: For human channel 126, the sample was labeled as Hs126 to run. This was done for channels 126 - 131 respectively, excluding channel 130. For *E. coli* channel 126, the sample was labeled as Ec126 to run. This was done for channels 126 - 131 respectively, excluding channel 130. The sample mixtures are listed in the table below. The first column denotes the name of the sample. The second column denotes the amount of Buffer A. The third column denotes the amount of stock and type of stock being used in the sample. The fourth column denotes the concentration of protein in the final sample. For example, injecting 5μL of a 0.2μg/μL sample into the instrument corresponds to 1μg. The final column denotes the final volume of the sample. For Hs126 - Hs131, each Hs126, Hs127, Hs128, Hs129, Hs131 are all the same for what is listed. The same is true of Ec126 - Ec131.

Sample	Buffer A	Stock	Final concentration of protein in sample	Final Volume
Hs126 - Hs131	8.63μL	2.37μL Hs126 - Hs131	0.2μg/μL	11μL
Ec126 - Ec131	8.63μL	2.37μL Ec126 - Ec131	0.2μg/μL	11μL
Hs 1 All	5.75μL	2.37μL Hs126 - Hs131	0.125μg/μL	17.60μL
Ec 1 All	0μL	14.20μL Ec126 - Ec131	0.186μg/μL	71.0μL
Hs V	0μL	3.55μL Hs126		71.0μL
Hs V	0μL	14.20μL Hs127		
Hs V	0μL	35.4μL Hs128		
Hs V	0μL	14.20μL Hs129		

Sample	Buffer A	Stock	Final concentration of protein in sample	Final Volume
Hs V	0μL	3.55μL Hs131		
EcHs Mix	0μL	59.20μL Ec 1 All		118.40μL
EcHs Mix	0μL	59.20μL HsV		

The samples were run on a Thermo LTQ Orbitrap Velos Pro Mass Spectrometer orbitrap. Hs126, Hs127, Hs128, Hs129, Hs131, Ec126, Ec127, Ec128, Ec129, Ec131, Hs 1:1:1:1:1, Ec 1:1:1:1:1 all ran using the "High" method. Buffer A consists of 5% aqueous acetonitrile, 0.1% formic acid, 2% DMSO. Buffer B consists of 80% aqueous acetonitrile, 0.1% formic acid, 2% DMSO. Samples were loaded onto a homemade trapping column (0.1μm inner diameter, 1cm long, filled with 0.3μm C18 stationary resin), washed well with Buffer A, and then separated on a homemade analytical column (0.1 μm inner diameter, 20 cm long, 5 μm tip, filled with 0.3 μm C18 stationary resin). The samples were maintained at 25 °C, using the following gradient: start at 1% 80acetonitrile, 20H20 for to 7% for 5 minutes, from 7% to 36% for 135 minutes, from 36% to 100% for 50 minutes, down to 1% for 20 minutes. Samples were ionized into the mass spectrometer using electrospray (3.0 kV). Precursor ions were analyzed in the orbitrap at 30000 resolving power 500 to 1400 m/z. Ions were chosen for fragmentation on the basis of top 10 intensity, 2+ charge state, with 50 second dynamic exclusion (500 ion exclusion list). Fragmentation was performed using HCD (100 ms activation, stepped collision mode at 32%, 38%, and 44%). Fragmentation spectra were collected at the indicated resolving power of 7500, 15000, 30000, 60000, and 100000 from 100 m/z to double the precursor plus approximately 10 m/z m/z. Raw files were converted to mzml files using MSConvert version 3.0.21105-a3d3afac6, and searched for peptides against a Uniprot human fasta database and *E.*

coli fasta database concatenated (downloaded 2021-02-23, longest isoform of each protein and including contaminants and reversed decoy set for 51438 total sequences) using MSFragger in the FragPipe pipeline.

Raw files were searched for peptides using FragPipe version 17.1. MS data type was regular MS. Variable modifications included Methionine and Lysine. Fragment mass tolerance was 20 PPM, cleavage enzymatic with N-term M clipping. The isobaric label type was TMT-6 with quant level 2. Min PSM probability was 0.9, min purity 0.5, and min intensity 0.05.

For TMTc+ analysis, we ran samples which were labeled with a single TMT label so that we could experimentally measure the isotopic impurities of our lot. These impurities were then provided to TMTc+ as an input matrix to allow deconvolution of the chemical impurity for the reagents. The samples were run with TMTc+ using the following settings:

1) The isolation window was set to 0.4.

2) The impurities were changed to reflect what was previously mentioned.

3) The raw files for each sample were searched using FragPipe.

4) All other settings are default.

Results and Discussion

One method of evaluating the quantitative accuracy of TMTc+ to TMT-MS2 was done by comparing the ratios of TMTc+ to that of the reporter ions. Lysates were measured separately to get the ratios of every channel. The relative ratios for peptides tagged with TMT reagents should correspond to how the channels were distributed in sample preparation. The channels were distributed in a 1 to 4 to 10 to 4 to 1 ratio of human cell lysate added into *E. coli* lysates. In other words, human cell lysate was added to *E. coli* lysates to create a 1 to 4 to 10 to 4 to 1 ratio of

human cell lysate to *E. coli*. For each of the mixed human cell samples and *E. coli* samples we determined the true ratios separately which are represented by the dashed line in the box plots of Figure 3. The actual ratio was determined to account for any errors in preparing the original samples.

In Figure 3, we look at the results of running TMTc+ at 100,000 nominal resolving power. The actual RP used by the orbitrap varies with the square root of m/z. We expect at this high RP TMTc+ will be able to resolve ion distributions more clearly. For each sample, when we look at the TMT complement ions we see that the complement ions generally look like the reporter ions. In other words, the averages of the complement ions are similar to that of the reporter ions and the variances are not very different.

In the first box plot of Figure 4, the results of running TMTc+ at 100,000 RP for human peptides are shown. The reporter ions are shown in the second box plot. On the x axis, each channel is labeled with respect to its human cell sample. The y axis is the ratio of the specific peptide for a given sample. Peptides with a sum of signal to noise of less than 100 were filtered out. Each channel for TMTc+ was calculated by dividing the ratio reported by TMTc+ by the sum of the ratios of TMTc+. Similarly, each channel for the reporter ions was calculated by dividing the ratio reported by the reporter ions by the sum of ratios of the reporter ions. The actual ratios for TMTc+ and reporter ions are shown in the legend as each colored line.
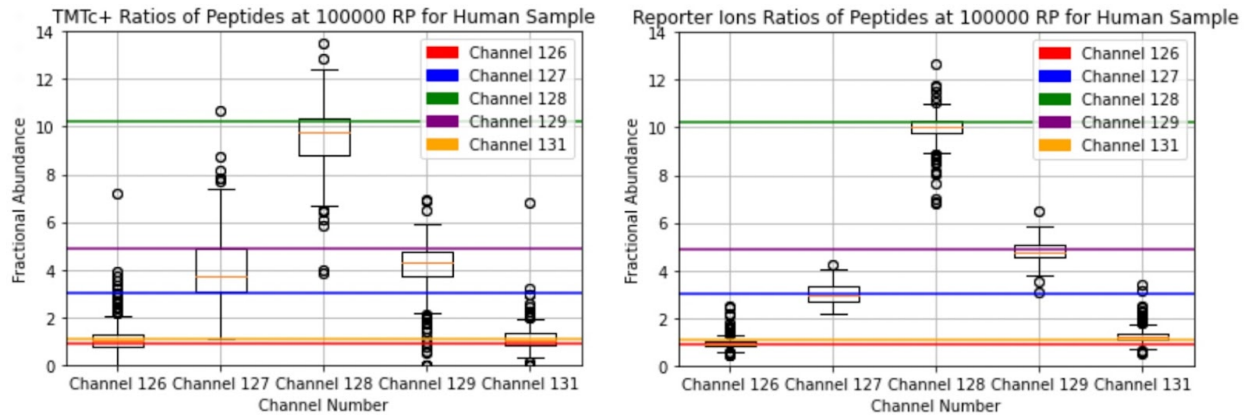
Figure 4: Box plots representing peptide ratios for human cell samples ran at 100,000 RP for TMTc+ and reporter ions.

As seen, the figure's median for each sample does come close to what we would expect for TMTc+. However, each channel has outliers indicating a heavy-tailed distribution. In other words, it is likely the distribution is skewed with channel 126 and 131 trending towards ratio compression. Furthermore, the box plot for TMTc+ shows a wider range of outliers than depicted by the reporter ions, and the reporter ions have a tighter fit on average values reported. This suggests that at 100,000 RP, the reporter ions are at minimum just as accurate at quantifying peptides as TMTc+ is.

In Figure 5, the results of running the samples at 7500 RP for each TMTc+ and reporter ions of human peptides are shown. The axis represents the same information as represented in Figure 4. The ratios were calculated the same as in Figure 4 as well.
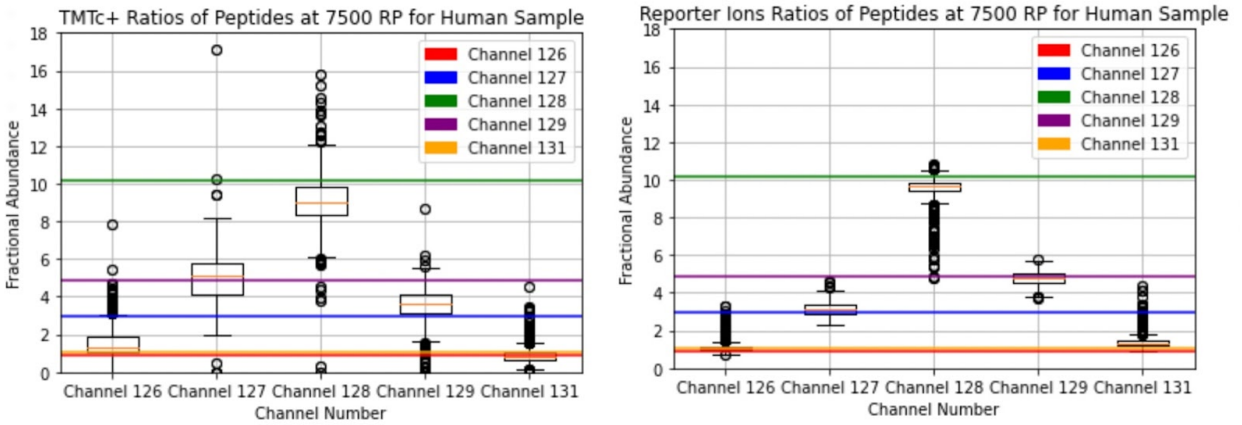
Figure 5: Box plots representing peptide ratios for human cell samples ran at 7500 RP for TMTc+ and reporter ions.

Once again, at 7500 RP the median ratios for each channel are close to what we would expect. Similarly, the results shown by TMTc+ have less tight fitting data than that of the reporter ions. Again, we see the data trending towards ratio compression suggesting TMTc+ is not resolving ion distributions better than the reporter ions at low RP as well. This suggests that at 7500 RP, the reporter ions are at minimum just as accurate at quantifying peptides as TMTc+ is just as seen at 100,000 RP.

In order to further quantify the accuracy of the reporter ions versus TMTc+, Figure 6 represents the deviation between the measured distribution of peptides and the actual distribution of peptides between the channels. The figure shows scatter plots of the peptide ratio values for reporter ions versus TMTc+ for 100,000 RP. The y-axis represents points quantified by TMTc+ and the x-axis represents points quantified by reporter ions. The scatter plots contain a unity line which is used to evaluate the accuracy of TMTc+ versus reporter ions. Points above the unity line had less accurate quantification for TMTc+ and points below the unity line were quantified less accurately by the reporter ions.
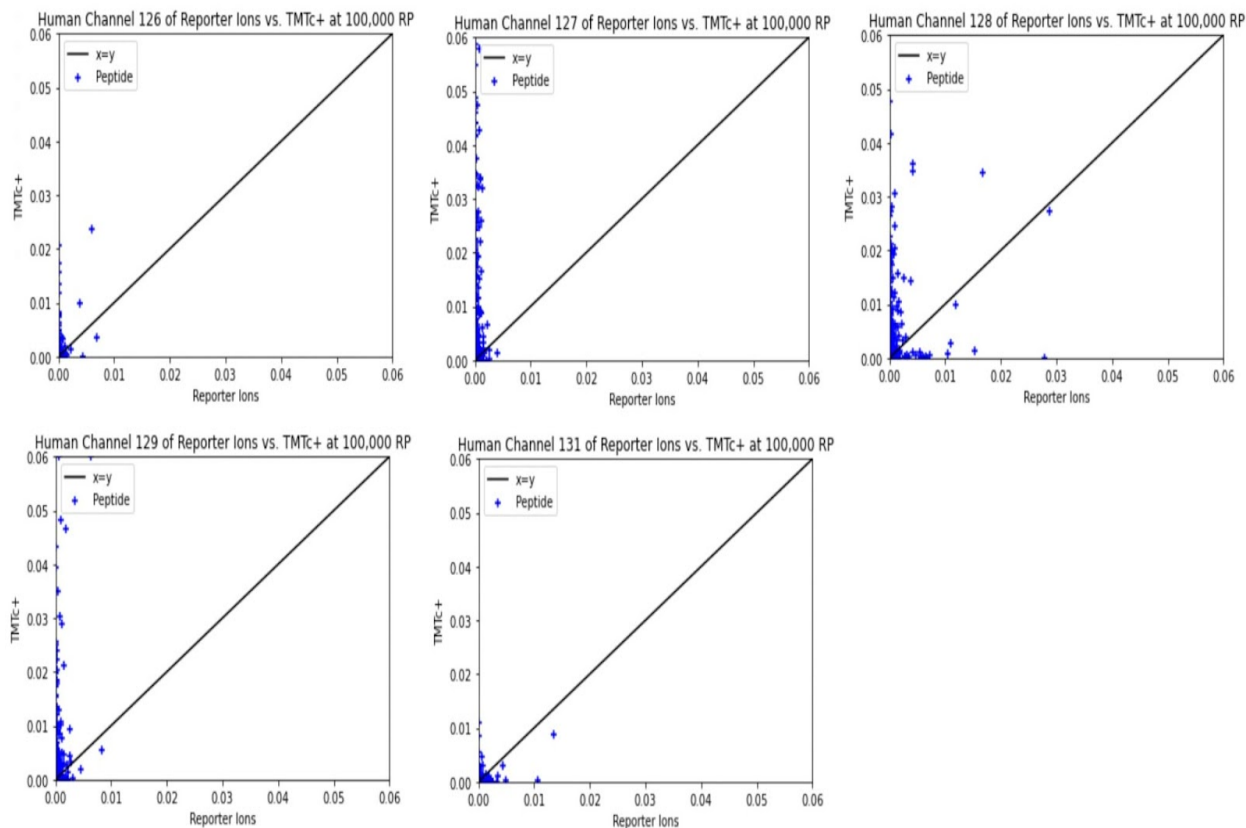
Figure 6: Scatter plots representing peptide ratio values of reporter ions versus TMTc+ at 100,000 RP for each sample.

As seen in Figure 6, channel 127, 128, and 129 all clearly show the majority of data above the unity line. This suggests that for these channels, TMTc+ is less accurate in quantification of peptides when compared to the reporter ions. However, as the scatter plots do not precisely quantify the accuracy of TMTc+ versus reporter ions, we show a sum of squares deviation bar graph across the human and *E. coli* samples for each resolution in Figure 7.

The sum of squares deviation precisely measures the variability of the quantification for TMTc+ versus reporter ions. The sum of squares deviation was calculated by taking the sum of squares of the TMTc+ values divided by the sum of squares of the reporter ion values. The y-axis represents the sum of squares deviation and the x-axis represents the resolving power at each sample. Sample's 1 through 5 correspond to each of channel's 126 through 131.
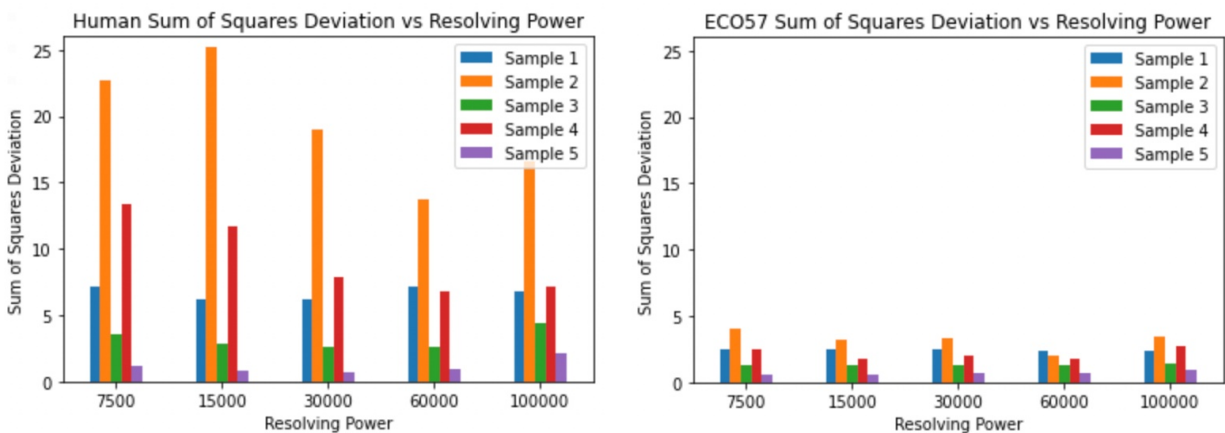
Figure 7: Bar plots representing deviations between the measured and actual fractional intensity across all channels and peptides.The fraction of labeling intensity for each peptide in each channel was subtracted from the actual fractional labeling intensity for that channel. Each plot presents the summed and squared deviation as determined by TMTc+, divided by the summed and squared deviation determined by reporter ion integration.

For the human cell sample, the deviation among channels 127 and 129 are high. For the *E. coli* sample, the deviation is highest in channel 127. These deviations are seen for every resolution in both the human and *E. coli* sample, suggesting that across the board TMTc+ is unable to outperform quantification done by the reporter ions.

Possible reasons for failure could be due to failures associated with data analysis. This could be a possible step in failure due to co-isolation, failures in the TMTc+ deconvolution algorithm, or other failures in interference. TMTc+ does not appear to be performing as well as reporter ions in the context of these given samples.

Conclusion

For most samples, there is no evidence of peptides present at a ratio that is different from the overall ratio. Furthermore, for all human cell samples TMTc+ and reporter ion ratio measurements show no correlation of peptides being present at ratios different from the overall

ratio. For each of the samples, the averages of the complement ions appear similar to that of the reporter ions displaying minor variances. Both reporter ions and TMTc+ shows data suggesting ratio compression, however TMTc+ displays a wider range of outliers suggesting underperformance to that of the reporter ions. Not only was this seen at 7500 RP, this appears again at 100,000 RP suggesting the increase in RP does not contribute a significant difference.

Furthermore, Figure 6 clearly shows TMTc+ being on average more inaccurate in peptide quantification than reporter ions. For each sample, the reporter ions are seen outperforming TMTc+ at every single RP. These results are supported by the deviations presented in Figure 7, where at each RP a high sum of squares deviation is seen with TMTc+ compared to reporter ions. Ultimately, it is seen that in the context of the samples created in this experiment, TMTc+ does not outperform reporter ions. Additionally, TMTc+ cannot be used as a benchmark to determine how bad ratio compression is for the given samples.

REFERENCES

Sonnett, Matthew, Eyan Yeung, and Martin Wühr. 2018. "Accurate, Sensitive, and Precise
Multiplexed Proteomics Using the Complement Reporter Ion Cluster." *Analytical Chemistry*
90 (8): 5032–39.

Ting, Lily, Ramin Rad, Steven P. Gygi, and Wilhelm Haas. 2011. "MS3 Eliminates Ratio
Distortion in Isobaric Multiplexed Quantitative Proteomics." *Nature Methods* 8 (11):
937–40.

Pappireddi, Nishant et al. "A Review on Quantitative Multiplexed Proteomics." *Chembiochem :
a European journal of chemical biology* vol. 20,10 (2019): 1210-1224.
doi:10.1002/cbic.201800650

Zecha, Jana et al. "TMT Labeling for the Masses: A Robust and Cost-efficient, In-solution
Labeling Approach." *Molecular & cellular proteomics : MCP* vol. 18,7 (2019): 1468-1478.
doi:10.1074/mcp.TIR119.001385

Savitski, Mikhail M et al. "Measuring and managing ratio compression for accurate
iTRAQ/TMT quantification." *Journal of proteome research* vol. 12,8 (2013): 3586-98.
doi:10.1021/pr400098r

Roumeliotis, Theodoros I et al. "Evaluation of a Dual Isolation Width Acquisition Method for
Isobaric Labeling Ratio Decompression." *Journal of proteome research* vol. 18,3 (2019):
1433-1440. doi:10.1021/acs.jproteome.8b00870

Roumeliotis, Theodoros I et al. "Evaluation of a Dual Isolation Width Acquisition Method for Isobaric Labeling Ratio Decompression." *Journal of proteome research* vol. 18,3 (2019): 1433-1440. doi:10.1021/acs.jproteome.8b00870