

Evaluation of Advanced Precursor Determination for Tandem Mass Tag (TMT)-Based Quantitative Proteomics across Instrument Platforms

Samuel A. Myers,^{*,†} Susan Klaeger,[†] Shankha Satpathy,[†] Rosa Viner,[‡] Jae Choi,[§] John Rogers,[§] Karl Clauser,[†] Namrata D. Udeshi,[†] and Steven A. Carr^{*,†}

[†]The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, United States

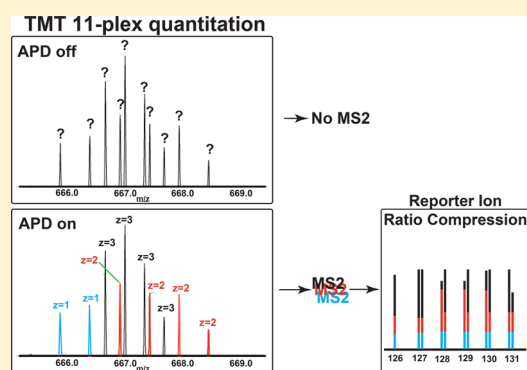
[‡]Thermo Fisher Scientific, San Jose, California, United States

[§]Thermo Fisher Scientific, Rockford, Illinois, United States

Supporting Information

ABSTRACT: Tandem mass tag (TMT)-based quantitation is a strong modality for quantitative proteomics, as samples can be multiplexed, creating large-scale data sets with high precision and minimal missing values. However, coisolation/cofragmentation of near isobaric, coeluting precursor peptide analytes has been well-documented to show ratio compression, compromising the accuracy of peptide/protein quantitation. Advanced peak determination (APD) is a new peak-picking algorithm that shows improved identification of peak detection in survey scans (MS1) to increase the number of precursors selected for unimolecular dissociation (MS2). To increase the number of these “features” selected for MS2 APD purposefully selects multiple peptide precursors of very similar m/z that often derive from different proteins—a major source of ratio compression in TMT quantification. Here, we evaluate the effects of various data acquisition parameters combined with APD on ratio compression. We find that data acquisition with APD enabled results in more coisolated precursors, more mixed spectra, and in turn, fewer peptide spectral matches, especially at standard on-column loads. We conclude that APD should not be utilized for isobaric tagging, MS2-based experiments.

KEYWORDS: isobaric tagging, ratio compression, advanced peak determination, TMT, precursor ion purity



INTRODUCTION

Isobaric peptide labeling for relative quantitation (tandem mass tags, TMT; isobaric tagging for relative and absolute quantitation, iTRAQ) has proven to be a valuable method for relative quantitation of peptides and proteins across a large number of samples in discovery proteomic analyses.^{1–4} Samples can be multiplexed, and the data can be acquired, together resulting in data sets with minimal missing values across experimental conditions.⁵ However, coisolation and cofragmentation of coeluting and near-isobaric precursor peptide ions can result in a convolution of reporter ion signals, resulting in a phenomenon commonly referred to as “ratio compression”.^{6–9} Ratio compression has been traditionally combated by off-line peptide mixture fractionation followed by liquid chromatography–mass spectrometry/mass spectrometry (LC-MS/MS) of fractions and MS2 level quantitation, or by synchronous precursor selection (SPS) followed by MS3 level quantitation.^{10–12}

Recently, advanced precursor determination (APD) has been implemented on the Orbitrap Fusion Lumos and Q Exactive HF-X mass spectrometers (Thermo Fisher Scientific).

The algorithm resolves overlapping, high charge states, or low-intensity MS1 isotope distributions “on the fly” for charge-state determination and subsequent selection for unimolecular dissociation (MS2 or MS3). For label-free peptide analysis, APD has been shown to increase the number of precursors selected and identified for MS2 by identifying charge states previously unidentifiable, particularly for ion trap MS2 scans.¹³

APD was developed to increase the number of peptides identified in label-free peptide experiments,¹³ but its utility for TMT-based peptide identification and quantitation has not been critically examined. Because the benefit of APD comes from its ability to identify overlapping isotope distributions for MS2 acquisition, one can hypothesize that this may affect the ratio compression of peptides analyzed during TMT-based experiments. Here, we evaluate the performance of APD during TMT 11-based data acquisition for low (10 ng) and high (1 μ g) on-column loads using a knockout TMT 11-plex standard, for various separation gradients on both the Q

Received: August 8, 2018

Published: October 23, 2018

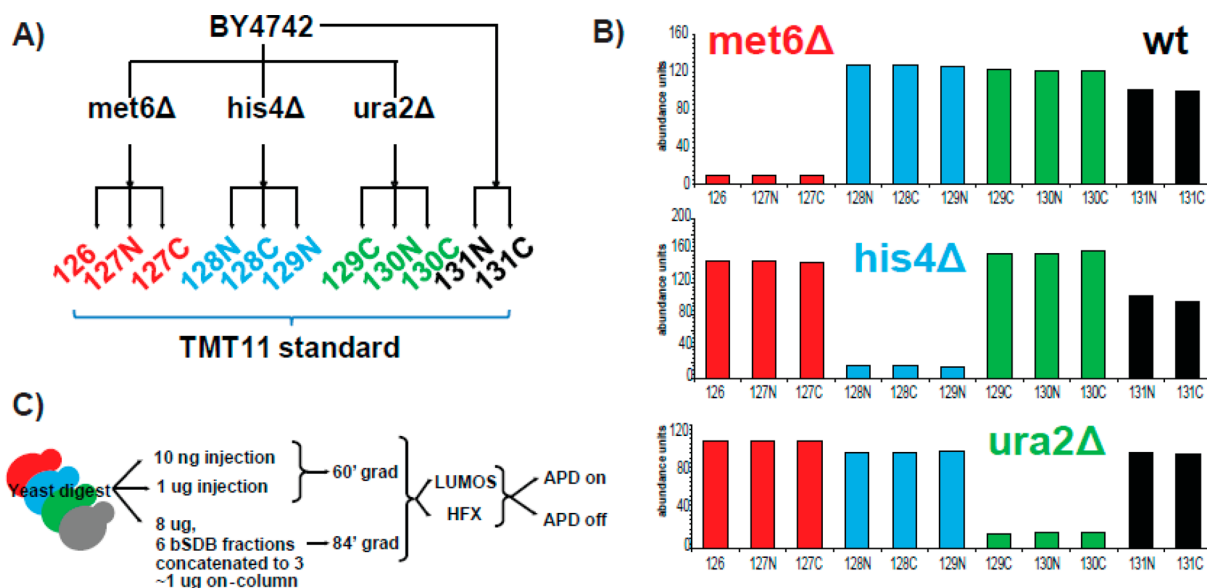


Figure 1. *S. cerevisiae* knockout strain standard for TMT 11-plex assessment and optimization. (A) TMT 11-plex design and genetic knockout strains used in this study. (B) TMT total signal-to-noise protein abundance ratios of respective proteins for initial characterization of TMT-11 plex standard. One μg on a 75 μm x 50 cm Thermo Scientific EASY-Spray column coupled to an Orbitrap Fusion Lumos instrument using SPS MS3. (C) Experimental design used to evaluate APD during data acquisition for multiple on-column loads, gradient lengths, and instrument platforms.

Exactive HF-X and Fusion Lumos. We find that APD should not be used in combination with MS2-based quantitative proteomic analyses employing isobaric mass tag labeling.

EXPERIMENTAL PROCEDURES

Saccharomyces cerevisiae knockout strain TMT 11-plex was created and initially assessed by Thermo Fisher Scientific (Figure 1A,B). Lyophilized peptides were resuspended in 50% acetonitrile (ACN) and 0.1% formic acid (FA) at 1 mg/mL and centrifuged at 16000g for 5 min at 4 °C, and the supernatant was transferred and stored at -80 °C. For single-shot analysis, peptides were transferred as is, to high-performance liquid chromatography (HPLC) vials, dried by vacuum centrifugation, and resuspended in the appropriate volume for 1 μL injections. For basic reversed-phase fractionation (bRP), a Stage tip¹⁴ was packed with sulphonateddivinylbenzene (SDB) mesh (two punches) using a 16-gauge needle. After the SDB was activated and equilibrated, 8 μg of the yeast standard in 0.1% FA were loaded as called for by the standard Stage tip protocol (50 μL of methanol, followed by 50 μL of solvent B (0.1% FA/90% ACN), then 75 μL of solvent A (0.1% FA twice). Ammonium formate (25 μL , pH 10) was used as a buffer switch and collected as part of fraction one. $\text{NH}_4\text{OAc}/5\%$ ACN (75 μL) was then used for the first elution, followed by 100 μL of increasing ACN percentage (10, 15, 20, 30, and 50%). Fractions were then concatenated (5 + 50, 10 + 20, and 15 + 30) to form three final fractions for subsequent analysis.

DATA ACQUISITION

Peptide mixtures were separated using a Thermo Scientific EASY-nLC1200 UPLC system at a flow rate of 200 nL/min. Peptides were separated at 50 °C on a 75 μm id PicoFrit (New Objective) column packed in-house with 1.9 μm AQ-C18 material (Dr. Maisch) to 20 cm in length. Single shot analyses were separated using a 60 min separation gradient from 2% to 30% solvent B, followed by a ramp to 50% B in 5 min. The

fractionated samples followed the same gradient, though the separation gradient was 84 min. Initial experiments were run using a gradient from 8 to 32% acetonitrile (v/v) in 50 min and 75 μm id x 50 cm an EASY-Spray column at a flow rate of 300 nL/min.

Mass spectrometry was performed on a Thermo Scientific Q Exactive HFX or a Lumos mass spectrometer. The Lumos has a toggle switch enabling or disabling APD, Tune version 3.0.2041. For the HFX, we were provided with a software patch allowing us to turn APD off in the tune file, Tune version 2.9.0.2926 (the latest version of Exactive Tune 2.9 sp2 for the HFX now has this option as standard). For the HFX, the precursor scan ranged from 350 to 1800 m/z at 60000 resolution with an automatic gain control (AGC) target of 3×10^6 and maximum injection time (IT) of 10 ms. The isolation window was set to 0.7 Th with no offset. The top 20 most intense multiply charged precursors (2–8) were selected for higher-energy collisional dissociation (HCD) at a resolution of 45000, an AGC target of $1e5$ and maximum IT of 96 ms. Peptide match was set to preferred, and dynamic exclusion was set for 20 s. For the Lumos, the precursor scan ranged from 350 to 1800 m/z at 60000 resolution with an AGC target of $4e5$ with a max IT of 50 ms. The isolation width was 0.7 Th with no offset. The top 20 most intense multiply charged ions (2–6) were selected for HCD at a resolution of 50000, a max IT of 105 ms, and an AGC of $1e5$.

Data were searched with Spectrum Mill (Agilent) using the UniProt yeast database. Fixed modifications were carbamidomethylation at cysteine. TMT-11 labeling was required at lysine residues, but peptide N-termini were allowed to be modified either by TMT-11 labeling or unmodified. Allowed variable modifications were acetylation of protein N-termini and oxidized methionine with a precursor MH⁺ shift range of -18 – 64 Da. The enzyme specificity was set to trypsin, and a maximum of three missed cleavages was allowed. The maximum precursor ion charge state was set to 6. Search parameters included parent and fragment mass tolerance of 20 ppm, minimum matched peak intensity of 50%, scoring, and

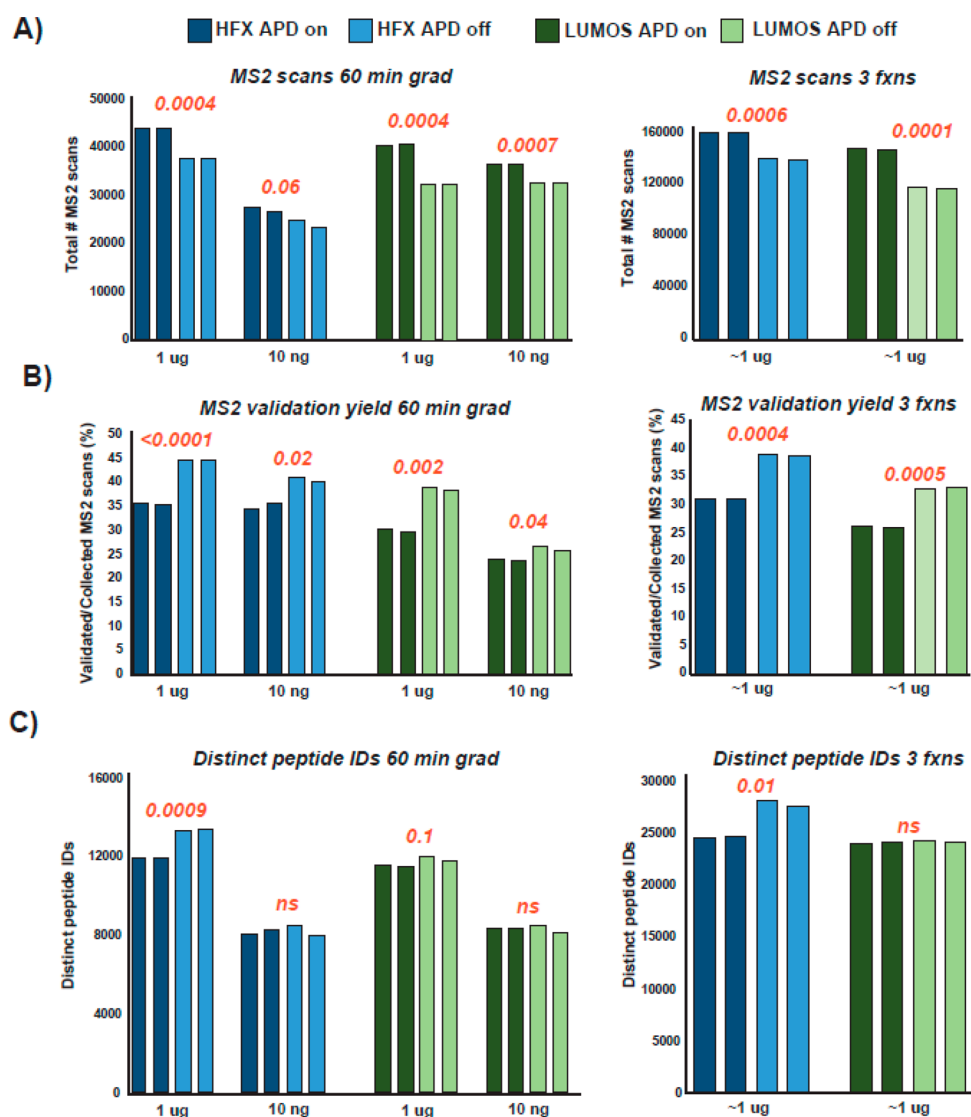


Figure 2. APD increases the number of MS2 events but not the number of peptides identified. (A) Total number of MS2 scans acquired with either 1 μ g or 10 ng on-column separated by a 60 min gradient (left) or three fractions separated by an 84 min gradient (right). Blue indicates HFX data; green indicates Lumos data. Dark colors represent APD on; lighter color APD is off. Unpaired Student's *t* test *P*-value is marked in red above the appropriate APD on vs off conditions, and it is only reported if $P \leq 0.1$. Each acquisition replicate is shown. "ns", not significant. (B) Percentage of MS2 validated as a peptide spectral match (PSM) divided by the number of MS2 scans acquired. Colors and *P*-values the same as in (A). (C) Number of distinct peptide identifications for each acquisition condition. Colors and *P*-values the same as in (A).

peak detection parameters for high-resolution HCD spectra ("ESI QExactive HCD 35"), and calculation of reversed database scores was enabled. Peptide and protein false discovery rates (FDRs) were both calculated to be less than 1%.

Spectrum Mill's Precursor Isolation Purity (PIP) metric was calculated for ions present in the precursor isolation window (0.7 Th). PIP is the ratio, expressed as a percentage, of the sum of the intensities of the peaks in the precursor ion's isotope cluster to the sum of all ion intensities in the isolation window. The Scored Peak Intensity (SPI) metric is the ratio, expressed as a percentage, of the sum of the intensities of the peaks annotated to the identified peptide by supported fragment ion types to the sum of all ion intensities in the MS2 spectrum. The intensities used are after peak detection has been done, which includes deisotoping, removal of residual precursor ion and its neutral losses, and thresholding to the top 25 peaks by

signal/noise. Our typical PIP cutoff of 50% for peptide quantification was not used in the current study.

Interference free index (IFI) was calculated as previously described,¹¹ except instead of signal to noise (S/N) we used protein abundance as a measure of ratio compression. Protein abundance was determined by taking the summed MS1 intensity for all peptides from a particular protein, multiplied by the ratio of one TMT channel's contribution to the total MS1 intensity divided by the sum of all 11 channels.¹⁵

RESULTS AND DISCUSSION

To evaluate the effect of APD on TMT-based data acquisition and quantitation, we utilized a commercially available digested and TMT 11-plex-labeled yeast standard from Thermo Fisher in which the genes encoding Met6p, His4p, and Ura2p had been deleted. The TMT 11-plex Yeast Digest Standard is based on the triple knockout standard previously shown to provide useful metrics for instrument performance and ratio

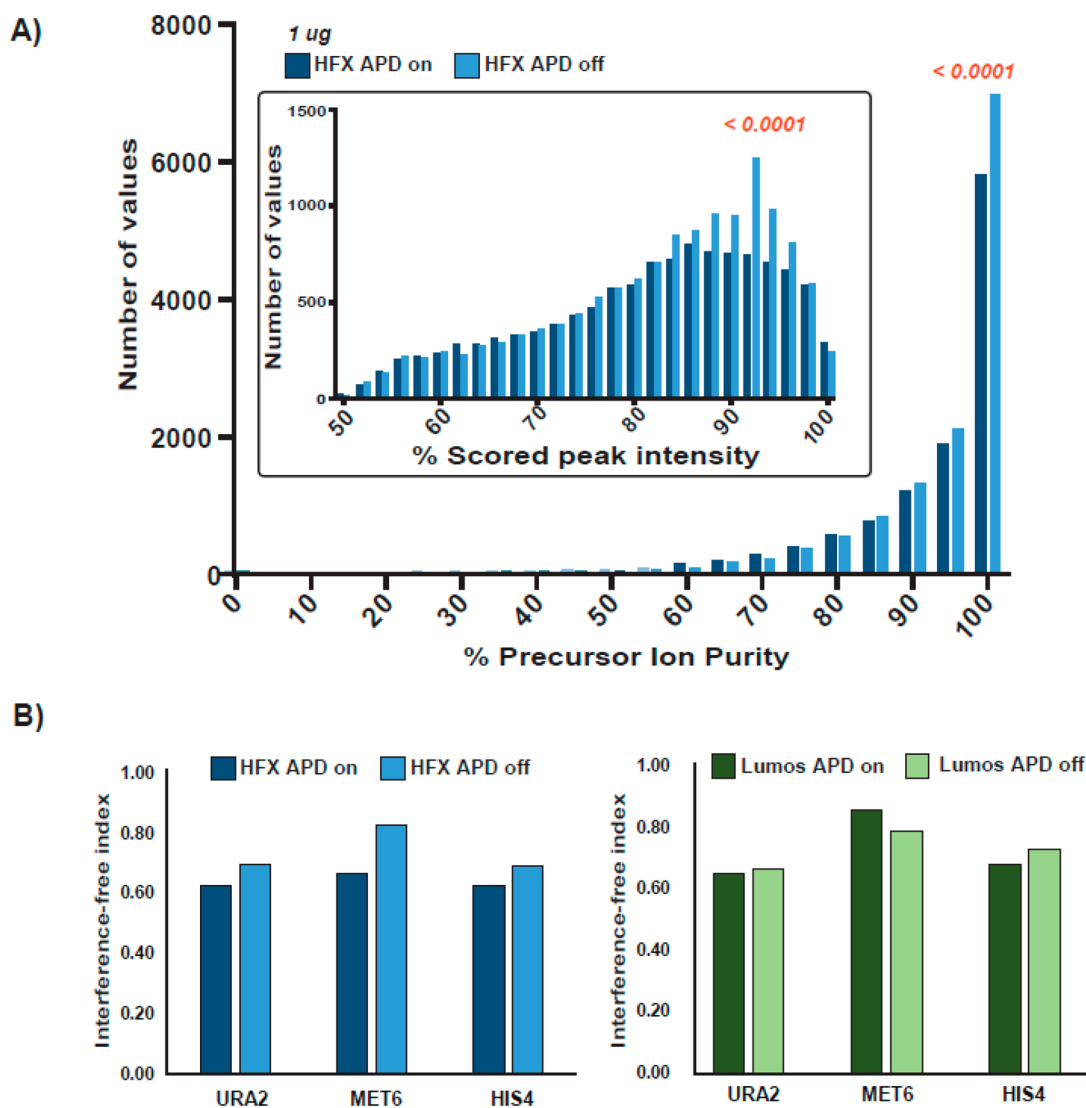


Figure 3. APD results in less pure MS1 and MS2 scans leading to more ratio compression. (A) Histogram for precursor ion purity for the 1 µg on-column, 60 min gradient experiment, as a proxy for purity of coisolation. (inset) A histogram for the percentage of product ions for a given MS2 scan that match the interpreted peptide (higher is better). Wilcoxon test P values comparing the distributions between APD on (dark blue) to the APD off (light blue) are shown in red. (B) Average IFI for the three gene products knocked out in the yeast strains analyzed during 1 µg on-column data acquisition. IFI for each replicate for each experiment is shown in Figure S1.

compression.¹¹ The new standard differs from previous iterations, as it is 11-plex rather than nine, contains two channels of wild-type *S. cerevisiae* strain BY4742, and has *His4* knocked out instead of *Pfk2* (Figure 1A). Initial testing of this standard showed marked loss in signal-to-noise (S/N) ratios for certain proteins consistent with the genetic knockout strain (Figure 1B). We chose to assess APD for TMT across multiple conditions, including high and low on-column loads (1 µg and 10 ng, respectively), with and without off-line BRP fractionation, and on two separate mass spectrometers (HFX and Lumos; Figure 1C). We focused on MS2-based quantitation, because the HFX is not capable of SPS-MS3, which has been extensively studied on Tribrid Thermo mass spectrometers.^{10,11,16}

Enabling APD significantly increased the number of acquired MS2 scans (unpaired two-tailed t test P value < 0.05) across peptide loads, instrument platforms, and fractionation extents (Figure 2A). This increase is consistent with previous assessments of APD and is likely due to the

ability of APD to deconvolute overlapping MS1 isotope envelopes, thus allowing for more “peptide” analytes matching criteria for MS2.¹³ The increase in the number of MS2 scans acquired was less prominent for the 10 ng on-column loads. This is likely due to the fact that over 85% of MS2 scans reached their maximum injection time (96 ms for HFX, 105 for Lumos), whereas for the 1 µg loads, the highest median MS2 fill time was 31.6 ms (Table S1).

While APD provided an increase in MS2 scans, the ratio of peptide spectral matches to spectra collected was significantly less than with APD off (Figure 2B). The decrease in identification rate was observed across all experimental conditions tested in this study and is consistent with previous results.¹³ This decrease is likely due to the higher resolution needed for TMT 11-plex analysis, which results in a slower duty cycle. With respect to the absolute number of peptides identified, our data showed that enabling APD resulted in a modest but statistically significant decrease in the number of peptides identified for 1 µg on-column loads, (11% and 4%

decrease for HFX and Lumos, respectively; Figure 2C). These results were mainly seen on the HFX for both unfractionated and fractionated samples. At 1 μg on-column level (\pm fractionation), the Lumos also identified fewer peptides, but the differences detected were not statistically significant. At the 10 ng load level, no differences were observed in the numbers of peptides identified on either instrument. These data indicate that, from a peptide identification standpoint, enabling APD had a minor detrimental, or no, effect on proteome coverage depending on the amount of peptide injected.

Because coisolation and cofragmentation of isobarically labeled precursors is a major concern for TMT-based proteomics, we investigated the effect of APD on metrics for ratio compression.^{6–8} PIP is a measurement of the total MS1 ion current in the isolation window that is attributed to the ultimate peptide identified in that window. We found that disabling APD led to a statistically significant increase in the number of “pure” precursors isolated and subsequently identified as measured by PIP (Wilcoxon test, P -value < 0.05; Figure 3A). This increase in precursor isolation purity was corroborated by the percentage of product ions observed that could be assigned to the peptide identified (scored peak intensity [SPI]; see Methods). A higher percentage of the product ions in a given spectrum could be assigned to the peptide in question when APD was disabled (Figure 3A, inset), indicating that “APD off” generates fewer mixed spectra. While methods to identify both peptides in mixed MS2 spectra would improve the number of identifications,^{17,18} we reasoned that quantitation would be compromised owing to the decreased PIP and, therefore, not helpful for the overall interpretation of the experiment.

If more mixed spectra are produced, we hypothesized that this would lead to ratio compression and distortion (i.e., less apparent knockout) of the true ratios for the knocked-out yeast proteins Met6p, His4p, and Ura2p. We calculated the IFI¹¹ for the peptides of the three knocked-out genes in their respective TMT channels. We found that when APD was enabled on the HFX, we consistently observed lower levels of IFI for the three knocked-out proteins His4p (0.63 vs 0.69), Ura2a (0.63 vs 0.70), and Met6p (0.67 vs 0.83) (Figure 3B). The Lumos showed similar trends but to a lesser extent, where His4p (0.69 vs 0.73) and Ura2p (0.65 vs 0.67) had slightly better IFI with APD disabled, though Met6p showed a higher IFI with APD enabled (0.86 vs 0.79). All IFI calculations are shown in Figure S1. Our results suggest that APD has a less adverse effect on Lumos mass spectrometers compared to the HF-X, where APD causes consistent underperformance compared APD-disabled data acquisition.

CONCLUSIONS

Successful differential proteomic experiments strive to balance depth and quantitative precision and accuracy for a large number of peptides over a wide dynamic range. The accuracy of TMT-based quantitation depends on the ability to isolate and fragment “pure” analytes for quantitation. While isobaric labeling with TMT and iTRAQ has been proven to be a powerful method for relative quantitation, its power to determine differential protein abundance can be hampered by coisolating analytes.

Here, we show that utilizing APD for TMT-based data acquisition can be detrimental. Our results indicate that APD takes more MS2 scans, due to its ability to identify more peptide precursors in survey scans; however, the increase in

MS2 scans of TMT-labeled peptide precursors does not translate to more identified peptides. We also found more unassigned product ions in the MS2 spectra when APD was used. These unassigned MS2 peaks likely derive from coeluting, nearly isobaric precursor ions that are coisolated with the primary analyte for MS2. Reporter ions derived from this collection of precursors increases interference (as measured by the IFI) and results in a higher overall level of ratio compression when APD is on. These results are less prominent on the Lumos mass spectrometer and on either instrument, at 10 ng on-column loads. On the basis of the experiments described here, we do not recommend using APD for TMT-based experiments.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.8b00611.

IFI for each replicate of each experiment conducted for this study (PDF)

Unpaired Student's t test P -value is reported in red above if only reported if $P \leq 0.1$. ns indicates not significant. (PDF)

Data acquisition metrics for all experiments in this study (XLSX)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: sammyers@broadinstitute.org. (S.A.M.)

*E-mail: scarr@broad.mit.edu. (S.A.C.)

ORCID

Samuel A. Myers: 0000-0001-7636-0887

Shankha Satpathy: 0000-0002-6040-4854

Author Contributions

N.U., S.A.M., and S.A.C. conceived the study. S.A.M., S.K., and S.S. designed and implemented the research. J.C. and J.R. developed and prepared the yeast standard. R.V. initially characterized the yeast standard. S.A.M. performed the data analyses. K.C. developed the metrics for PIP and SPI; R.V. initially characterized the yeast standard. S.A.M., S.S., S.K., N.U., R.V., and S.A.C. wrote the manuscript.

Notes

The authors declare the following competing financial interest(s): R.V., J.C., and J.C.R. are employees of Thermo Fisher Scientific, developer and distributor of the Orbitrap Fusion / Lumos and Q-Exactive mass spectrometers.

ACKNOWLEDGMENTS

We are thankful to K. Strupat at Thermo Fisher Scientific, Bremen, Germany, for the APD toggle on the HFX. We are also grateful to Joao Paulo for his helpful discussions. S.K. is a Cancer Research Institute/Hearst Foundation fellow. The raw mass spectrometry data can be found on MassIVE with the identifier MSV000083005.

REFERENCES

(1) 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.

(2) Paek, J.; Kalocsay, M.; Staus, D. P.; Wingler, L.; Pascolutti, R.; Paulo, J. A.; Gygi, S. P.; Kruse, A. C. Multidimensional Tracking of GPCR Signaling via Peroxidase-Catalyzed Proximity Labeling. *Cell* **2017**, *169* (2), 338–349.

(3) Myers, S. A.; Wright, J.; Peckner, R.; Kalish, B. T.; Zhang, F.; Carr, S. A. Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-mediated proximity labeling. *Nat. Methods* **2018**, *15*, 437–439.

(4) Mertins, P.; Mani, D. R.; Ruggles, K. V.; Gillette, M. A.; Clauser, K. R.; Wang, P.; Wang, X.; Qiao, J. W.; Cao, S.; Petralia, F.; Kawaler, E.; Mundt, F.; Krug, K.; Tu, Z.; Lei, J. T.; Gatz, M. L.; Wilkerson, M.; Perou, C. M.; Yellapantula, V.; Huang, K. L.; Lin, C.; McLellan, M. D.; Yan, P.; Davies, S. R.; Townsend, R. R.; Skates, S. J.; Wang, J.; Zhang, B.; Kinsinger, C. R.; Mesri, M.; Rodriguez, H.; Ding, L.; Paulovich, A. G.; Fenyo, D.; Ellis, M. J.; Carr, S. A. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **2016**, *534* (7605), 55–62.

(5) O'Connell, J. D.; Paulo, J. A.; O'Brien, J. J.; Gygi, S. P. Proteome-wide evaluation of two common protein quantification methods. *J. Proteome Res.* **2018**, *17* (5), 1934–1942.

(6) Shiran, S. L.; Botting, C. H. A comparison of the accuracy of iTRAQ quantification by nLC-ESI MSMS and nLC-MALDI MSMS methods. *J. Proteomics* **2010**, *73* (7), 1391–403.

(7) 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), 5347–55.

(8) 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), 1885–97.

(9) Savitski, M. M.; Mathieson, T.; Zinn, N.; Sweetman, G.; Doce, C.; Becher, I.; Pachel, F.; Kuster, B.; Bantscheff, M. Measuring and managing ratio compression for accurate iTRAQ/TMT quantification. *J. Proteome Res.* **2013**, *12* (8), 3586–98.

(10) Ting, L.; Rad, R.; Gygi, S. P.; Haas, W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat. Methods* **2011**, *8* (11), 937–40.

(11) Paulo, J. A.; O'Connell, J. D.; Gygi, S. P. A Triple Knockout (TKO) Proteomics Standard for Diagnosing Ion Interference in Isobaric Labeling Experiments. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (10), 1620–5.

(12) 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), 2341–6.

(13) Hebert, A. S.; Thoing, C.; Riley, N. M.; Kwieci, N. W.; Shiskova, E.; Huguet, R.; Cardasis, H. L.; Kuehn, A.; Eliuk, S.; Zabrouskov, V.; Westphall, M. S.; McAlister, G. C.; Coon, J. J. Improved Precursor Characterization for Data-Dependent Mass Spectrometry. *Anal. Chem.* **2018**, *90* (3), 2333–2340.

(14) Rappsilber, J.; Mann, M.; Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2007**, *2* (8), 1896–1906.

(15) McAlister, G. C.; Huttlin, E. L.; Haas, W.; Ting, L.; Jedrychowski, M. P.; Rogers, J. C.; Kuhn, K.; Pike, L.; Grothe, R. A.; Blethrow, J. D.; Gygi, S. P. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* **2012**, *84* (17), 7469–78.

(16) Zecha, J.; Meng, C.; Zolg, D. P.; Samaras, P.; Wilhelm, M.; Kuster, B. Peptide level turnover measurements enable the study of proteoform dynamics. *Mol. Cell. Proteomics* **2018**, *17* (5), 974–992.

(17) Trnka, M. J.; Baker, P. R.; Robinson, P. J.; Burlingame, A. L.; Chalkley, R. J. Matching cross-linked peptide spectra: only as good as the worse identification. *Mol. Cell. Proteomics* **2014**, *13* (2), 420–34.

(18) Zhang, B.; Pirmoradian, M.; Chernobrovkin, A.; Zubarev, R. A. DeMix workflow for efficient identification of cofragmented peptides in high resolution data-dependent tandem mass spectrometry. *Mol. Cell. Proteomics* **2014**, *13* (11), 3211–23.