Optimization of RP-LC-MS Top-down Protein Analysis on an Orbitrap Fusion Lumos Tribrid MS with the Advanced Peak Determination Algorithm

Vlad Zabrouskov[1], Luca Fornelli[2], Ryan Fellers[2], Kristina Srzentic[2], Joshua Silveira[1], Helene Cardasis[1], Graeme C. McAlister[1], Christian Thoeing[3], Derek Bailey[1], Andreas Kuehn[3], Mike Senko[1], Shannon Eliuk[1], Romain Huguet[1], Neil Kelleher[3]
[1]Thermo Fisher Scientific, San Jose, CA; [2]Northwestern University, Evanston, IL; [3]Thermo Fisher Scientific, Bremen, Germany

ABSTRACT

Purpose: Top-down data-dependent analysis workflows are typically plagued by repeated interrogation of different charge states of the same protein precursor. Through the development of an improved precursor charge state determination algorithm, we are able to better associate all the ions representing a single proteoform in real-time. This added information improves the data-dependent decisions we can make during the workflow, which in turn, maximizes the depth of analysis. We demonstrate here the impact of this new algorithm on top-down analysis using a Thermo Scientific™ Orbitrap Fusion™ Lumos™ Tribrid™ mass spectrometer.

Methods: A commercial intact protein standard mixture and E.coli lysate were analyzed by top-down data-dependent analysis using a Thermo Scientific™ Vanquish™ UHPLC and an Orbitrap Fusion Lumos MS. Multiple parameter settings were evaluated using both the traditional precursor detection algorithm (SPD) and the newly developed Advanced Precursor Determination (APD) algorithm.

Results: Data dependent parameters were optimized to minimize redundant sampling of each protein and to improve sampling depth in both the 'low-high' and the 'high-high' modes of operation. The APD algorithm has a significant impact on the system's ability to make intelligent data-dependent decisions during top-down analysis, and represents a major step forward for top-down LC-MS analysis.

INTRODUCTION

Top-down data-dependent analysis has traditionally been complicated by aspects of the data that are inherent of intact protein MS analysis, but not encountered during bottom-up analysis. First, the electrospray process generates a number of charge states for a single protein species, diluting the MS1 signal in any one *m/z* precursor. Additionally, the number of potential sites of fragmentation increases with precursor MW, which further dilutes the signal that can be accumulated in any one given product ion in the MS2 spectra. Spectral summing, gas phase purification by ion-ion chemistry, and 'peak parking' are means of overcoming this challenge, though often at the cost of speed and throughput. A second aspect of intact protein detection in FT-generated MS1 is the inverse relationship between MW and resultant achievable resolution, as well as the removal of preliminary beats during FT signal processing that represent most if not all of the available signal from high MW protein ions. To overcome this, MS1 analysis is typically performed at low resolution (short transient collection) in order to detect, and in turn trigger MS2, on a wider range of MW precursors. Low resolution detection has traditionally precluded charge state assignment in MS1 however, resulting in the inability to apply data-dependent filters such as "one charge state per precursor" to this type of analysis. In turn, different charge states of the same protein are often triggered on sequentially, in an intensity-dependent manner, resulting in redundant sampling of the most intense proteins, and limited depth of analysis. Even when resolution is sufficient to assign the proper charge state, overlapping isotopic envelopes previously presented a challenge to proper charge assignment. Recently, an algorithm was developed that greatly improves upon charge state assignment for both overlapping isotope distributions and non-isotopically resolved peaks that can be clustered into so-called 'charge envelopes'. We demonstrate here the impact of this algorithm, referred to as "Advanced Precursor Determination" (APD), on top-down data-dependent analysis.

MATERIALS AND METHODS

Sample Preparation

Thermo Scientific™ Pierce™ Intact Protein Standard Mix (P/N: A33536) and *E. Coli* Protein Sample (Bio-Rad Laboratories, P/N: 1632110) were used in these analyses. Each vial of the Pierce sample was reconstituted in 100ul of water as indicated by the Pierce instructions. The *E. Coli* Protein Sample was reconstituted to a final concentration of 1ug/µl in solvent A (0.1% formic acid in water).

LC-MS Setup

LC-MS experiments were performed using an Orbitrap Fusion Lumos Tribrid MS with APD (Figure 1) coupled with a Vanquish UHPLC system (Solvent A: Water with 0.1% Formic Acid (v/v); Solvent B: Acetonitrile with 0.1% Formic Acid (v/v)). The Pierce Intact Protein Standard Mix sample was separated with a 10 cm Thermo Scientific™ MAbPac™ RP column (P/N 088647) for 20 minutes at 200 µL/min (Figure 2 a and b). The *E. Coli* Protein Sample was analyzed for 60 min at 200 µL/min (Figure 2 c and d). The optimized 'low-high' and the 'high-high' MS methods are shown in Figure 2 e.

Figure 1. APD is a capability of the Orbitrap Fusion Lumos Tribrid MS for Proteomics.

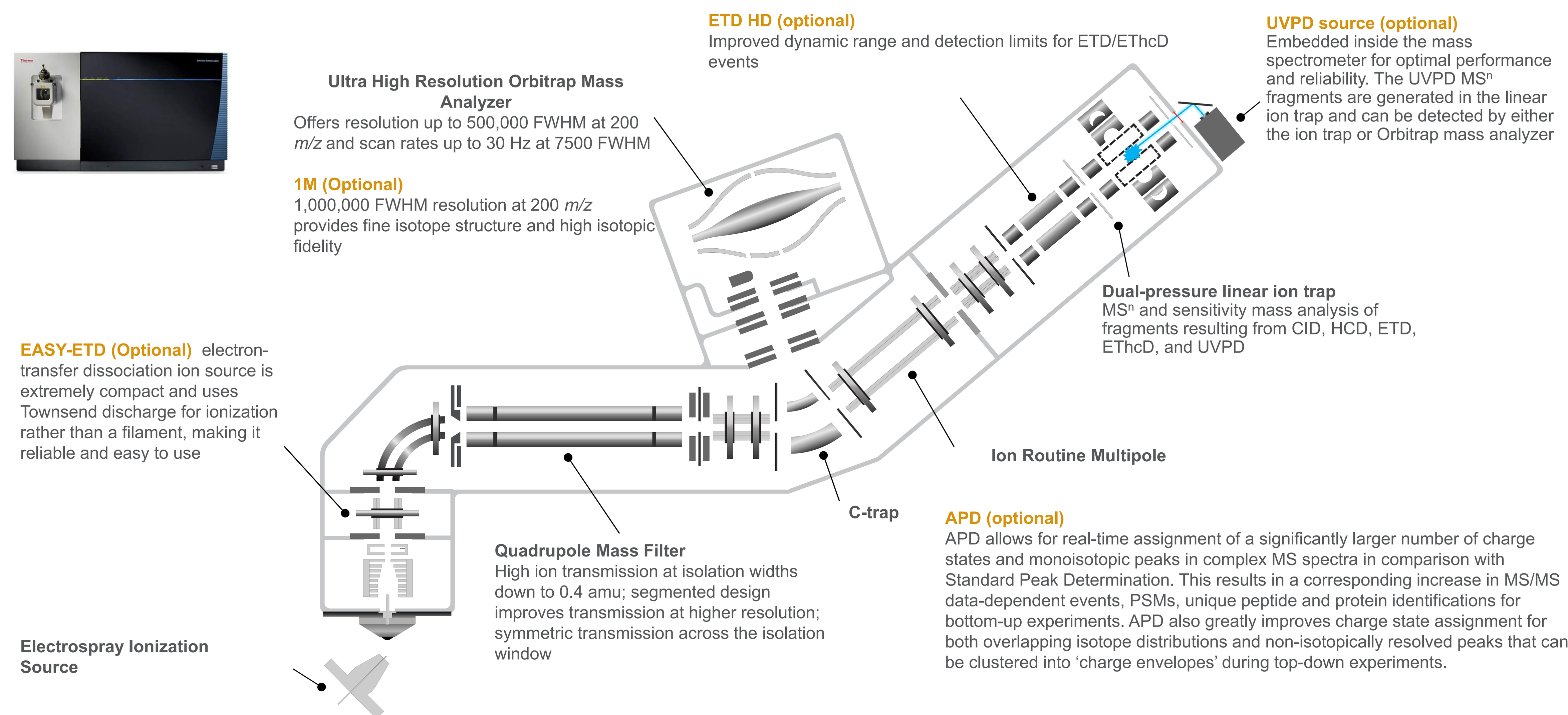
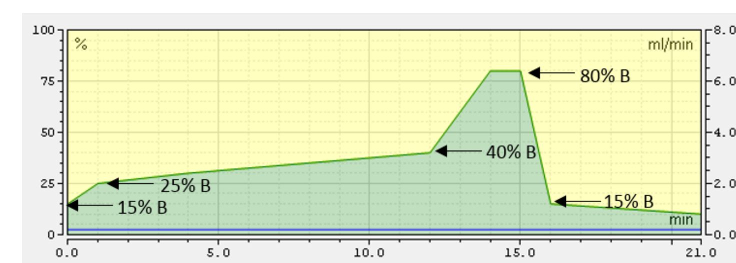
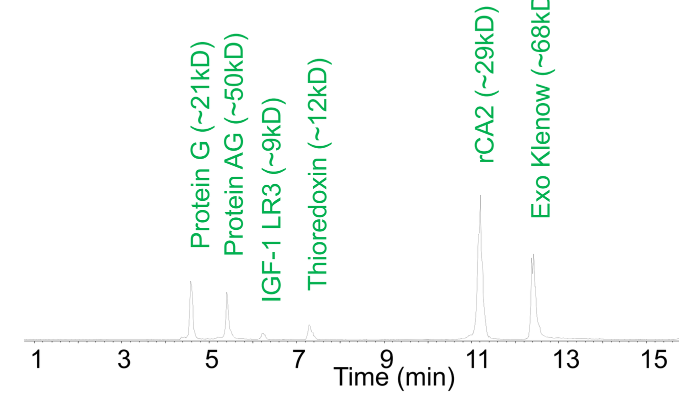


Figure 2. LC-MS experiment methods

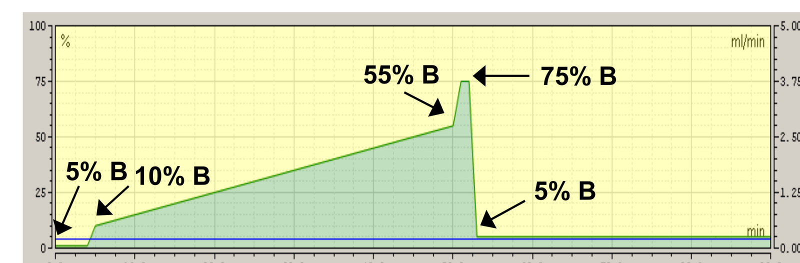
a) Gradient profile used for Intact Protein Standard Mix



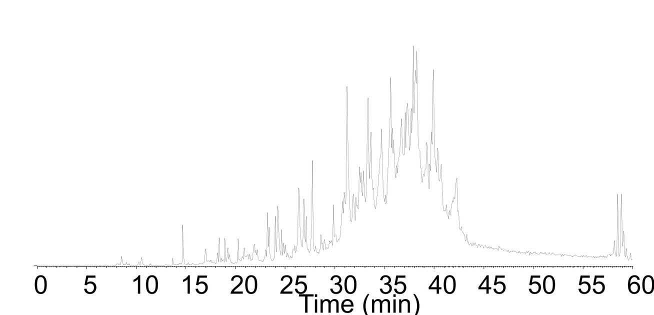
b) Chromatogram for Pierce intact protein standard mix



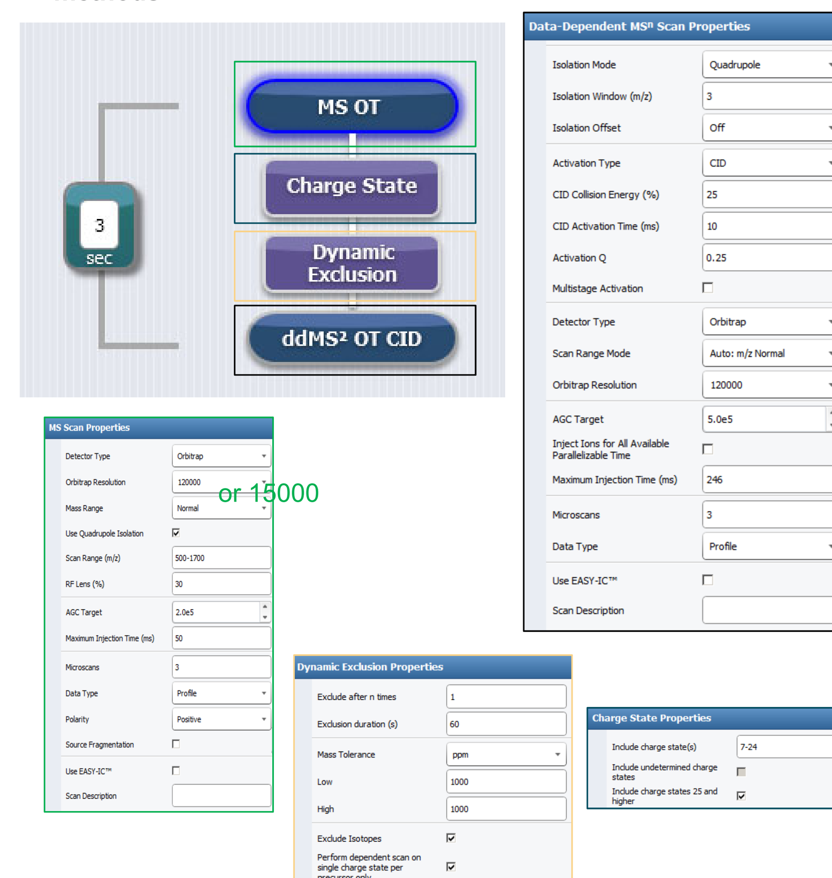
c) Gradient profile used for *E.Coli* Protein Sample



d) Chromatogram for *E. Coli* Protein Sample



e) 'low-high' (MS1 at 15,000 FWHM – MS2 at 120,000 FWHM) MS and 'high-high' (MS1 and MS2 at 120,000 FWHM at *200 m/z*) MS methods



Data Analysis

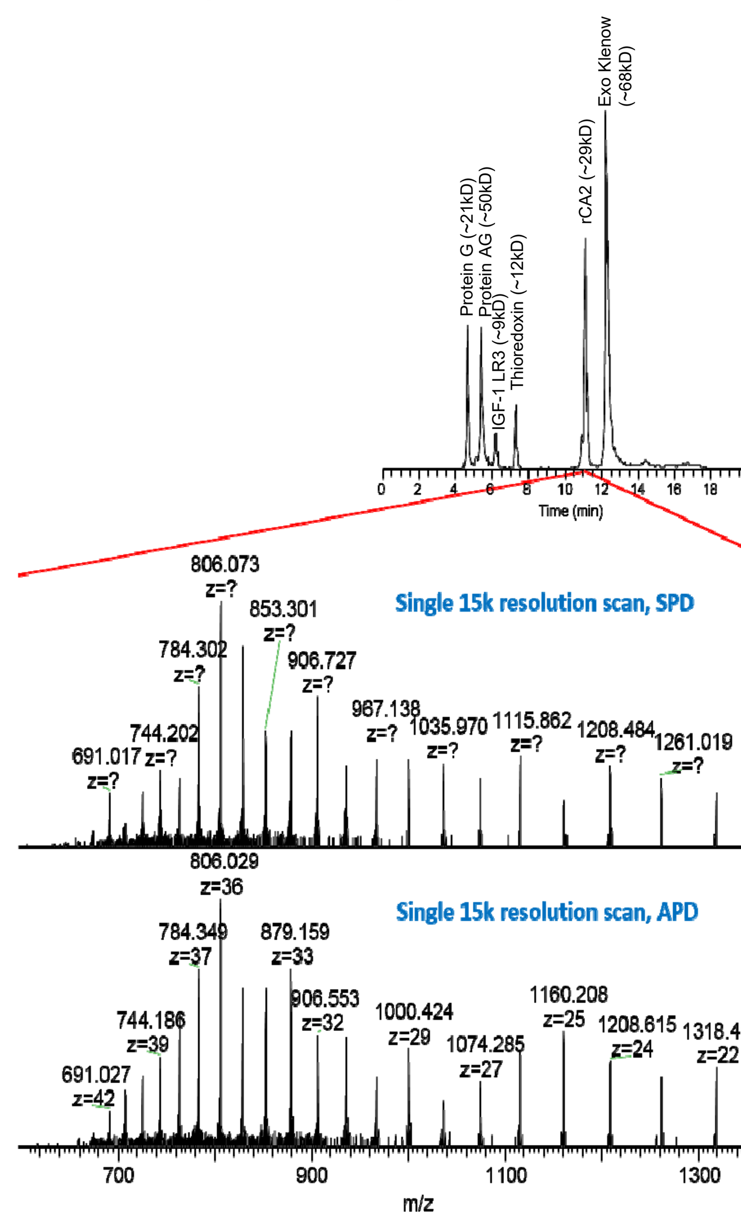
Data were analyzed using Thermo Scientific™ Proteome Discoverer™ 2.2 software using the standard ProSightPD Low-High or High-High Three Tier processing templates and the K12 *E.coli* database (.pcsw) downloaded directly from the Proteinacious database warehouse (http://proteinaceous.net/database-warehouse).

RESULTS

Analysis of Pierce Intact Protein Standard Mixture

Initial studies were done with the Pierce Intact Protein Standard Mixture in order to gauge effectiveness of the APD algorithm across a wide range of protein MWs. The Pierce Intact Protein Standard Mixture contains 6 proteins ranging in size from 9kD to 68kD.

Figure 3. Charge states are accurately assigned to the entire charge state envelope of carbonic anhydrase in a single low resolution scan with APD on.



Pierce intact protein standard mix was run here using a short 20 minute gradient to separate the 6 proteins. The same method was run with and without APD enabled. The top panels shows a typical chromatogram of this mixture, whereas the middle panel shows a single 15,000 resolution scan of carbonic anhydrase collected with APD off (standard settings). The bottom panel shows a spectra of the same resolution setting and same retention time, but collected with APD turned on. All charge states are properly assigned in the "APD on" spectra, despite being collected at a resolution setting that does now allow for isotopic resolution of each charge state. This should enable the usage of more advanced filters during data-dependent acquisition in so called "low-high" mode of operation (low resolution full scan, high resolution MS2 scan).

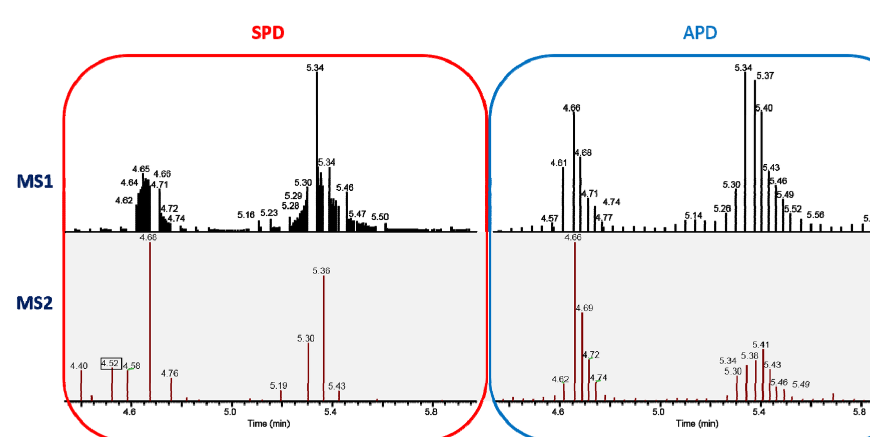


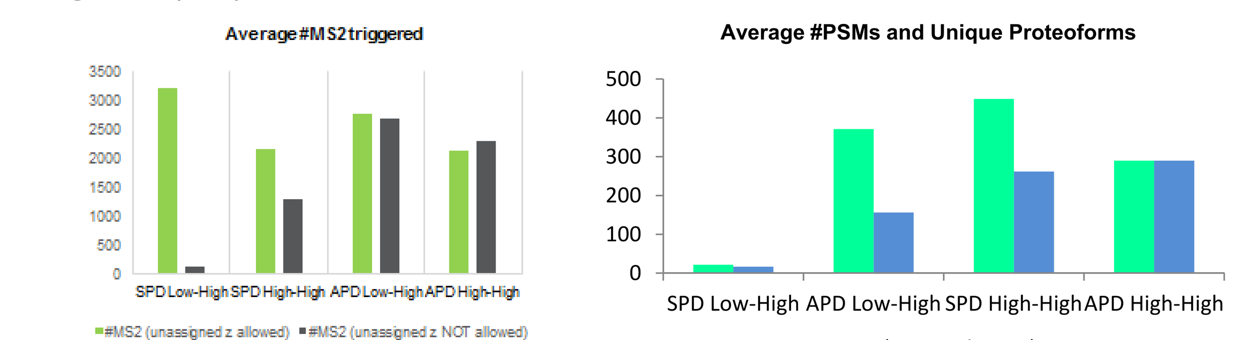
Figure 4. Data dependent acquisition is now possible from low resolution full scan data without "allow unassigned charge states" selected.

Since legacy algorithms did not assign intact protein charge states in real time from low resolution full scans, the option to "allow unassigned charge states" to trigger MS2 was necessary in order to collect any valuable MS2 data. Unfortunately, this option disables other data-dependent filtering and prioritization options (charge state range, 'one charge state per precursor', MW range, etc.), and allows triggering on low mass, low charge chemical noise peaks that are of no interest. In Figure 4, we show the effect of turning "allow unassigned charge states" off in a standard 'low-high' data dependent analysis of Pierce Intact Protein Standard Mix. We are zoomed in here on only two eluting proteins. Each line here represents a scan triggered. One can see that the distribution of MS1 (top sub-panel) to MS2 (bottom sub-panel) scans in unbalanced, with MS2 only being triggered occasionally. Manual analysis of these MS2 scans indicate that the precursors were not the eluting proteins themselves, but rather low intensity, low charge peptide fragments (data not shown). Conversely, data collected using the same method with "allow unassigned charge states" off but APD enabled are shown in the right hand panel. Here one can see an even distribution between MS1 and MS2 scans, indicating that charge states are being properly assigned in the low resolution MS1 full scan, and MS2 are triggered on these major protein peaks.

Analysis of *E.coli* lysate

In order to understand the impacts of different data-dependent filters on MS2 triggering dynamics with APD on, we moved to analysis of a more complex sample, *E.coli* lysate, in both "low-high" (low resolution full scan, high resolution MS2 scan) and "high-high" (high resolution full scan, high resolution MS2 scan) modes of analysis.

Figure 5. With APD enabled, MS2 triggering is not hindered by the addition of charge state filters in either Low-High or High-High modes of operation, as it does when the legacy algorithm (SPD) is used.



Here we can see that with the legacy algorithm (SPD), disallowing precursors with unassigned charge states significantly hinders triggering in both Low-High and High-High modes of operation. With APD on, the number of MS2 triggered are preserved. This is interesting especially when comparing the high-high analysis. Here the APD algorithm is clearly improving the number of charge assigned precursors when even resolution is high enough to otherwise provide isotopic resolution of each independent precursor. APD with charge state filters (unassigned z in not allowed and only a single charge state per precursor (1zpp) can be selected for MS2), provided near 10 times more unique proteoforms identification over SPD in 'Low-High' experiments and over 10% more in 'High-High' experiments.

CONCLUSIONS

APD greatly improves the charge state assignment of both isotopically and non-isotopically resolved precursors that can be clustered into charge envelopes. Therefore, this new feature with the addition of charge state filters allows optimal data dependent acquisition of simple and complex samples by collecting only valuable MS2 data in both 'Low-High' and 'High-High' modes of operation for unmatched LC-MS top-down performances.

TRADEMARKS/LICENSING

© 2018 Thermo Fisher Scientific Inc. All rights reserved. ProSightPD is a trademark of Proteinaceous, Inc. All other trademarks are the property of Thermo Fisher Scientific and its subsidiaries. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.

PO65304-EN 0518S

ThermoFisher SCIENTIFIC