Improvements in Protein Identifications in Data-Dependent Acquisition Experiments by Use of Novel Precursor Selection and Fragmentation Strategies

ASMS 2010

Christine Miller, Jose Meza, Norton Kitagawa, Wilfred Tang, Joseph Roark, Javier Satulovsky, Patrick D. Perkins; Agilent Technologies, Inc., Santa Clara, CA USA

Improvements in Protein Identifications in Data-Dependent Acquisition Experiments by Use of Novel Precursor Selection and Fragmentation Strategies

Introduction

Efficient and comprehensive protein identification depends on several conditions during the acquisition, among them: the number of precursors examined per unit time, selection of the most promising precursors for fragmentation, and application of the appropriate fragmentation conditions to yield the highest quality product ion spectrum during acquisition. This work describes changes to the precursor selection and fragmentation steps to select the precursors most likely to produce good quality peptide MS/MS spectra, and to increase the quality of those spectra.

Experimental

Instrument

The instrument consisted of an Agilent Technologies LC/MSD system with 1200 Series nanoflow and capillary HPLC pumps, microdegassers, micro wellplate autosampler with thermostat, HPLC-Chip/MS interface, and 6530 QTOF mass spectrometer.

LC/MS Conditions (shorter gradient)

*Column: Zorbax 300SB-C18, 0.075 x 150 mm Chip, 40 nL enrichment column, 5 µm particle size (Agilent Technologies part #G4240-62006); Mobile phase: A = 0.1% formic acid, B 0.1% formic acid in 90% acetonitrile; Nanoflow gradient (%B): 3% at 0 min, 5% at 0.1 min, 8% at 2 min, 35% at 40 min, 40% at 47 min, 90% at 47.1 min, 90% at 51 min, 3% at 51.1 min; Stop Time: 60 minutes; Nanopump flow rate: 300 nL/min; Capillary pump gradient: 3% B isocratic; Capillary pump flow: 2.5 µL/min; Chip valve position: enrichment at 51.2 min; Ionization mode: positive electrospray; Capillary voltage: various but typically -1850 V; Drying gas flow: 5 L/min; Drying gas temperature: 325 °C; Fragmentor: 175 V; Skimmer: 65 V; Acq. Mode: autoMS/MS; Max. precursors/cycle: 10; Scan range: 275-1700 m/z (MS), 50-1700 m/z (MS/MS); Acq. rate: 8 Hz (MS), variable (MS/MS), see results; Isolation width (MS/MS): medium (4 m/z); Collision energy: energy (V) = -4.8 + 3.6 * (precursor m/z/100); Active exclusion: on, exclude after one spectrum, release after 0.5 min; Charge state preference: 2, 3 >3, and unknown, sorted by abundance only; Reference mass: 1221.990637 m/z from continuous addition of trace amounts of HP-1221 calibrant into the ionization region.*

LC/MS Conditions (longer gradient)

*Column: Zorbax 300SB-C18, 0.075 x 150 mm Chip, 160 nL enrichment column, 5 µm particle size (Agilent Technologies part #G4240-62010); Mobile phase: A = 0.1% formic acid, B 0.1% formic acid in 90% acetonitrile; Nanoflow gradient (%B): 3% at 0 min, 5% at 0.1 min, 8% at 2 min, 35% at 60 min, 40% at 67 min, 90% at 67.1 min, 90% at 71 min, 3% at 71.1 min; Stop Time: 91 minutes; Nanopump flow rate: 300 nL/min; Capillary pump gradient: 3% B isocratic; Capillary pump flow: 2.5 µL/min; Chip valve position: enrichment at 71.2 min; Ionization mode: positive electrospray; Capillary voltage: various but typically -1850V; Drying gas flow: 5 L/min; Drying gas temperature: 325 °C; Fragmentor: 175 V; Skimmer: 65 V; Acq. Mode: autoMS/MS; Max. precursors/cycle: 10; Scan range: 275-1700 m/z (MS), 50-1700 m/z (MS/MS); Acq. rate: 8 Hz (MS), variable (MS/MS), see results; Isolation width (MS/MS): medium (4 m/z); Collision energy: energy (V) = -4.8 + 3.6 * (precursor m/z/100); Active exclusion: on, exclude after one spectrum, release after 1 min; Charge state preference: 2, 3 and >3, sorted by abundance only; Reference mass: 1221.990637 m/z from continuous addition of trace amounts of HP-1221 calibrant into the ionization region.*

Software

A custom pre-release version of MassHunter Acquisition B.04.00 was used for instrument control and real-time data-dependent decisions. Protein/peptide identifications were done by database searching using a prerelease version of Spectrum Mill MS Proteomics Workbench B.04.00 with a global false discovery rate of 1%. The Swiss-Prot database was used, specifying the E. coli K-12 proteome.

Sample preparation

E. coli protein lysate (BioRad Laboratories, Hercules, CA) was denatured in 50% 2,2,2-trifluoroethanol/50% 50 mM ammonium bicarbonate, reduced with DTT, alkylated with iodoacetamide, and digested with trypsin 50/1 (w/w) in 50 mM ammonium bicarbonate (37 °C, overnight) [ref]. The sample was cleaned up by spin filtration through a 10 kDa MWCO filter, then diluted to the final concentration in 6% acetonitrile/0.1% formic acid. Injection volumes were adjusted so that typically 500 ng was loaded on-column.

Improvements in Protein Identifications in Data-Dependent Acquisition Experiments by Use of Novel Precursor Selection and Fragmentation Strategies

Results and Discussion

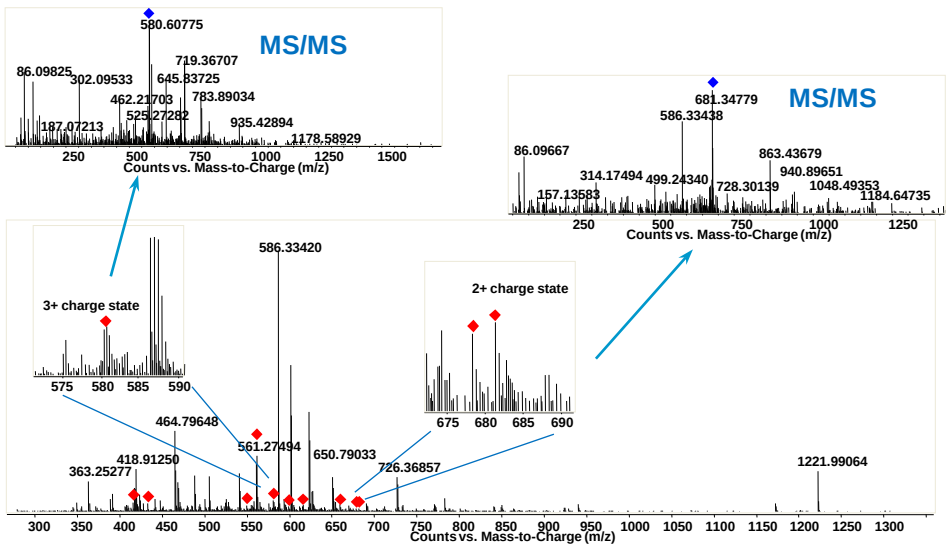
Areas investigated

The areas investigated for improvement fell into three categories:

- *Better recognition of peptide isotope patterns*
- *Better selection of precursors to fragment*
- *Optimum duration for collection of higher quality product ion spectra*

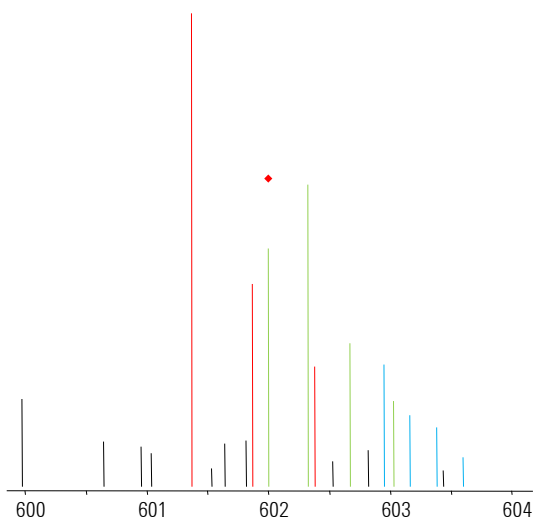
Better recognition of peptide isotope patterns

Better detection of peptides in the sample was achieved by optimization of the abundance thresholds and abundance ratios in the code, in combination with using a peptidic model for the isotope pattern. As a result, charge states are properly assigned, and the monoisotopic peak is preferentially selected, allowing smarter precursor selection.



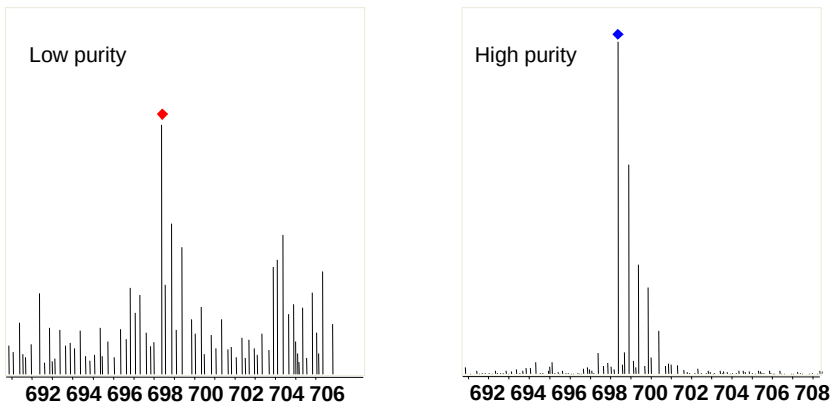
Better selection of precursors to fragment

During analysis of complex samples, it is common to isolate more than one precursor in the transmission window of the quadrupole. If these precursors are two (or more) peptides, then it is likely a mixed product ion spectrum will result a chimeric peptide ID. Most database searches do not fare well with these mixed spectra.

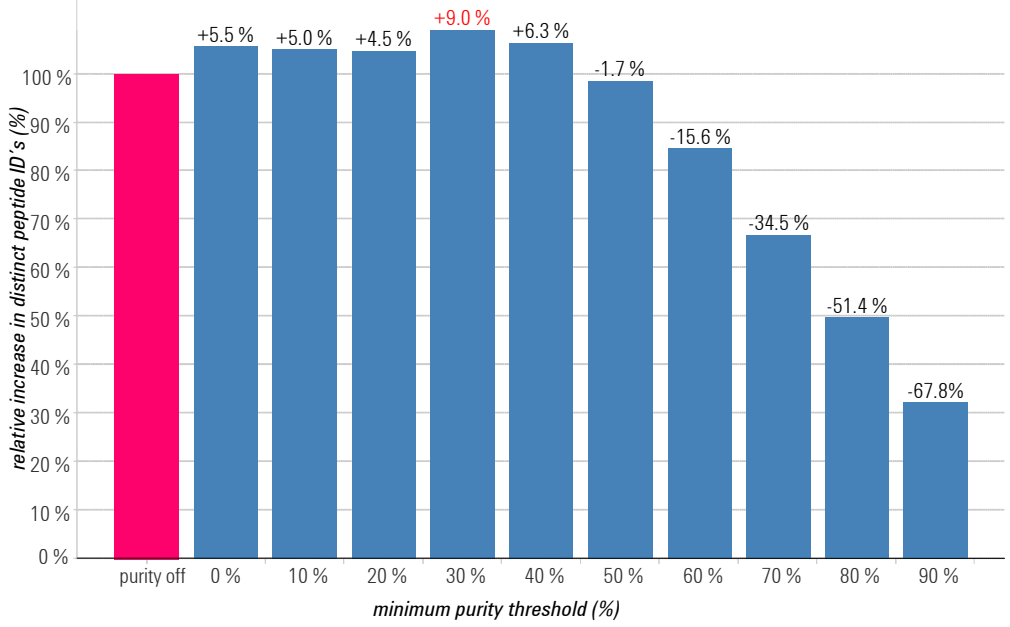


In the example to the left, the precursor in green is the intended target for MS2 analysis. However, the dominant peptides ([M+2H] to the left in red and [M+4H] to the right, in red) confound the identity of the targeted peptide

The new precursor selection process includes a "precursor purity" routine which assigns a higher priority to precursor candidates which are a high percentage of the total abundance in the isolation window, along with a parameter for rejecting precursors below a threshold purity value. This isotopic purity rejection threshold is user-adjustable.



The effect of the isotopic purity rejection threshold on peptide identifications is shown below. Higher values for this parameter increase the stringency, filtering out precursor candidates being isolated simultaneously with greater amounts of other signals. At low values for this parameter, the routine exhibits the desired effect of selecting "cleaner" precursor candidates for fragmentation, and the number of identifications increases. At higher values of this parameter, precursor candidates which would have given good identifications are eliminated from consideration for fragmentation, and the total number of identifications falls. The balance point was found to be at a setting of about 30% for this parameter.

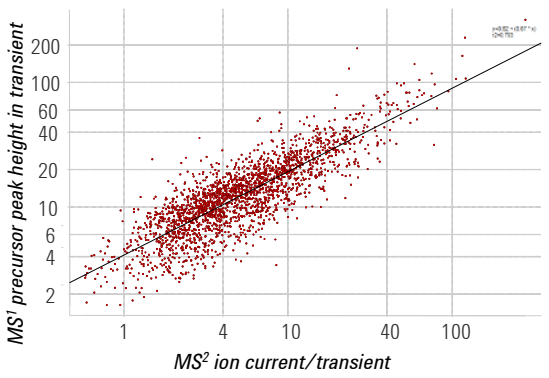


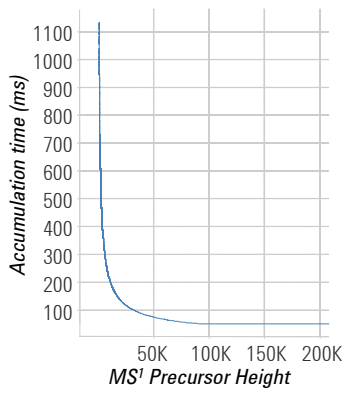
Improvements in Protein Identifications in Data-Dependent Acquisition Experiments by Use of Novel Precursor Selection and Fragmentation Strategies

Results and Discussion

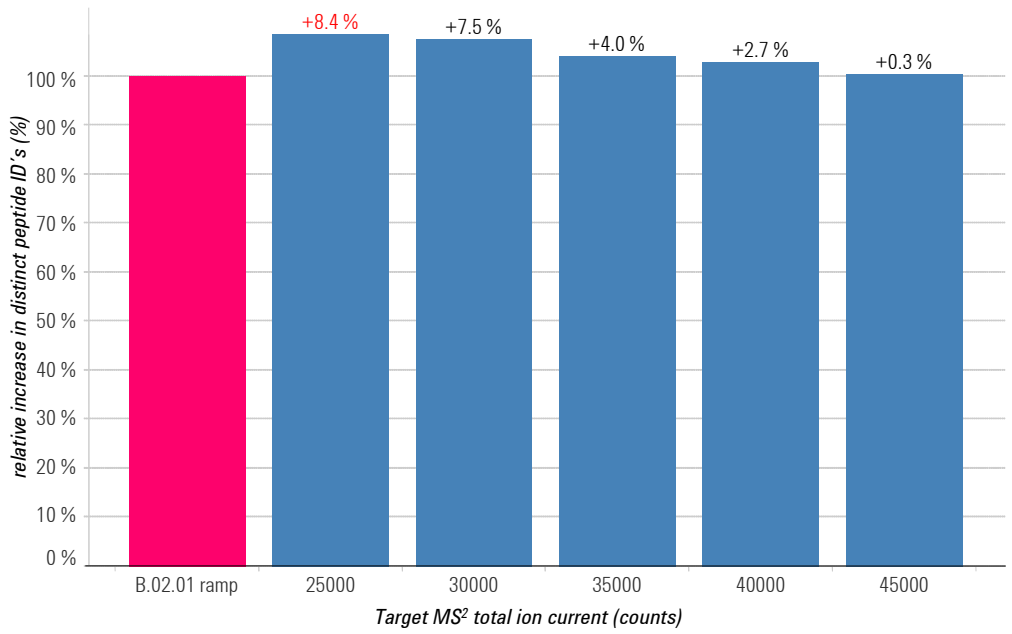
Optimum duration for collection of high quality product ion spectra

The currently available acquisition software uses a "variable number of transients" feature to increase the acquisition time of the product ion spectrum for precursors with low abundance. The accumulation time is determined by a linear relationship to MS^1 precursor abundance. A precursor with two-fold lower abundance in the MS-only scan is thus fragmented for twice the duration to create the product ion spectrum. Upon fragmentation of lower abundance precursors, an even longer accumulation time appeared to be needed to yield a usable product ion spectrum, so a power function relationship between MS^1 precursor abundance and MS^2 accumulation time was derived.



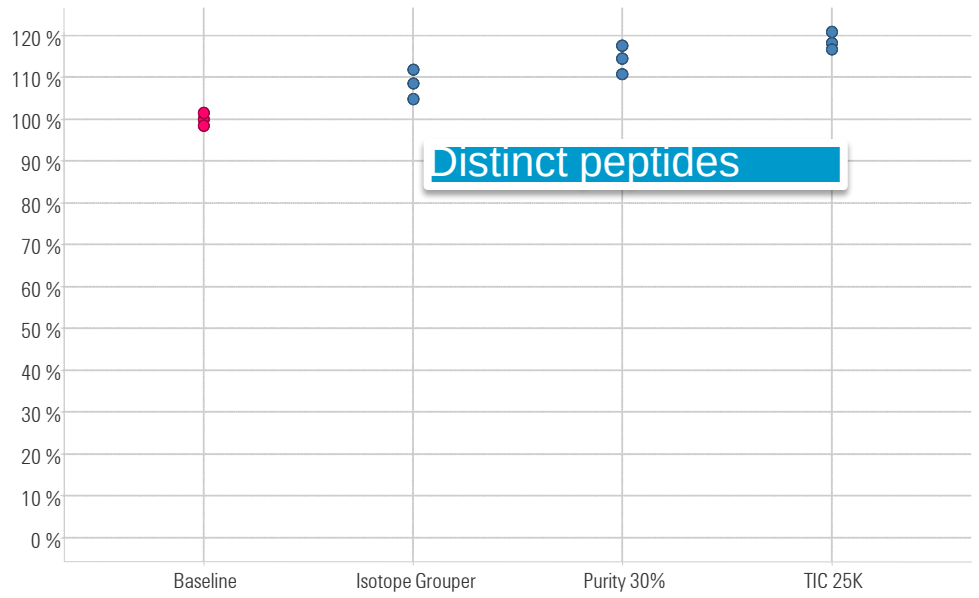


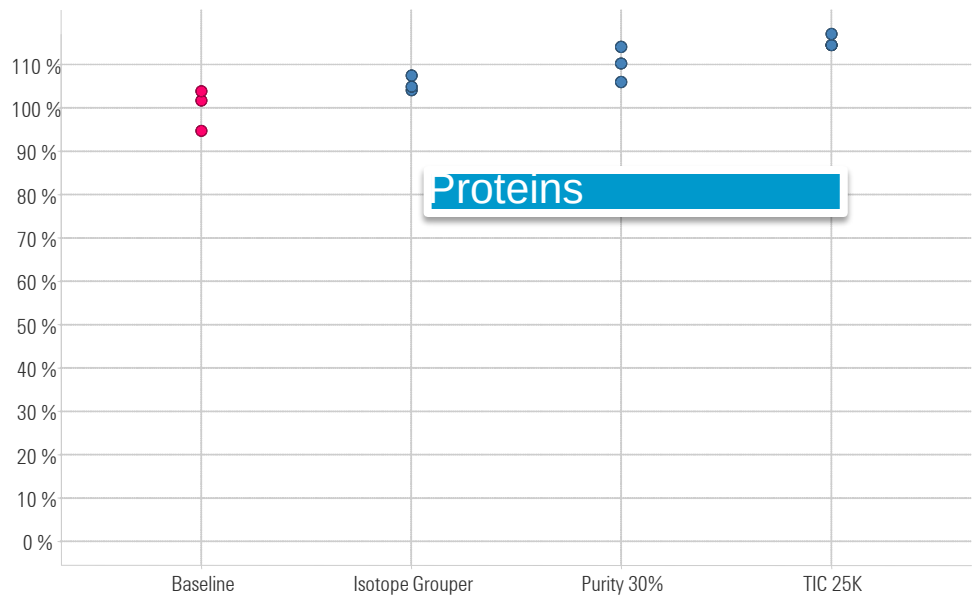
Similar to the optimization of isotopic purity mentioned previously, a tradeoff situation exists for the variable MS^2 acquisition time. A longer duration of MS^2 acquisition increases the quality and identification scores for low-abundance peptides, but it also decreases the number of precursors examined per unit time. Experiments showed that a target abundance value of 25,000 for the MS^2 TIC achieved a balance between these two aims. This parameter is adjustable by the user.



Cumulative improvement

Combining all the improvements resulted in a 20% increase in peptide and 15% increase in protein identifications.





Conclusions

Examination of a number of interdependent variables in the acquisition software has led to significant gains in the number of peptide and proteins identified in data-dependent experiments. This investigation is ongoing.

Reference

1. *Meza, J. E., Miller, C. A., Fischer, S. M. "Improved Tryptic Digestion of Proteins Using 2,2,2-Trifluoroethanol (TFE)", ABRF conference 2004 poster.*