Performance Evaluation of the Q Exactive HF Hybrid Quadrupole-Orbitrap Mass Spectrometer for High-Throughput Top-Down Proteomics

Eugen Damoc,[1] Ping Yip,[2] Leena Valmu,[3] Alexander Cherkassky,[2] Bernard Delanghe,[1] Eduard Denisov,[1] Helene Cardasis,[2] Jason Neil,[2] Alexander Makarov,[1] Jim Stephenson[2]
[1]Thermo Fisher Scientific, Bremen, Germany; [2]Thermo Fisher Scientific, Cambridge, MA, USA;
[3]Thermo Fisher Scientific, Vantaa, Finland

Poster Note 64455

Overview

Purpose: Evaluation of the Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap mass spectrometer for high-throughput top-down proteomics.

Methods: Top-down analysis of an *Escherichia coli* extract using the data dependent "TopN" method with and without chromatographic separation.

Results: We demonstrate utility and applicability of the Q Exactive HF mass spectrometer to perform high-throughput top-down proteome analysis.

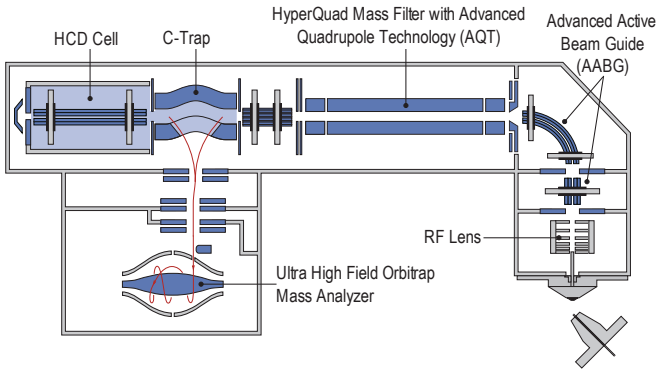
Introduction

Major goals in every top-down proteomics experiment are protein identification and characterization. The strategy used to achieve these goals involves high-resolution mass measurement of intact protein ions followed by their fragmentation and analysis in the mass spectrometer. In spite of enormous improvements in terms of speed and sensitivity in FTMS instrumentation over the last few years, top-down LC-MS/MS in large scale proteome analyses will further benefit if high resolution analysis at higher detection speed would be possible. Furthermore, improvement to the current generation of charge assignment and protein deconvolution algorithms to handle complex top-down data will lead to more efficient, complete, and accurate protein identification. Here we demonstrate the improved performance of the Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer in a series of high-throughput top-down proteomics experiments in conjunction with a new algorithm for charge assignment and protein deconvolution. Furthermore, a multiplex SIM approach to isotopically resolve multiple charge states of proteins up to 50 kDa at LC timescale is presented.

Methods

Direct infusion experiments using intact carbonic anhydrase II were carried out to evaluate the ability of the Q Exactive HF instrument to perform top-down analysis. Also, top-down microbial proteome analysis was performed by LC-MS/MS or direct static nanospray utilizing an *E. coli* extract. 1–2 µg of protein sample was loaded onto a Thermo Scientific™ PepSwift™ Monolithic PS-DVB (200 µm × 25 cm) EASY-Spray™ column, and four different LC gradients (5, 15, 30, and 60 min) were run on a Thermo Scientific™ EASY-nLC™ 1000 system. A data-dependent "Top-N" method using the "high-high" approach was employed to deliver high resolution and high mass accuracy in both MS and MS/MS modes, using the Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer. Proteoforms were identified using a new charge assignment and protein deconvolution algorithm. Furthermore, the high-throughput top-down proteomics data was analyzed using Thermo Scientific™ ProSightPC 3.0 software. Multiplexed SIM experiments were performed using LC/MS analysis of intact enolase.

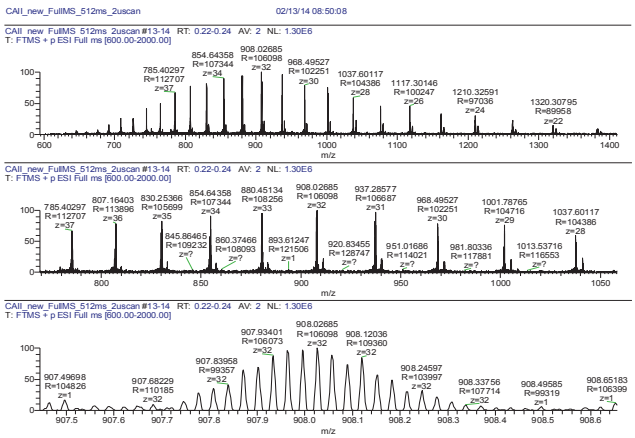
FIGURE 1. The Q Exactive HF instrument layout.



Results

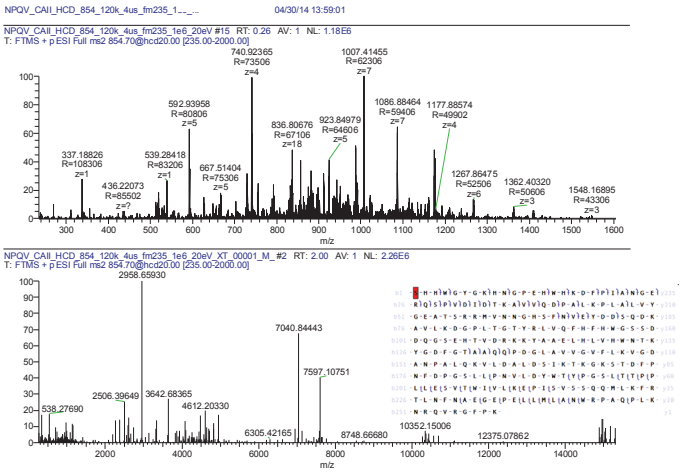
With the implementation of the compact ultra-high field Thermo Scientific™ Orbitrap™ analyzer on the Q Exactive HF instrument (see Figure 1), the resolving power has been increased by 1.8 fold over that of the previous Orbitrap detector. This enables high-resolution analysis at high detection speed which makes the HF instrument more suitable for top-down analysis at LC time scale. The novel Intact Protein Mode allows adjustment of the trapping gas pressure and optimizes the control logic of the instrument to analyze intact proteins with masses up to 50 kDa with isotopic resolution. Carbonic anhydrase II with a molecular mass of 29 kDa was used to evaluate the ability of the Q Exactive HF instrument to perform top-down analysis. Figure 2 shows results of an experiment, where full MS scans were recorded at a resolving power setting of 240,000 (FWHM at *m/z* 200) and AGC target value of 3e6. The figure shows an averaged spectrum over 2 seconds, where the isotopes are baseline resolved and the charge states are properly assigned. Figure 3 is retrieved from a second experiment, where an MS/MS scan with higher-energy collisional dissociation (HCD) of the charge state 34+ at a collision energy of 20 eV was performed. The AGC target value was 1e6 at a resolving power setting of 120,000 (FWHM at *m/z* 200) with 4 µscans in 1 second acquisition time. 36 *b*-type and 28 *y*-type fragment ions were identified using ProSight PC 3.0 software.

FIGURE 2. Full-MS spectrum of intact carbonic anhydrase II (2 × 2 µscans @ 240k res. pwr. → acq. time: 2 seconds) with baseline resolution of the isotopic pattern.



Thermo SCIENTIFIC

FIGURE 3. Top: HCD fragmentation spectrum of carbonic anhydrase II (4 µscans @ 120k res. pwr. → acq. time: 1 second). Bottom: Deconvoluted HCD spectrum and ProSight PC results.



The increased performance in high-throughput top-down proteomics experiments was evaluated using a complex *E. coli* protein extract. Intact proteins from *E. coli* were purified using solid-phase extraction and analyzed by direct static nanospray or LC-MS/MS. Without chromatographic separation, 66 unique proteoforms (Figure 4), could be unambiguously identified in less than 2 minutes' acquisition time by using a "Top-N" "high-high" method. ProsightPC analysis results are shown in Table 1. In this case, an *E. coli* sample solution of 1 µg/µl was directly infused at a flow rate of about 140 nl/min. Furthermore, with chromatographic separation, we demonstrate that the number of proteoforms identified grow linearly with LC gradient duration (Figures 5a and 5b). For 5, 15, 30, and 60 min gradients we were able to identify 722, 996, 1395, and 1964 proteoforms, respectively. ProsightPC analysis was carried out for each LC data set. Figure 6 shows top-down identification of glutamine-binding periplasmic protein from the *E. coli* extract separated by using a 5 min LC gradient.

FIGURE 4. Full MS spectrum of the SPE C4 purified *E. coli* sample, obtained by averaging eighty microscans in direct static nanospray mode.

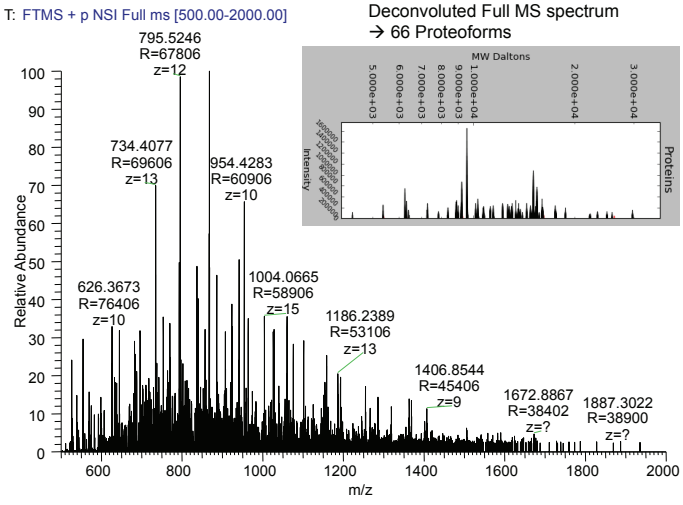
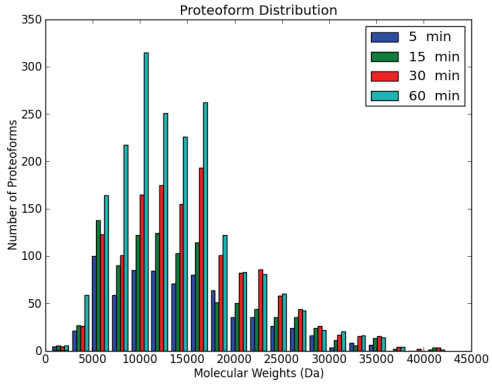


TABLE 1. List of top 30 proteins by E-value identified in the C4 purified *E. coli* sample using the direct static nanospray and data dependent "TopN" method.

	E Value	Number	Observed	Mass Diff Da	Mass Diff ppm	Protein Description	P Score	B-ions	Y-ions
2	2.74E-68	67	9529.229	0.0400253	4.20027	>sp|P0ACF1|DBHA_ECOL6 DNA-binding protein HU-alpha OS=Escherichia coli O6:H	1.32E-74	33	34
3	5.66E-30	39	15037.93	14.0854	936.659	>sp|P68240|HBA_SHEEP Hemoglobin subunit alpha-1/2 OS=Ovis aries PE=1 SV=2	2.72E-36	12	27
4	3.37E-29	31	6311.393	0.00167531	0.265442	>sp|P0A7N4|RL32_ECOU 50S ribosomal protein L32 OS=Escherichia coli (strain K12	1.62E-35	11	20
5	6.92E-29	36	9219.999	0.00415531	0.450684	>sp|P0ACF5|DBHB_ECOL6 DNA-binding protein HU-beta OS=Escherichia coli O6:H1	3.33E-35	20	16
6	2.69E-27	31	10292.67	-0.00185469	-0.180195	>sp|P0A7U4|RS19_ECOL6 30S ribosomal protein S19 OS=Escherichia coli O6:H1 (str	1.29E-33	16	15
7	2.05E-25	32	12217.78	-0.990275	-81.0519	>sp|P61176|RL22_ECOL6 50S ribosomal protein L22 OS=Escherichia coli O6:H1 (stra	9.84E-32	13	19
8	2.58E-21	34	15399.01	0.0115053	0.752341	>sp|P0ACF9|HNS_ECOL6 DNA-binding protein H-NS OS=Escherichia coli O6:H1 (str	1.24E-27	16	18
9	1.07E-20	27	8869.033	0.00040531	0.0456953	>sp|P0A7M9|RL29_ECOL6 50S ribosomal protein L29 OS=Escherichia coli O6:H1 (str	5.16E-27	12	15
10	1.13E-16	25	12960.24	0.00435531	0.336052	>sp|P0A7T0|RS13_ECOL6 30S ribosomal protein S13 OS=Escherichia coli O6:H1 (str	5.45E-23	14	11
11	2.8E-14	23	11208.3	0.0163153	1.45565	>sp|Q8FD03|RL24_ECOL6 50S ribosomal protein L24 OS=Escherichia coli O6:H1 (str	1.35E-20	10	13
12	1.82E-13	26	9184.949	-0.00084469	-0.0919646	>sp|P0A7T4|RS16_ECOL6 30S ribosomal protein S16 OS=Escherichia coli O6:H1 (str	8.77E-20	11	15
13	3.64E-13	24	10685.77	0.0105653	0.988727	>sp|Q8FFS1|RL25_ECOL6 50S ribosomal protein L25 OS=Escherichia coli O6:H1 (stra	1.75E-19	5	19
14	1.83E-11	24	14715.98	0.00002531	0.0017199	>sp|P0A7X4|RS9_ECOL6 30S ribosomal protein S9 OS=Escherichia coli O6:H1 (strain	9.1E-18	9	15
15	1.04E-10	20	11192.23	0.00231531	0.206868	>sp|P0AD21|RL23_ECOL6 50S ribosomal protein L23 OS=Escherichia coli O6:H1 (str	5.01E-17	19	1
16	3.09E-09	20	6250.595	28.0314	4484.59	>sp|A9R676|RL33_YERPG 50S ribosomal protein L33 OS=Yersinia pestis bv. Antiqua	1.48E-15	3	17
17	3.71E-09	20	11442.23	0.00738531	0.645443	>sp|P0AG59|RS14_ECOU 30S ribosomal protein S14 OS=Escherichia coli (strain K12	1.78E-15	10	10
18	1.38E-07	18	10130.47	-0.992145	-97.9367	>sp|P0AD24|RS15_ECOU 30S ribosomal protein S15 OS=Escherichia coli (strain K12	6.62E-14	3	15
19	5.34E-07	16	8363.708	0.0224553	2.68485	>sp|P68680|RS21_ECOL6 30S ribosomal protein S21 OS=Escherichia coli O6:H1 (stra	2.57E-13	6	10
20	2.38E-06	18	14956.31	-0.999505	-66.8283	>sp|P66071|RL15_ECOL6 50S ribosomal protein L15 OS=Escherichia coli O6:H1 (stra	1.15E-12	5	13
21	2.87E-06	18	9529.179	-18.1009	-1899.53	>sp|P0A7U8|RS20_ECOL6 30S ribosomal protein S20 OS=Escherichia coli O6:H1 (str	1.38E-12	13	5
22	4.99E-06	18	8868.83	19.0239	2145.03	>sp|P0A7T8|RS18_ECOL6 30S ribosomal protein S18 OS=Escherichia coli O6:H1 (str	2.4E-12	2	16
23	5.18E-06	17	13641.45	-714.204	-52355.4	>sp|P0AG45|RL17_ECOL6 50S ribosomal protein L17 OS=Escherichia coli O6:H1 (str	2.49E-12	13	4
24	5.4E-06	13	5377.075	131.039	24370	>sp|P0A7P6|RL34_ECOL6 50S ribosomal protein L34 OS=Escherichia coli O6:H1 (str	2.6E-12	5	8
25	6.67E-06	15	11192.23	-996.493	-89034.4	>sp|A6TEW7|RL22_KLEP7 50S ribosomal protein L22 OS=Klebsiella pneumoniae su	3.21E-12	12	3
26	8.1E-06	16	14715.98	0.00072531	0.0492872	>sp|P0A7X4|RS9_ECOL6 30S ribosomal protein S9 OS=Escherichia coli O6:H1 (strain	3.9E-12	7	9
27	1.43E-05	18	9547.286	-19.9567	-2090.31	>sp|P0AG64|RS17_ECOL6 30S ribosomal protein S17 OS=Escherichia coli O6:H1 (str	6.9E-12	11	7
28	0.000257	14	15316.43	20.6845	1350.48	>sp|Q8FF53|NDK_ECOL6 Nucleoside diphosphate kinase OS=Escherichia coli O6:H	1.24E-10	3	11
29	0.000303	14	7269.001	-0.0181347	-2.4948	>sp|B4F1J2|RL29_PROMH 50S ribosomal protein L29 OS=Proteus mirabilis (strain H	1.46E-10	6	8
30	0.000345	14	16063.32	0.00994531	0.619132	>sp|P02075|HBB_SHEEP Hemoglobin subunit beta OS=Ovis aries GN=HBB PE=1 SV=	1.66E-10	2	12

FIGURE 5. Proteoform MW distributions (A) and cumulative distributions (B) for 5, 15, 30, and 60 min LC gradients.



LC Gradient	5 min	15 min	30 min	60 min
Total Proteoforms	722	996	1395	1964

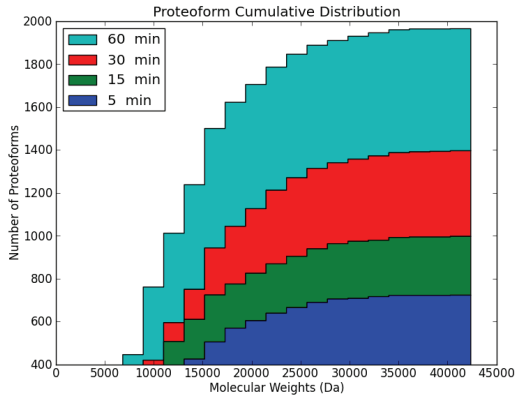
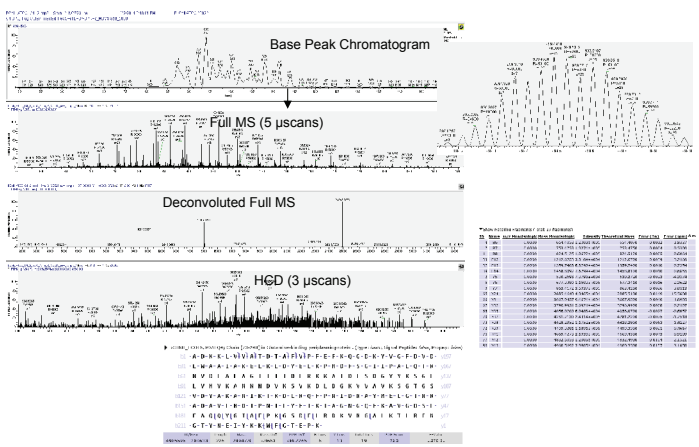
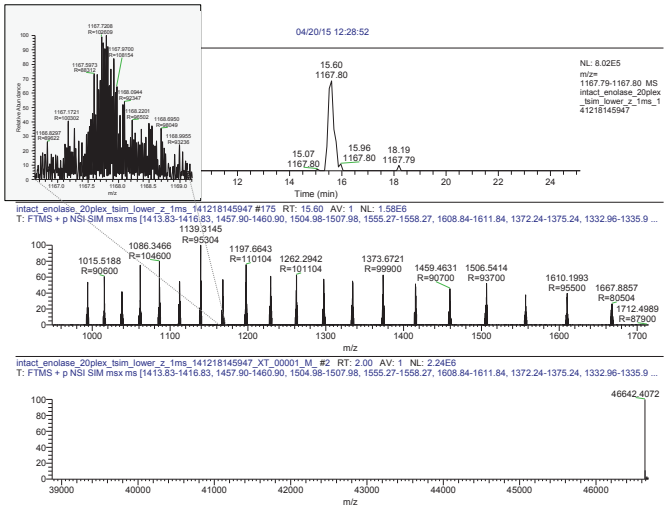


FIGURE 6. Example of top-down protein identification using ProSight PC: Glutamine-binding periplasmic protein (MW: 24.9 kDa).



Most of the proteins identified using “Top-N” “high-high” method have molecular weights < 35 kDa, which is why a multiplex SIM approach was tested to see whether the mass range of isotopically resolved proteins can be extended beyond this limit. With this approach, different charge states of the same protein or of different proteins can be first selected using the quadrupole, then trapped in the HCD cell, and detected all together with the Orbitrap analyzer. Using intact enolase we could demonstrate that proteins up to about 50 kDa can be analyzed with isotopic resolution at LC time scale (see Figure 7).

FIGURE 7. Multiplex SIM spectrum of 20 consecutive charge states of intact enolase (10 µscans @ 240k res. pwr. → acq. time: ~ 5 seconds).



Conclusion

- Q Exactive HF mass spectrometer with its Intact Protein Mode and 1.8 fold increase in resolving power enables high-res analysis at high detection speed which makes it more suitable for high throughput top-down analysis.
- Aided by a new charge assignment/deconvolution algorithm, Q Exactive HF MS provides significant proteoform and protein coverage, even from a single direct infusion spectra.
- A multiplex SIM approach allows analysis of intact proteins up to about 50 kDa with isotopic resolution at LC timescale.

www.thermoscientific.com

Africa +43 1 333 50 34 0
Australia +61 3 9757 4300
Austria +43 810 282 206
Belgium +32 53 73 42 41
Canada +1 800 530 8447
China 800 810 5118 (free call domestic)
400 650 5118
Denmark +45 70 23 62 60
Europe-Other +43 1 333 50 34 0
Finland +358 10 3292 200
France +33 1 60 92 48 00
Germany +49 6103 408 1014
India +91 22 6742 9494
Italy +39 02 950 591
Japan +81 45 453 9100
Korea +82 2 3420 8600
Latin America +1 561 688 8700
Middle East +43 1 333 50 34 0
Netherlands +31 76 579 55 55
New Zealand +64 9 980 6700
Norway +46 8 556 468 00
Russia/CIS +43 1 333 50 34 0
Singapore +65 6289 1190
Spain +34 914 845 965
Sweden +46 8 556 468 00
Switzerland +41 61 716 77 00
UK +44 1442 233555
USA +1 800 532 4752

Thermo SCIENTIFIC
A Thermo Fisher Scientific Brand

PN64455-EN 0615S