Journal of proteome research

Technical Note

pubs.acs.org/jpr

Rapid and Deep Proteomes by Faster Sequencing on a Benchtop Quadrupole Ultra-High-Field Orbitrap Mass Spectrometer

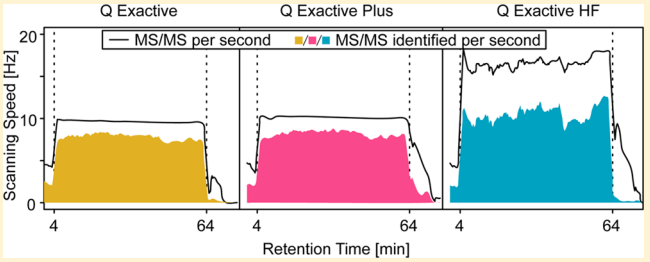
Christian D. Kelstrup,[†] Rosa R. Jersie-Christensen,[†] Tanveer S. Batth,[†] Tabiwang N. Arrey,[‡] Andreas Kuehn,[‡] Markus Kellmann,[‡] and Jesper V. Olsen*,[†]

[†]Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen, Denmark

[‡]Thermo Fisher Scientific (Bremen) GmbH, Hanna-Kunath-Strasse 11, 28199 Bremen, Germany

Supporting Information

ABSTRACT: Shotgun proteomics is a powerful technology for global analysis of proteins and their post-translational modifications. Here, we investigate the faster sequencing speed of the latest Q Exactive HF mass spectrometer, which features an ultra-high-field Orbitrap mass analyzer. Proteome coverage is evaluated by four different acquisition methods and benchmarked across three generations of Q Exactive instruments (ProteomeXchange data set PXD001305). We find the ultra-high-field Orbitrap mass analyzer to be capable of attaining a sequencing speed above 20 Hz, and it routinely exceeds 10 peptide spectrum matches per second or up to 600 new peptides sequenced per gradient minute. We identify 4400 proteins from 1 μg of HeLa digest using a 1 h gradient, which is an approximately 30% improvement compared to that with previous instrumentation. In addition, we show that very deep proteome coverage can be achieved in less than 24 h of analysis time by offline high-pH reversed-phase peptide fractionation, from which we identify more than 140 000 unique peptide sequences. This is comparable to state-of-the-art multiday, multienzyme efforts. Finally, the acquisition methods are evaluated for single-shot phosphoproteomics, where we identify 7600 unique HeLa phosphopeptides in one gradient hour and find the quality of fragmentation spectra to be more important than quantity for accurate site assignment.



KEYWORDS: *Orbitrap, Q Exactive HF, shotgun proteomics, HCD, HeLa, phosphoproteomics, single-shot analysis, high-pH reversed-phase fractionation, parallel acquisition, deep proteome coverage*

INTRODUCTION

The field of mass spectrometry-based proteomics has evolved tremendously over the past decade, and shotgun proteomics in particular has made a great impact in many biological areas. It is now possible to analyze thousands of proteins and their post-translational modifications in a few hours on a routine basis.[1−5] The fundamental principle of shotgun proteomics is based on proteolytic digestion of a protein sample into shorter peptides, which are subsequently separated online by nanoscale liquid chromatography, ionized by electrospray ionization, and directly analyzed in a tandem mass spectrometer followed by bioinformatic interpretation. Although the technology is becoming more mature and easier to use, there is a continuous interest in optimizing and improving every part of the proteomics workflow, as it is still expensive, time-consuming, and incomplete in the number of peptides that are analyzed.[6−9]

A key component of the workflow is the tandem mass spectrometer (MS). The most important feature of an MS instrument in shotgun proteomics is its ability to acquire high numbers of readily identifiable tandem mass (MS/MS) spectra on a chromatographic time scale. As a consequence of this, the Orbitrap mass analyzer has become one of the most popular mass analyzers for shotgun proteomics, as it provides a unique combination of sequencing speed, high resolving power, dynamic range, sensitivity, and mass accuracy.[10] The family of MS instruments with an Orbitrap mass analyzer is constantly growing, of which the first benchtop quadrupole Orbitrap mass spectrometer introduced was the Q Exactive[11] followed by the Q Exactive Plus and, recently, the Q Exactive HF. The new hardware feature on the Q Exactive HF compared to that on the Q Exactive Plus is the incorporation of the smaller ultra-high-field Orbitrap mass analyzer also present in the recently introduced Orbitrap Fusion instrument.[12] The axial frequency of harmonic oscillations of a given ion in an ultra-high-field Orbitrap is 1.8 times higher than that in the standard size high-field Orbitrap analyzer. Because the resolving power of an Orbitrap mass analyzer is proportional to this frequency multiplied by the acquisition time or transient length of the measurement, the ultra-high-field Orbitrap cell theoretically provides 1.8 times higher resolution at the same transient length. Alternatively, the ultra-high-field Orbitrap mass analyzer

Received: September 21, 2014
Published: October 28, 2014

   dx.doi.org/10.1021/pr500985w | J. Proteome Res. 2014, 13, 6187−6195

can perform MS and MS/MS analysis much faster yet at an almost identical resolution compared to that of a normal Orbitrap mass analyzer readout.

A very interesting feature for shotgun proteomics is the new short scans of only a 32 ms transient available in the Q Exactive HF. This short transient is half of the fastest transients available on previous Orbitrap mass analyzer generations and enables much faster scanning speeds, which is especially interesting for fragmentation scans. However, to utilize this scanning speed improvement in an optimal way, it is important to remember that the Q Exactive HF, like its predecessors, is capable of parallel acquisition. This is a feature that results in faster scan cycles through accumulation and preparation of ions for injection into the Orbitrap mass analyzer simultaneous with analysis in the Orbitrap mass analyzer of the previous ion package. Therefore, in order to take full advantage of the Orbitrap acquisition speed gain on the Q Exactive HF, the ion accumulation and preparation step must also be done faster in the parallel mode of operation. This can be accomplished only by using lower fill times, which reduce the number of ions in each scan proportionally at a constant ion flux. Hence, the increase in speed comes at a cost of lower ion abundances, and the challenge can be seen as finding the right balance between quantity and quality of scans.

Here, we set out to systematically define and test a set of optimized parallel acquisition methods for shotgun proteomics and identify the best method in terms of peptides and proteins identified by single-shot liquid chromatography (LC)−MS/MS analyses. Different resolution settings in combination with optimized/parallel injection times were tested across the three generations of Q Exactives. In addition, we also investigated the depth of proteome coverage obtainable in less than 1 day of total MS measurement time by analyzing peptide fractions from offline high-pH reversed-phase chromatography using the fastest scanning method on the Q Exactive HF. Finally, we tested the different acquisition methods on a complex phosphopeptide mixture to determine which instrument parameters are the most important for maximizing coverage and site localization in phosphoproteomics.

MATERIALS AND METHODS

HeLa Lysis and Digestion

Adherent HeLa cervical carcinoma cells were cultured in Dulbecco's modified Eagle's medium (Gibco, Life Technologies) with L-glutamine supplemented with 10% fetal bovine serum (Gibco, Life Technologies) and 100 U/mL penicillin/streptomycin at 37 °C in a humidified 5% CO_2 atmosphere. The cells were harvested at ∼80% confluence by washing twice with PBS (Gibco, Life technologies) and subsequently adding boiling lysis buffer[13] (6 M guanidinium hydrochloride (GndCl), 5 mM tris(2-carboxyethyl)phosphine, 10 mM chloroacetamide, 100 mM Tris, pH 8.5) directly to the plate. The cell lysate was collected by scraping the plate and boiled for additional 10 min followed by micro tip probe sonication (Vibra-Cell VCX130, Sonics, Newtown, CT, USA) for 2 min with pulses of 1 s on and 1 s off at 50% amplitude. Protein concentration was estimated by Bradford, and the lysate was digested with LysC (Wako) in an enzyme/protein ratio of 1:100 (w/w) for 1 h followed by dilution with 25 mM Tris, pH 8.5, to 2 M GndCl and further digested overnight with a protease that cleaves C-terminal to arginine and lysine 1:100 (w/w).[14,15] Protease activity was quenched by acidification with trifluoroacetic acid (TFA) to a final concentration of approximately 1%, and the resulting peptide mixture was concentrated on Sep-Pak (C18 Classic Cartridge, Waters, Milford, MA, USA). Elution was done with 2 mL 40% acetonitrile (ACN), 0.1% TFA followed by 4 mL 60% ACN, 0.1% TFA. The combined eluate was reduced to 1 mL by SpeedVac (Eppendorf, Germany). The final peptide concentration was estimated by measuring absorbance at 280 nm on a NanoDrop (NanoDrop 2000C, Thermo Scientific, Germany).

High-pH Fractionation

HeLa peptides from two replicates (5 mg each) were fractionated in turn using a reversed-phase XBridge BEH130 C18 3.5 μm 4.6 × 250 mm column (Waters, Milford, MA, USA) on an Ultimate 3000 high-pressure liquid chromatography (HPLC) system (Dionex, Sunnyvale, CA, USA) operating at 1 mL/min. Because the Dionex HPLC is a binary pump system, a Rheodyne MXII pump (IDEX Corporation, Rohnert Park, CA, USA) was coupled to the HPLC and used to inject sample onto the column and run simultaneously with the Dionex pump. This was necessary since the sample volume exceeded the maximum capacity of the injection loop on the HPLC autosampler. Buffer A (water) and buffer B (water with 90% ACN) were adjusted to pH 10 with ammonium hydroxide. Peptides were separated by a linear gradient from 5% B to 35% B in 55 min followed by a linear increase to 70% B in 8 min. A total of 63 fractions were collected, which were concatenated to 14 fractions, as described.[16,17] For nanoflow LC−MS/MS, the loading amount was kept constant at 1 μg per injection, estimated by measuring absorbance at 280 nm on a NanoDrop.

TiO_2 Enrichment

For phosphoproteomics experiments, HeLa cells were serum-starved overnight and stimulated with 10% fetal bovine serum for 10 min before lysis, as described above. After digestion, sample volume was doubled by addition of 80% ACN in 12% TFA and subsequently enriched with TiO_2 beads (5 μm, GL Sciences Inc., Tokyo, Japan) as previously described, however, with slight modifications.[18,19] Briefly, the beads were suspended in 20 mg/mL 2,5-dihydroxybenzoic acid (DHB), 80% ACN, and 6% TFA. The sample was incubated with the beads in a sample to bead ratio of 1:2 (w/w) in batch mode for 15 min with rotation. The beads were washed and collected on C8 STAGE-tips[20] with first a 10% ACN and 6% TFA wash followed by a 40% ACN and 6% TFA wash and finally by 80% ACN and 6% TFA. Elution of phosphopeptides was accomplished with 5% NH_3 followed by 10% NH_3 in 25% ACN, which were finally evaporated in a SpeedVac. The enriched phosphopeptides were acidified with TFA to a final concentration of 1%, loaded on preconditioned C18 STAGE-tips, and stored at 5 °C. Just prior to analysis, peptides were eluted from STAGE-tips with 20 μL of 40% ACN, 0.1% TFA followed by 10 μL 60% ACN, 0.1%TFA and reduced to 3 μL by SpeedVac.

Nanoflow LC−MS/MS

The peptide solution was adjusted in volume to an appropriate concentration and kept in loading buffer (5% ACN and 0.1% TFA) followed by autosampling onto an in-house packed 15 cm capillary column with 1.9 μm Reprosil-Pur C18 beads (Dr. Maisch, Ammerbuch, Germany) using an EASY-nLC 1000 system (Thermo Scientific, Odense, Denmark). The column temperature was maintained at 40 °C using an integrated column oven (PRSO-V1, Sonation GmbH, Biberach, Ger-

dx.doi.org/10.1021/pr500985w | J. Proteome Res. 2014, 13, 6187−6195

many) and interfaced online with the mass spectrometer. Formic acid 0.1% was used to buffer the pH in the two running buffers used. The total gradient was 60 min followed by a 17 min washout and re-equilibration. In detail, the flow rate started at 250 nL/min and 8% ACN with a linear increase to 24% ACN over 50 min followed by 10 min linear increase to 36% ACN. The washout followed with a flow rate set to 500 nL/min at 64% ACN for 7 min followed by re-equilibration with a 5 min linear gradient back down to 4% ACN. Finally, the flow rate was set to 250 nL/min for the last 5 min.

The Q Exactive, Q Exactive Plus, and Q Exactive HF instruments (all Thermo Scientific, Bremen, Germany) were freshly cleaned and calibrated using Tune (version 2.3 SP1 build 1788) instrument control software. Spray voltage was set to 2 kV, S-lens RF level at 50, and heated capillary at 275 °C. Full scan resolutions were set to 70 000 at m/z 200 (Q Exactive and Q Exactive Plus) and 60 000 at m/z 200 (Q Exactive HF). Full scan target was 3×10^6 with a maximum fill time of 15 ms. Mass range was set to 375−1500. Target value for fragment scans was set at 1×10^5, and intensity threshold was kept at 1×10^5. Isolation width was set at 1.2 Th unless otherwise specified. A fixed first mass of 100 was used. Normalized collision energy was set at 28. Peptide match was set to off, and isotope exclusion was on. All data was acquired in profile mode using positive polarity.

Raw Data Processing and Analysis

All raw LC−MS/MS data were analyzed by MaxQuant v1.4.1.4 using the Andromeda Search engine and searched against the human Swiss-Prot database, the reviewed part of UniProt, without isoforms (May 2014 release with 20 213 protein sequences).[21] Two analysis groups were made in MaxQuant, enabling one combined analysis of all proteome and phosphoproteome data. Carbamidomethylation of cysteine was specified as fixed modification for both groups. For the proteome data, variable modifications considered were oxidation of methionine, protein N-terminal acetylation, and pyro-glutamate formation from glutamine. The phosphoproteome data was additionally searched with phosphorylation as a variable modification of serine, threonine, and tyrosine residues. An experimental design was used where each raw file was considered an independent experiment except for the high-pH reversed-phase fractionation studies, where fractions were specified and each original sample were considered an experiment. The match between run feature and the second peptide option was disabled, and everything else was set to the default values, including the false discovery rate limit of 1% on both the peptide and protein levels. Phosphorylation sites were considered localized at a site localization probability above 75%. Analysis of data was performed in R[22] with the assistance of scripting in Perl for filtering of the very large result files. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium[23] via the PRIDE partner repository with the data set identifier PXD001305.

RESULTS AND DISCUSSION

Initial Optimization of the Q Exactive HF Mass Spectrometer

To make the most of the faster sequencing capabilities of the Q Exactive HF compared to that of previous Q Exactive instrument generations, we carefully optimized its performance for shotgun proteomics of whole-cell HeLa digests analyzed using a 1 h LC−MS/MS gradient. Building on our previous experiences with the Q Exactive,[24] an initial round of optimization was done on the Q Exactive HF starting from the default settings. For MS full scans, the default values seemed to be optimal, as a higher number of identifications was not achieved by using microscans above one, multiple ranges, a different resolution, or different target value than the default (data not shown). It is noteworthy that, for the analysis of complex mixtures, the best setting for maximum sensitivity, dynamic range, and mass accuracies in full scans was achieved with an AGC target of 3×10^6 on all instruments. We found the optimal full scan mass range was 375−1500 with very little difference in signal-to-noise or number of observable peaks compared to that of other slightly wider or narrower mass windows. The best setting we found for the dynamic exclusion was 20 s, which is equivalent to the widest peaks (base-to-base) observed on this column for 1 h gradients. This value is, of course, expected to be slightly different for other gradient/LC/column combinations; however, the optimum was set to avoid repeated triggering of fragment scans on the same precursor.

The optimal collision energy for a peptide depends on the mass-to-charge (m/z) ratio and charge state of the precursor ion for higher-energy collisional dissociation (HCD).[25] We observed a slight improvement in fragmentation by setting the normalized collision energy (NCE) to 28% compared to that at slightly higher or lower values. Interestingly, at lower NCE values, we observed that doubly charged precursors were often not fragmented sufficiently, as the precursor ion was the base peak. This was not observed for higher charge state precursors. We therefore anticipate that a future improvement is possible by fine tuning the charge-dependent scaling of the normalized collision energy. Switching the peptide match and apex feature settings on were found to give less identifications and were therefore switched off for all experiments. The exclusion of isotopes and charge states 1+ and 6+ were, as many studies have previously confirmed, the best settings.

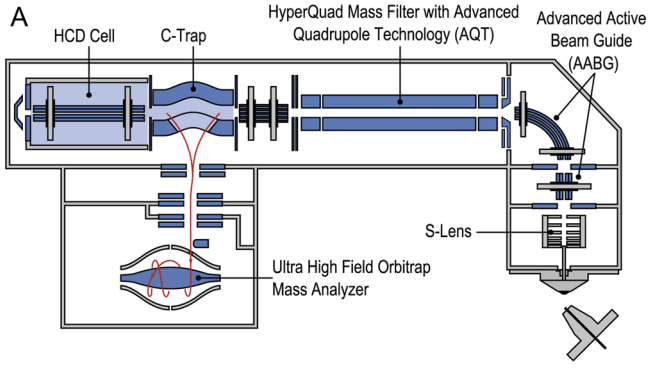
The underfill ratio is a parameter that is used to define the minimum number of ions to be collected for fragment scans, as it is a percentage of the set MS/MS target value. This directly translates into a minimum intensity threshold for triggering fragmentation depending on the allowed maximum injection time. It turned out to be difficult to optimize the underfill ratio for different target values because we continuously obtained ambiguous results that were difficult to interpret. This is likely due to other variable factors such as mass, charge, and peptide sequence differences, which are all important features for the lower limit of signal needed for identification. We therefore decided to keep the method's intensity threshold at 1×10^5 for peak picking in all experiments and set the underfill ratio accordingly. Although this, in our experience, is a low cutoff, it is sufficiently high to avoid triggering sequencing attempts of low-intensity noise peaks. This low-intensity threshold effectively made charge state recognition the limiting factor for the instrument control software to select a precursor ion peak for fragmentation in the samples analyzed.

Acquisition Method Design on the Q Exactive HF

The Q Exactive HF features a new compact ultra-high-field Orbitrap mass analyzer compared to that of the Q Exactive Plus (Figure 1A). This Orbitrap cell is the fastest scanning Orbitrap mass analyzer currently available, featuring a 5 kV central electrode voltage and a maximum inner diameter of 20 mm (detailed comparison in Figure 1B). Compared directly to the Q Exactive and Q Exactive Plus, the Q Exactive HF is roughly

dx.doi.org/10.1021/pr500985w | J. Proteome Res. 2014, 13, 6187−6195

A



B

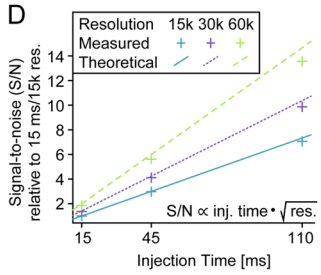
Instrument	Orbitrap central electrode voltage [kV]	Orbitrap analyzer size [mm]	eFT	Specified resolution @ m/z 200	Transient length [ms]	†Resolution per millisecond @ m/z 200
LTQ Orbitrap	3.5	30		140,000	1536	91
LTQ Orbitrap XL	3.5	30		140,000	1536	91
Orbitrap Velos	3.5	30		140,000	1536	91
Orbitrap Velos Pro	3.5	30		140,000	1536	91
Orbitrap Elite	3.5	20	+	336,000	768	438
Exactive	5.0	30		100,000	700	143
Exactive Plus	5.0	30	+	140,000	512	273
Exactive Plus EMR	5.0	30	+	140,000	512	273
Q Exactive	5.0	30	+	140,000	512	273
Q Exactive Plus	5.0	30	+	140,000	512	273
Q Exactive HF	5.0	20	+	240,000	512	469
Orbitrap Fusion	5.0	20	+	480,000	1024	469

†The resolution per millisecond is calculated from rounded values and may be inaccurate

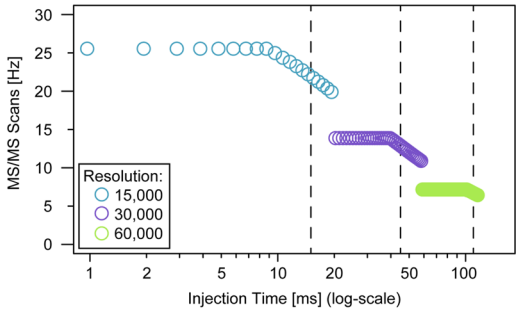
C

Transient Length [ms]	Q Exactive (Plus) Resolution @m/z 200	Q Exactive HF Resolution @m/z 200
32 ms	-	15,000
64 ms	17,500	30,000
128 ms	35,000	60,000
256 ms	70,000	120,000
512 ms	140,000	240,000

D



E



F

	Acquistion Method		
	'Faster'	'Fast'	'Sensitive'
Resolution	15k	30k	60k
Injection Time	15ms	45ms	110ms
MS/MS Scan Time	43ms	73ms	138ms
Scan Numbers	3.2x	1.9x	1x
Signal-to-noise (S/N)	1x	~4x	~15x

S/N ∝ Injection Time • √Resolution

Figure 1. Q Exactive HF mass spectrometer. (A) Schematic of the Q Exactive HF with the ultra-high-field Orbitrap mass analyzer. (B) Overview of the different Orbitrap MS instruments, showing the differences in Orbitrap mass analyzers. (C) Overview of the resolution setting and the time (i.e., transient length) the ions are measured in the Orbitrap mass analyzer. (D) Crosses depicts the measured relationship between signal-to-noise (S/N, measured as intensity divided by noise) for three different resolutions and injection time settings relative to the lowest (15 ms injection time and 15 000 resolution) for direct infusion of the 445.12 peak. Dotted lines indicate theoretical relationship, i.e., linear dependence on injection time and square-root dependence on resolution. (E) Direct injection measurements of HCD scans of the 525 m/z peak (z = 1+, peptide MRFA) at variable injection time and injection time settings. Dotted vertical lines are set at 15, 45, and 110 ms. (F) Three methods are designed with variable MS/MS scan times and resolution settings, giving differences in the quantity of scans and quality of individual scans defined by the signal-to-noise (S/N) relation shown.

1.8 times faster (Figure 1C). Validation of the performance of ultra-high-field Orbitrap mass analyzer was done by varying the resolution and injection time for a constant flux of a singly protonated dodecamethylcyclohexasiloxane ion at m/z 445.12 from ambient air (Figure 1D). As expected, an almost linear increase in intensity divided by noise (~signal-to-noise or S/N) was observed for the peak with increasing injection times, whereas S/N increased proportional to the square root of the increase in resolution.

The detailed speed limits of parallel acquisition were uncovered by measuring the scan-to-scan time for 3 different resolutions (transient lengths) for a series of injection times using direct infusion of the calibration MRFA peptide at m/z 524 with HCD fragmentation enabled (Figure 1E). As expected, this analysis showed that to take advantage of the faster 32 ms transient a much lower injection time was needed; otherwise, the Orbitrap mass analyzer would be idle, waiting for the injection time to be reached. However, with a lowered injection time optimized for full parallel operation, the instrument is capable of a sequencing speed exceeding 25 fragment scans per second (at 524 m/z). The 15 000 resolution scan (32 ms transient) had an optimal point of parallelization at an injection time setting of 9 ms. The 30 000 resolution scan (64 ms transient) was found to operate in parallel at an injection time setting of 41 ms, which is exactly equivalent to the additional 32 ms longer transient time. As expected, the 60 000 resolution (128 ms transient) had a parallel injection time setting an additional 64 ms higher at 105 ms. A small but measurable effect of varying the isolated precursor m/z was observed, where a fragment scan of the 196 m/z ion could be measured 2 ms faster than precursors at 524 m/z, allowing an 11 ms injection time to be parallel with the 32 ms transient. We therefore decided on pragmatic values for the injection time slightly above the values for perfectly parallel operation. Our choice was 15, 45, and 110 ms injection times for the corresponding resolutions (transient lengths) of 15 000 (32 ms), 30 000 (64 ms), and 60 000 (128 ms). These injection time values are indicated with vertical dotted lines in Figure 1E.

From a theoretical perspective, these combinations of injection times and resolutions can be used to create three different sets of fragment scans that differ in how the speed of analysis is balanced against the overall spectrum quality in the individual scans. This is outlined in Figure 1F, where the three settings of injection time and resolution are termed faster, fast, and sensitive in relation to how they differ between quality and quantity. As indicated, the faster method is roughly 3 times faster than that of the sensitive method. The previously described S/N relationship can be used to estimate a difference of roughly 15-fold in what S/N can be expected from a given intensity precursor using the faster versus sensitive methods.

dx.doi.org/10.1021/pr500985w | J. Proteome Res. 2014, 13, 6187−6195

Evaluation of HeLa Dilution Series and Acquisition Method Comparison

To evaluate the performance of the new 32 ms transient against previously available transient lengths, we tested four different acquisition methods against a concentration range of a proteolytic digest derived from a HeLa whole-cell lysate using a 1 h LC−MS/MS gradient. The four methods consisted of the three previously described above, fast, faster, and sensitive, and a fourth method based on a 32 ms transient but not optimized for parallel acquisition and therefore allowing for longer fill times up to 110 ms, making use of the automatic gain control (AGC) feature. The rationale behind the default method is to reach a specific target value of ions required for HCD by predicting the optimal fill time based on precursor ion intensities in previous full scans. We term this method pAGC to underline the predictive nature of the method design. The comparison between the four acquisition methods was done as triplicate technical replicates (Figure 2A). To evaluate and compare the methods, all raw files were processed and analyzed together using MaxQuant software (www.maxquant.org). The number of fragment scans was found to increase for all methods with higher peptide load, as shown on Figure 2B. This indicates that the number of detectable precursors is a limiting factor at lower peptide loads. We also observed that the difference in the number of scans acquired between faster, fast, and sensitive methods corresponds to the theoretical difference in quantity of scans, as outlined previously with roughly a factor of 2 in the number of HCD scans between the fast and sensitive scanning methods and a factor of 3 between the fastest and sensitive methods (Figure 1F). Interestingly, we found that the pAGC method showed a quite different behavior.

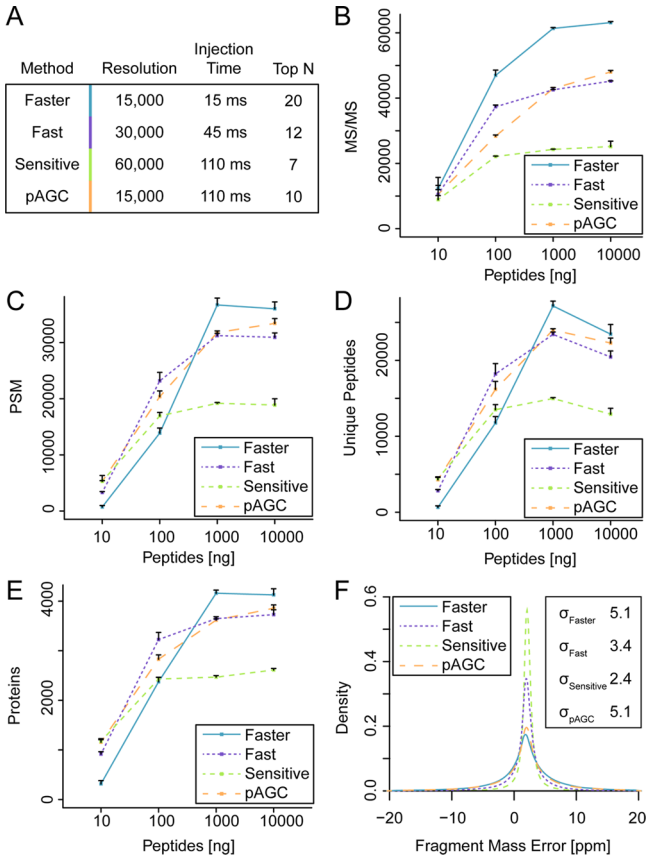


Figure 2. Method comparison and dilution series. All error bars depict the standard deviation of a triplicate measurement. (A) Overview of the four methods; cycle time was kept as constant as possible, and they therefore differ in the number of fragment scans per MS/MS. (B) Number of MS/MS events triggered in each of the method and dilution combinations. (C) Number of peptide spectrum matches (PSMs) in each of the method and dilution combinations. (D) Number of peptides with a unique sequence in each of the method and dilution combinations. (E) Number of inferred proteins from the identified peptides in each of the method and dilution combinations. (F) Standard deviation and Gaussian kernel density distributions of fragment ion mass error are shown for each of the four methods for one of the 1000 ng peptide measurements.

To establish which acquisition methods performed best at the different peptide loads, we looked closer into the overall peptide and protein identifications. For all methods except the pAGC we observed a saturation effect, where the highest load (10 000 ng of peptides) reached only a similar or even lower number of identifications compared to that with 1000 ng, shown in Figure 2C−E. This is likely due to chromatography limitations, as the capacity of the column used is compromised at 10 μg loads, where abundant peaks show increased peak tailing. Interestingly, the faster, fast, and sensitive methods each provided the best identification number for 1000, 100, and 10 ng loads, respectively, whereas the pAGC was constantly the second best method for all peptide loads. This shows the potential of increased robustness of a method that adapts to the sample, yet, at the same time, it leaves room for improvement. As expected from theoretical considerations, we find that when analyzing the fragment mass error distributions the use of higher resolutions provides better mass accuracy, as shown in Figure 2F. We did not take advantage of this in this analysis, as a ±20 ppm search window was used, but a narrower fragment ion match tolerance could be used in the future to improve the confidence of a peak assignment. Regardless, these dilution experiments established that when sample amounts are not limited the new fastest 32 ms sequencing method outperforms all of the others by identifying significantly more peptides and proteins.

Comparison among Q Exactive Generations

To compare the performance of the Q Exactive HF to that of the two previous generations, the Q Exactive Plus and the Q Exactive, we analyzed 1 μg of the same HeLa lysate in triplicate using 3 different isolation widths on all three instruments using the same 1 h LC−MS/MS gradient as described above. In total, this resulted in 27 raw files. To achieve the fairest comparison, all three instruments were carefully cleaned, calibrated, and optimized for best performance. We also made use of the same nano-LC system and analytical column that were moved in succession between the instruments. For all instruments, the fastest scanning method available was used. The method setup, as outlined in Figure 3A, where a different top n was used, ensured comparable scan cycle times of approximately 1.2 s, including full scan and n number of HCD scans. This is shown in Figure 3B, where a zoom of a total ion chromatogram (TIC) view of three resulting raw MS files around the 50th retention minute after gradient elution started are shown. Importantly, all methods finish their full top n in 1.2 s. We observed that the Q Exactive HF generates on the order of 70% more HCD-MS/MS scans than that of the Q Exactive, which results in 38% more identified peptide−spectrum matches (PSMs) and 32% more unique peptide sequences and proteins, as shown on Figure 3C−F. This again underscores the benefits of faster

dx.doi.org/10.1021/pr500985w | J. Proteome Res. 2014, 13, 6187−6195

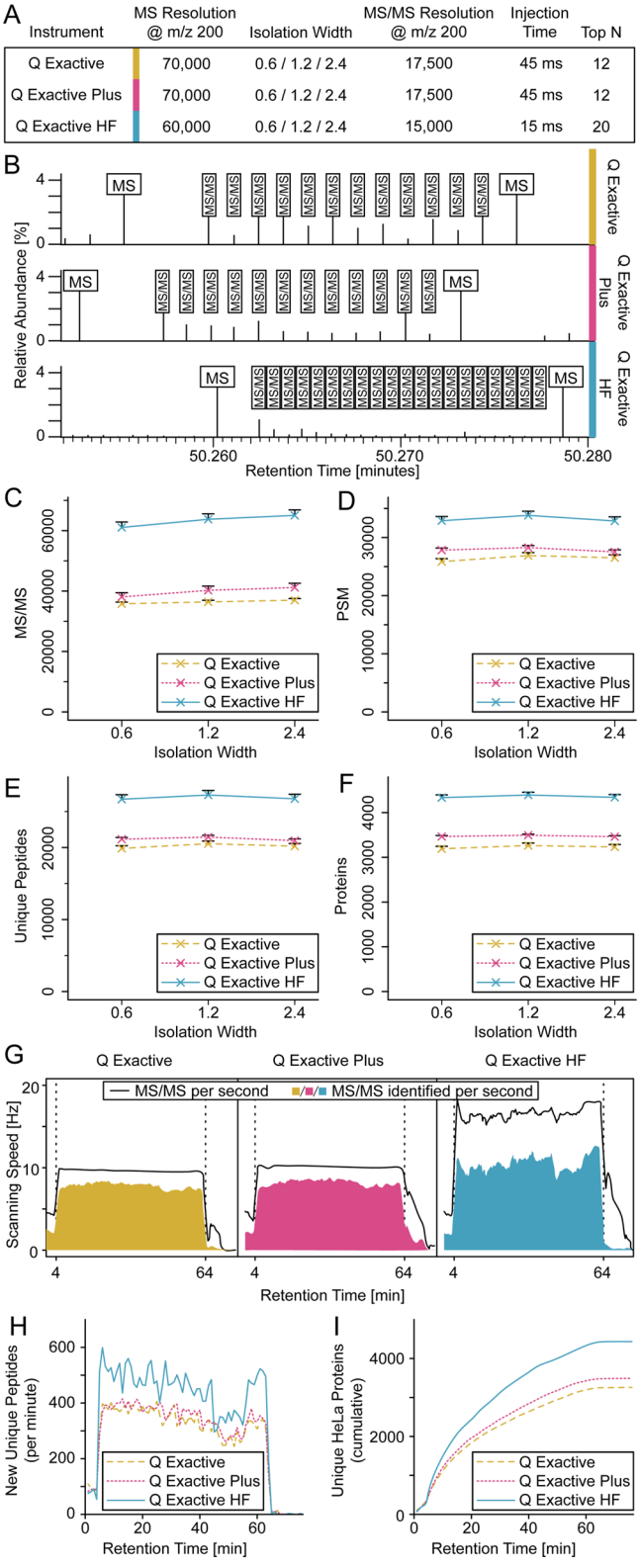


Figure 3. Instrument comparison using 1 μg HeLa peptides on a 1 h gradient. All error bars depict the standard deviation of triplicate measurements. (A) Instrument method layout; cycle time was kept as constant as possible, and they therefore differ in the number of fragment scans per MS scan. (B) Depiction of the cycle time for a full MS-to-MS scan cycle on each instrument around the 50th minute in retention time. (C) Number of MS/MS events generated for each instrument across each isolation width setting. (D) Number of peptide spectrum matches (PSMs) for each instrument across each isolation width setting. (E) Number of peptides with a unique sequence for each instrument across each isolation width setting. (F) Number of inferred proteins from the identified peptides for each instrument across each isolation width setting. (G) Comparison of scanning speed and identification speed for each of the three instruments across the acquired raw files. The dotted vertical lines show the 1 h elution from the 4th to the 64th minute. (H) Number of new unique peptides per minute of gradient elution is depicted across the gradient for each instrument. (I) The accumulating number of proteins across elution is shown for each instrument.

peptide sequencing in shotgun proteomics. The observed difference in identification rate among the instruments is expected, as the number of ions in each scan is lower due to the 3-fold lower maximum injection time for the faster method on the Q Exactive HF. Between the first two generations, the Q Exactive Plus was found to outperform the Q Exactive by 7–10% in the number of identifications.

When analyzing the performance of the different Q Exactive generations across retention time, under extremely optimized conditions, both the Q Exactive and the Q Exactive Plus triggered MS/MS events at a very constant 10 scans per second (10 Hz) during the full 1 h elution, as shown on Figure 3G, whereas the Q Exactive HF reaches 18 scans per second (18 Hz) and shows a higher degree of fluctuations. This is most likely due to the instrument control software running out of new candidates for sequencing in the full scans, as all isotope clusters have already been picked previously and are therefore dynamically excluded by the software. This phenomenon was also observed on the cousin instrument, the Orbitrap Fusion.[12] Despite running out of candidates, the instrument is still reaching close to its maximum sequencing speed, as the shorter scan cycles obviously increase the total number of scan cycles.

The identification rates of both the Q Exactive and Q Exactive Plus were close to 8 PSMs per second, whereas the Q Exactive HF generally achieved above 10 PSMs per second, indicating that the number of identified spectra on the Q Exactive HF surpassed the maximum number of spectra acquired on the previous generations of Q Exactive instruments. Analyzing the number of new unique peptides (based on sequence), this corresponds to upward of 600 new peptides per minute of gradient, whereas the older instruments are peaking around 400 new peptides per minute, as shown on Figure 3H. The accumulated number of assembled proteins that can be differentiated based on the peptide evidence increases throughout the 1 h gradient, reaching almost 4500 on the Q Exactive HF compared to 3500 on the Q Exactive Plus and 3300 on the Q Exactive, Figure 3I.

Deep Proteome Coverage with Offline Peptide Fractionation

To assess the Q Exactive HF for in-depth proteome profiling, we evaluated the use of fastest scanning speed with an offline high-pH reversed-phase peptide fractionation scheme, which has been shown to achieve both high resolution and good orthogonal separation power compared to that of the traditional online low-pH reverse-phase used in LC–MS/MS.[16] Our setup was similar to those previously published, as outlined in Figure 4A, where the concatenation scheme reduces the original 63 fractions to 14, which are subsequently analyzed in turn using a 1 h gradient and the faster scanning speed method previously described. It is noteworthy that the

dx.doi.org/10.1021/pr500985w | J. Proteome Res. 2014, 13, 6187–6195

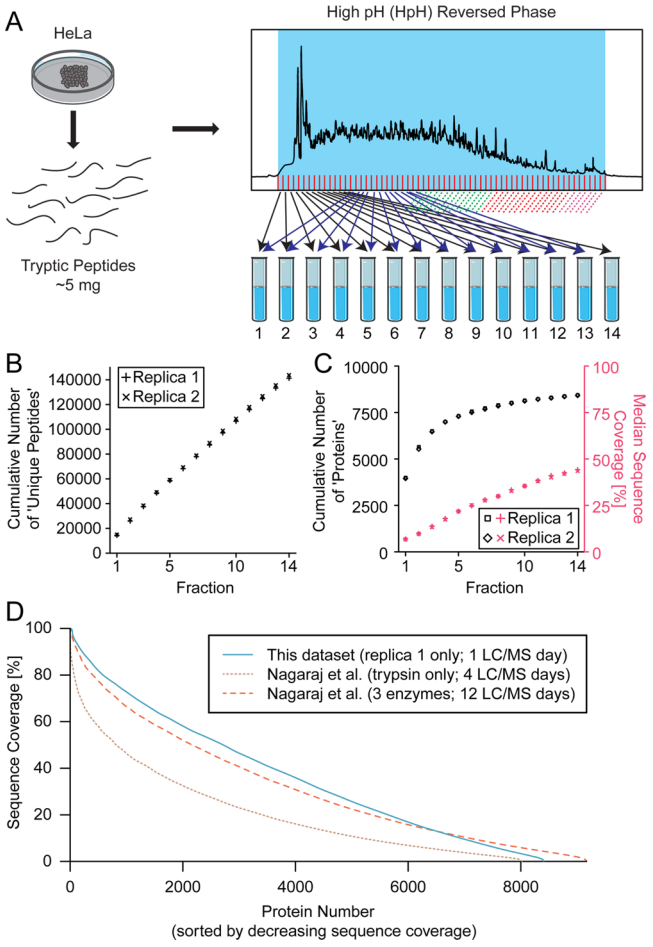


Figure 4. Deep proteome measurements on the Q Exactive HF. (A) Fractionation scheme where HeLa lysate is digested to peptides and fractionated by high-pH reverse phase. Fractions are concatenated to give 14 fractions. (B) Comparison of the cumulative number of unique peptides for two replicate measurements. (C) Comparison of the cumulative number of proteins and median protein sequence coverage for two replicate measurements. (D) Comparison of sequence coverage obtained here to results reported by Nagaraj et al.[26] (using the human SwissProt database without isoforms as template).

combined analysis time of all fractions including everything (standards for quality control, re-equilibration/cleaning of the column, and pickup/loading of the samples) was kept below 1 day of total MS measurement time.

The high performance of the offline peptide separation power could be confirmed by the small overlap of identifications between fractions. This can be seen in Figure 4B, as a near linear increase in the cumulative unique number of unique peptide sequences identified up to more than 140 000 unique peptides in less than 24 h of LC–MS/MS measurements. The technical reproducibility was good, as the two biological replicates analyzed had almost identical numbers. Investigating the identified number of corresponding proteins, it can be seen from Figure 4C that the first fraction alone had peptide evidence for more than 4000 out of the total number of 8400 proteins identified. As such, peptides from just a single fraction contain evidence for close to half of the proteins found across all fractions, which could, perhaps, be useful for experiments where the maximum number of proteins needs to be identified in the shortest possible time. Not surprisingly, the protein number did not increase linearly at the same rate as peptides when all fractions were analyzed. Interestingly, we found that the sequence coverage obtained in this study with less than 1 day of instrument time is comparable in coverage to that of previous state-of-the-art deep HeLa proteome using three different enzymes and 12 days of measurement time, as shown in Figure 4D.[26] This is a significant improvement over previous efforts and highlights the benefits of simple offline peptide fractionation combined with very fast sequencing speed to achieve deep proteome coverage in relatively short time frames.

Q Exactive HF for PTM Analysis: HeLa Phosphoproteome Acquisition Method Comparison

To investigate and find the best balance between quality and quantity of HCD fragment scans for post-translational modification studies, an investigation into the different acquisition strategies was repeated for a single-shot phosphoproteomics sample. The same four acquisition methods were used as in first method comparison (Figure 2) to analyze phosphopeptide mixtures enriched individually from 1 mg of HeLa digests by TiO_2, as shown in Figure 5A. As expected, the

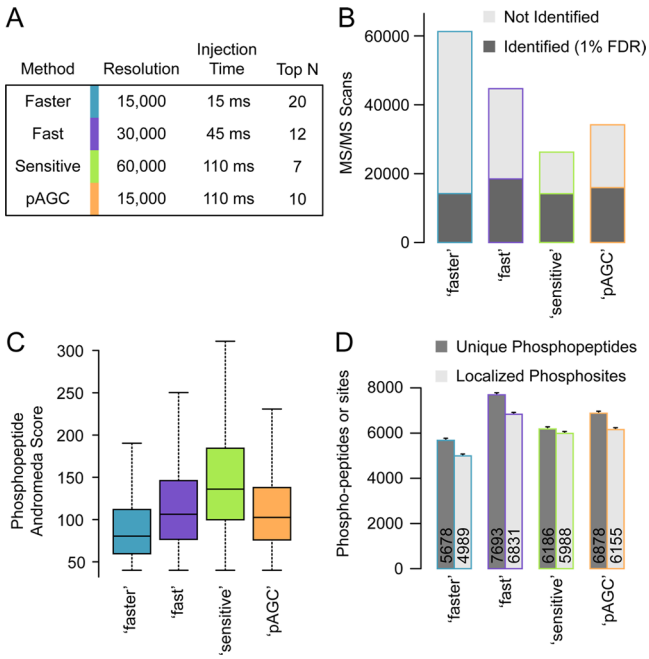


Figure 5. Method comparison for phosphoproteome analysis on the Q Exactive HF. (A) The four methods used in the comparison. (B) The average number of triggered and identified fragment scans in the four methods. (C) Boxplots depicting the Andromeda score distribution for phosphopeptides identified in the four methods. (D) Comparison of identified or localized phosphopeptides in each of the four methods. Error bars indicate the residual standard error from an analysis of variance (ANOVA) model.

faster method generated the most fragment scans followed, in descending order, by the fast, pAGC, and sensitive methods, as shown in Figure 5B. This is similar to the findings from the proteome results presented previously. However, despite generating the most fragment scans, the faster method now performed the worst when comparing the number of PSMs identified for each acquisition method. We could not explain this by a difference in ion count in the fragment scans, as this was found to be comparable to results for the 1 μg peptide load (shown in Supporting Information Figure 1). Instead, we

dx.doi.org/10.1021/pr500985w | J. Proteome Res. 2014, 13, 6187–6195

suspect that this is due to the quality of the spectra generated, as the faster method generated the lowest scoring identifications, as shown in Figure 5C. As expected, the sensitive method generated by far the best quality fragment scans with the highest score distribution of identified phosphopeptides. This can also be seen when investigating the localized phosphorylation sites compared to just looking at identified peptides with a possible ambiguity in site assignment, as shown in Figure 5D. In total, it seems that phosphoproteomic investigations place stricter requirements on fragment ion coverage in fragmentation spectra compared to those of proteome experiments and that the use of methods that allow higher quality spectra due to longer injection time are beneficial.

CONCLUSIONS

The ultra-high-field Orbitrap enables a faster scanning mode that can be taken advantage of when the injection time is lowered. This allows for data-dependent methods in which full-scan and 20 high-resolution HCD scans can be performed in 1.2 s, which, on previous generations of Q Exactive instruments, allow for only 12 HCD events in the same time frame. A 1 h gradient analysis of a complex 1 μg HeLa digest provided the best results for this faster mode, where 70% more fragmentation scans were generated, giving 30% more identifications compared to that with the Q Exactive instrument. This enhancement of the Q Exactive HF was found to provide a much larger improvement than the difference between the Q Exactive Plus and the Q Exactive, where the difference was found to be 7−10%. Under these optimal conditions and using the Q Exactive HF, it is now possible to exceed 10 PSMs per second or to identify up to 600 new peptides per minute of LC gradient, reaching, in total, almost 4500 proteins in 1 h of gradient. Combined with offline peptide fractionation by high-pH reversed-phase chromatography, we were able to achieve a sequence coverage median above 40% of more than 8400 proteins in less than a day of MS instrument time, which is equivalent to state-of-the-art results from multienzyme, 12 day LC−MS efforts published previously. Finally, we foundthat phosphoproteome studies still present an additional challenge in accurate site localization. Consequently, higher quality acquisition methods gave the highest numbers, with 7600 unique phosphopeptides identified in 1 h of gradient elution. We hope that this study serves as a valuable resource for the design of future experiments that require rapid and deep proteome coverage.

ASSOCIATED CONTENT

Supporting Information

Boxplot comparison of the ion count in all MS/MS scans for three dilutions from the dilution experiments and the phosphoproteome experiment for each of the four different methods used. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: +45-353-25022. E-mail: jesper.olsen@cpr.ku.dk.

Notes

The authors declare the following competing financial interest(s): T.N.A., A.K., and M.K. are employees of Thermo Fisher Scientific, the manufacturer of the Q Exactive, the Q Exactive Plus, and the Q Exactive HF instrument used in this research.

ACKNOWLEDGMENTS

Work at the Center for Protein Research is supported by a generous donation from the Novo Nordisk Foundation. This work was also supported by the Sapere Aude Research Leader grant to J.V.O. Part of this work has been funded by PRIME-XS, a seventh Framework Programme of the European Union (contract no. 262067-PRIME-XS). The authors would like to thank Mathias Mueller from Thermo Fischer Scientific and members of the Proteomics Program for critical input on the manuscript.

ABBREVIATIONS

ACN, acetonitrile; AGC, automatic gain control; ANOVA, analysis of variance; ASP, advanced signal processing; HCD, higher-energy collisional dissociation; LC, liquid chromatography; *m/z*, mass-to-charge; MS, mass spectrometry; MS/MS, tandem mass spectrometry; NCE, normalized collision energy; PSMs, peptide spectrum matches; S/N, signal-to-noise; TFA, trifluoroacetic acid

REFERENCES

(1) Nagaraj, N.; Kulak, N. A.; Cox, J.; Neuhauser, N.; Mayr, K.; Hoerning, O.; Vorm, O.; Mann, M. System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Mol. Cell. Proteomics* **2012**, *11*, M111.013722.

(2) Lundby, A.; Secher, A.; Lage, K.; Nordsborg, N. B.; Dmytriyev, A.; Lundby, C.; Olsen, J. V. Quantitative maps of protein phosphorylation sites across 14 different rat organs and tissues. *Nat. Commun.* **2012**, *3*, 876.

(3) Hebert, A. S.; Richards, A. L.; Bailey, D. J.; Ulbrich, A.; Coughlin, E. E.; Westphall, M. S.; Coon, J. J. The one hour yeast proteome. *Mol. Cell. Proteomics* **2014**, *13*, 339−47.

(4) Yin, X.; Liu, X.; Zhang, Y.; Yan, G.; Wang, F.; Lu, H.; Shen, H.; Yang, P. Rapid and sensitive profiling and quantification of the human cell line proteome by LC−MS/MS without prefractionation. *Proteomics* **2014**, *14*, 2008−16.

(5) Pirmoradian, M.; Budamgunta, H.; Chingin, K.; Zhang, B.; Astorga-Wells, J.; Zubarev, R. A. Rapid and deep human proteome analysis by single-dimension shotgun proteomics. *Mol. Cell. Proteomics* **2013**, *12*, 3330−8.

(6) Altelaar, A. F.; Munoz, J.; Heck, A. J. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nat. Rev. Genet.* **2013**, *14*, 35−48.

(7) Zubarev, R. A. The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics* **2013**, *13*, 723−6.

(8) Kalli, A.; Smith, G. T.; Sweredoski, M. J.; Hess, S. Evaluation and optimization of mass spectrometric settings during data-dependent acquisition mode: focus on LTQ-Orbitrap mass analyzers. *J. Proteome Res.* **2013**, *12*, 3071−86.

(9) Randall, S. M.; Cardasis, H. L.; Muddiman, D. C. Factorial experimental designs elucidate significant variables affecting data acquisition on a quadrupole Orbitrap mass spectrometer. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1501−12.

(10) Zubarev, R. A.; Makarov, A. Orbitrap mass spectrometry. *Anal. Chem.* **2013**, *85*, 5288−96.

(11) Michalski, A.; Damoc, E.; Hauschild, J. P.; Lange, O.; Wieghaus, A.; Makarov, A.; Nagaraj, N.; Cox, J.; Mann, M.; Horning, S. Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol. Cell. Proteomics* **2011**, *10*, M111.011015.

(12) Senko, M. W.; Remes, P. M.; Canterbury, J. D.; Mathur, R.; Song, Q.; Eliuk, S. M.; Mullen, C.; Earley, L.; Hardman, M.; Blethrow, J. D.; Bui, H.; Specht, A.; Lange, O.; Denisov, E.; Makarov, A.; Horning, S.; Zabrouskov, V. Novel parallelized quadrupole/linear ion trap/Orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates. *Anal. Chem.* **2013**, *85*, 11710−4.
(13) Kulak, N. A.; Pichler, G.; Paron, I.; Nagaraj, N.; Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **2014**, *11*, 319−24.
(14) Olsen, J. V.; Ong, S. E.; Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* **2004**, *3*, 608−14.
(15) Poulsen, J. W.; Madsen, C. T.; Young, C.; Poulsen, F. M.; Nielsen, M. L. Using guanidine-hydrochloride for fast and efficient protein digestion and single-step affinity-purification mass spectrometry. *J. Proteome Res.* **2013**, *12*, 1020−30.
(16) Wang, Y.; Yang, F.; Gritsenko, M. A.; Wang, Y.; Clauss, T.; Liu, T.; Shen, Y.; Monroe, M. E.; Lopez-Ferrer, D.; Reno, T.; Moore, R. J.; Klemke, R. L.; Camp, D. G., II; Smith, R. D. Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics* **2011**, *11*, 2019−26.
(17) Batth, T. S.; Francavilla, C.; Olsen, J. V. Off-line high pH reversed-phase fractionation for in-depth phosphoproteomics. *J. Proteome Res.* **2014**, DOI: 10.1021/pr500893m.
(18) Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **2006**, *127*, 635−48.
(19) Lundby, A.; Andersen, M. N.; Steffensen, A. B.; Horn, H.; Kelstrup, C. D.; Francavilla, C.; Jensen, L. J.; Schmitt, N.; Thomsen, M. B.; Olsen, J. V. In vivo phosphoproteomics analysis reveals the cardiac targets of beta-adrenergic receptor signaling. *Sci. Signal* **2013**, *6*, rs11.
(20) Rappsilber, J.; Ishihama, Y.; Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **2003**, *75*, 663−70.
(21) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*, 1367−72.
(22) *The R Project for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria; http://www.R-project.org.
(23) Vizcaino, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Rios, D.; Dianes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; Binz, P. A.; Xenarios, I.; Eisenacher, M.; Mayer, G.; Gatto, L.; Campos, A.; Chalkley, R. J.; Kraus, H. J.; Albar, J. P.; Martinez-Bartolome, S.; Apweiler, R.; Omenn, G. S.; Martens, L.; Jones, A. R.; Hermjakob, H. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **2014**, *32*, 223−6.
(24) Kelstrup, C. D.; Young, C.; Lavallee, R.; Lund Nielsen, M.; Olsen, J. V. Optimized fast and sensitive acquisition methods for shotgun proteomics on a quadrupole Orbitrap mass spectrometer. *J. Proteome Res.* **2012**, *11*, 3487−97.
(25) Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. Higher-energy C-trap dissociation for peptide modification analysis. *Nat. Methods* **2007**, *4*, 709−12.
(26) Nagaraj, N.; Wisniewski, J. R.; Geiger, T.; Cox, J.; Kircher, M.; Kelso, J.; Paabo, S.; Mann, M. Deep proteome and transcriptome mapping of a human cancer cell line. *Mol. Syst. Biol.* **2011**, *7*, 548.