Simplified high-throughput methods for deep proteome analysis on the timsTOF Pro

Jarrod J Sandow[1,2,3,*], Giuseppe Infusini[1,2,3], Laura F Dagley[1,2], Rune Larsen[1,2], Andrew I Webb[1,2,3,*]

1. Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Melbourne, Victoria 3052, Australia
2. Department of Medical Biology, University of Melbourne, Parkville, Melbourne, Victoria 3010, Australia
3. Ion Opticks, 1G Royal Parade, Parkville, Melbourne, Victoria 3052, Australia

*co-corresponding authors

Corresponding author details: Jarrod J Sandow (sandow@wehi.edu.au); Andrew I Webb (webb@wehi.edu.au)

Abstract

Recent advances in mass spectrometry technology have seen remarkable increases in proteomic sequencing speed, while improvements to dynamic range have remained limited. An exemplar of this is the new timsTOF Pro instrument, which thanks to its trapped ion mobility, pushes effective fragmentation rates beyond 100Hz and provides accurate CCS values as well as impressive sensitivity. Established data dependent methodologies underutilize these advances by relying on long analytical columns and extended LC gradients to achieve comprehensive proteome coverage from biological samples. Here we describe the implementation of methods for short packed emitter columns that fully utilize instrument speed and CCS values by combining rapid generation of deep peptide libraries with enhanced matching of single shot data dependent sample analysis. Impressively, with only a 17 minute gradient separation (50 samples per day), the combination of high performance chromatography and CCS enhanced library based matching resulted in an average of 6,690 protein identifications within individual samples, and 7,797 proteins cumulatively across replicates from HeLa cell tryptic digests. Additionally, an ultra-high throughput setup utilizing 5 min gradients (180 samples per day) yielded an average of 2,800 protein identifications within individual samples and 4,254 proteins cumulatively across replicates. These workflows are simple to implement on available technology and do not require complex software solutions or custom made consumables to achieve high throughput and deep proteome analysis from biological samples.

Introduction

The orchestration of essentially all biological processes is accomplished by proteins, the actors that create and control any given phenotype. The large-scale identification and quantification of proteins – termed proteomics – has facilitated researchers in studying and understanding complex phenotypes, encompassing a systems-wide view of the deep complexity of living systems. Further development in proteomics approaches hold exciting promise for advancing our understanding of cell biology and biological systems in health and disease. Mass spectrometry (MS)-based approaches to proteomics has become the technology of choice for systems-wide analysis of proteins, their post-translational regulation and interactions. Over the past few decades, technology and methodological improvements have pushed proteomics to the forefront of addressing complex questions in cell biology and biomedicine. These

improvements have been driven by advances in chromatography, informatics and primarily, developments that increase the accuracy, sensitivity and speed of MS instruments.

Proteomics approaches predominantly consist of a 'bottom-up' workflow, where proteins are extracted from a sample of interest and enzymatically hydrolyzed to create shorter peptides that are more amenable to high resolution chromatography and mass spectrometric analysis. The resulting highly complex mixture of peptides are separated via nano-flow ultra-high performance liquid chromatography (UHPLC) and introduced into the mass spectrometer by electrospray ionization. Operating in 'data dependent' mode of acquisition, mass spectrometers detect suitable peptide precursor ions (MS) and subjects them to higher energy collisions to induce fragmentation (MS/MS). High precision mass spectrometers detect in excess of 100,000 molecular features[1], of which, only a small proportion are identified. Improvements to shotgun proteomics approaches to increase the number of identifications in any given sample, are hampered by increases in multiplexed (chimeric) spectra. Thus, any improvements in sensitivity, dynamic range or chromatographic peak capacity all increase the number of co-eluting peptides, providing diminishing returns using traditional approaches.

Many MS analyzers have been employed in shotgun proteomics, however, time-of-flight (TOF) instruments in particular have properties ideal for the analysis of complex peptide mixtures. TOF instrument performance has steadily improved over the years providing resolving power >35,000 within <100µs of acquisition time[2]. The very fast acquisition rates of TOF instruments also permits coupling to fast orthogonal separation techniques such as ion-mobility spectrometry (IMS)[3]. IMS separates ions in the gas phase based in their collisional cross-sectional area (CCS, Ω), providing information about the size and shape of the molecules. IMS typically has separation times of 10-100 milliseconds. As the separation speed of IMS (milliseconds) sits between that of liquid chromatography (seconds) and MS TOF detection (micro seconds), the nesting of IMS between LC and MS provides an additional dimension of orthogonal information, drastically increasing the effective separation of co-eluting and near isobaric features. However, traditional IMS approaches such as drift tubes, have proved challenging due to device sizes, voltages required and limits to the proportion of the incoming ion-beam that can be utilized[4].

Trapped Ion Mobility (TIMS) provides a fundamentally new approach to IMS, using an electrostatic gradient tunnel to hold ions against the incoming gas stream. A molecules position within the TIMS device is a function of its CCS area, thus providing high resolution IMS in a short space at low voltages and has the additional benefit of concentrating ions from the incoming ion beam to increase sensitivity. The recent implementation of tandem, sequential accumulation and separation TIMS tunnels now allows for 100% duty cycle to be achieved[5]. The recently implemented 'Parallel Accumulation – SErial Fragmentation' (PASEF) method implemented on the timsTOF Pro takes further advantage of the ion concentration and IMS separation, staggering fragmentation with the near linear correlation of peptide mz and CCS (where low mz = low mobility and high mz = high mobility), improving the effective rate of precursor fragmentation beyond 100Hz[6].

Whilst improvements in MS sequencing speed provide more complete datasets, the overall performance and capability of LC-MS system still relies heavily on chromatographic separation performance. In the pursuit of ever-increasing depths of coverage, columns and separation gradients have tended to become longer. However, as MS dynamic range is one of the main limitations, these approaches produce diminishing returns, resulting in modest increases in proteome coverage at the expense of very large amounts of MS time and

consequently low through-put. The release of instruments with drastically improved sequencing speeds opens up new possibilities that can potentially provide much greater depth of information per unit of MS time[7], which will be of particular importance in the adoption of LC-MS into clinical utility. Thus, strategies that combine orthogonal pre-fractionation to increase the effective dynamic range, run on ultra-short gradients, is an effective method for overcoming the dynamic range limitation of MS instruments[8].

Here we present an optimized and easily implementable approach for generating deep proteomes in very short timescales enabling increased through-put. The method combines high pH reversed phase stage-tip fractionation for peptide feature library generation on short gradients on high peak capacity commercially available nano-columns to take full advantage of the timsTOF Pro's >100 Hz tandem MS scanning speed. Additionally, accurate CCS values are exploited[9] to provide deep library-based matching of individual samples running on short gradients for drastically improved utilization of MS analysis time and increasing peptide and protein identifications.

Results

Traditionally, single injection shotgun proteomics workflows have utilized liquid chromatography gradients of >60min to ensure that chromatographic peaks are wide enough to allow for multiple MS1 measurements across the peaks facilitating accurate label-free quantitation. In data-dependent acquisition modes, the minimum duty cycle of the instrument dictates the required peak widths. With the advent of the timsTOF Pro mass spectrometer and PASEF acquisition modes[7] we now have the ability to perform data-dependent acquisition using sub-one second duty cycles allowing for significantly reduced peak widths and shorter LC gradients. Shotgun proteomic workflows designed for >60min gradients achieve high numbers of peptide identifications and robust quantitation but limit the number of samples analyzed to around 10-12 per day (Fig. 1A, Supp. Fig. 1). To maximize the efficiency of peptide and protein identification we took advantage of the reduced duty cycle time on the timsTOF Pro with PASEF by designing a series of short gradients that were paired with analytical columns of specific dimensions to optimize results from the reduced gradient lengths (Fig. 1B and C). We first utilized an analytical column configuration measuring 15cm X 75µm with 1.6um C18 particles packed into the emitter coupled to ultra-high-performance liquid chromatography (UHPLC) instrument capable of loading and equilibrating analytical columns under set backpressure conditions. Equilibrating and sample loading the analytical column using 980 bar backpressure allows for the combined steps to be reduced to 6 minutes for a 1µl sample injection. These UHPLC systems require a period of valve switching and gradient preparation between the equilibration/load steps and the sample gradient introducing a period of around 2.8 minutes of "dead time". A 17min sample gradient is then performed with an additional 3 minutes to allow for ramping and washing the column in a high concentration of organic mobile phase. Including sample pick-up, this workflow results in a sample through-put of approximately 50 samples per day (Fig. 1B, Supp. Fig. 2). Next, we tested a shorter column measuring 5cm X 150µm with 1.6um C18 particles packed into the emitter tip coupled to a UHPLC that retains a stable flowrate of 2µl/min and loads and injects samples by using an injection valve to add or remove the sample loop from the flow path. In this arrangement, the equilibration and sample injection steps are performed concurrently as the mobile phase continues to flow over the analytical column during a 1µl sample injection. The sample loop is then switched online, with the 1µl of sample and an additional 1µl of mobile phase passing through the loop in 1min before switching the loop off-line. A 5min gradient is then performed with an additional minute to allow for ramping and washing the column in a high concentration

of organic mobile phase. The total run time including sample injection is 8min equating to 180 samples per day (Fig. 1C, Supp. Fig. 3).

To determine the outcome of utilizing a 17min gradient (50 samples per day) method to increase sample throughput we performed 50 injections of 200ng Hela tryptic digest in addition to running 6 technical replicates of a Hela tryptic digest that was separated into 12 fractions using high pH reversed phase fractionation in a stage-tip format. Comparing the 50 single shot samples to the peptides observed in the 12 fractions and employing the 'matching between runs' feature in MaxQuant, we identified a median of 37,234 unique peptide sequences (cumulative total of 65,708) (Fig. 2A, Supp. Fig. 4A) and 6,690 proteins (cumulative total 7,797)(Fig. 2B, Supp. Fig. 4B) in each single shot replicate. When we analyzed cumulative peptide and protein identifications across all fractions, we observed a median of 51,249 unique peptide sequences (cumulative total of 67,565)(Fig. 2C, Supp. Fig 5) and 7,407 proteins (cumulative total of 7,850)(Fig. 2D, Supp. Fig. 6). We repeated this process using our 5min gradient (180 samples per day) by performing 180 injections of Hela tryptic digest in addition to running 12 reversed phase high pH fractions of a Hela tryptic digest to match to. From these samples we observe a median of 11,901 unique peptide sequences (cumulative total of 26,600)(Fig. 2A, Supp. Fig. 7A) and 2,858 proteins (cumulative total of 4,254)(Fig. 2B, Supp. Fig. 7B) in each single shot replicate. When we analyzed the cumulative peptide and protein identifications across all fractions, we observed 25,418 unique peptide sequences (Fig. 2C, Supp. Fig 8) and 4,354 proteins (Fig. 2D, Supp. Fig. 9). The identification numbers in the 5min gradient samples fell over time and we were able to attribute this to loss of peptides to the surface of the sample vial due to the dilute nature of the sample that was being injected (Supp. Fig. 9A and B). The identification rate recovered when a fresh sample was placed into the autosampler. Together, these results demonstrate that by utilizing fractions to create a peptide library to match identifications to, we can achieve high numbers of peptide and protein identifications using short sample gradients.

We next determined the effect of reduced gradient lengths and column configurations on peak characteristics. The median full width at half maximum (FWHM) of peaks on a 90min gradient was 10.1 seconds over three replicates. Shorter gradient lengths led to a reduction in the median FWHM peak with to 3.7 and 2.2 seconds for the 17min gradient and 5min gradients, respectively (Fig. 3A). Overlaying the chromatographic traces from multiple replicates demonstrated high inter-replicate reproducibility with very little variation observed between runs for both the 17min gradient (Fig. 3B) and the 5min gradient (Fig. 3C). To demonstrate robust performance over a larger sample set we mapped the retention time for multiple peptides from different protein groups across all replicates. Results from the 17min and 5min gradients demonstrated very little divergence in retention time across the complete datasets (Fig. 4A and B).

Having demonstrated an ability to achieve high numbers of identifications in a common standard laboratory sample, we next sought to analyze a sample with clinical relevance. Blood plasma is widely used for clinical diagnostics and is a rich source of potential biomarkers across a broad range of diseases. A single sample of Plasma was collected, processed and tryptically digested before being injected 180 times using the 5min gradient method. The same sample was also separated into 12 fractions using high pH revered phase fractionation in a stage-tip format to generate a peptide library. This analysis successfully identified a median of 1,119 unique peptide sequences (cumulative total 2,151)(Supp. Fig. 10) and 213 proteins (cumulative total 454)(Fig. 5) per sample. Once again due the low concentration of peptides in the sample

we observed a time dependent loss of peptide identifications. These results demonstrate that the high-throughput method can be utilized to characterize clinically relevant samples.

Discussion

A key limitation to DDA based shotgun proteomics is the combination of stochastic feature sampling and an ability to generate MS2 spectra intense enough to confidently make the identification (effective sequencing speed), generally causing large numbers of missing values across sample sets. Accurate matching from a library of identified MS1 features ('Accurate Mass Tagging' and 'match between runs'[10,11] has widely been implemented over the last decade even though stringent control of false discovery rates has yet to be fully implemented. Here we demonstrate the effectiveness of combining high peak capacity chromatography on short gradients with the timsTOFs speed and CCS enhanced selectivity to provide an extremely high yielding proteomics discovery workflow.

The high speed and sensitivity of the timsTOF Pro utilizing the PASEF mode of operation has been clearly demonstrated previously[7]. However, workflows that fully take advantage of the reduction in cycle times, increased sequencing speed and CCS assisted matching has yet to ascertained. Here we evaluate the ability of the PASEF shotgun proteomics workflow, utilizing rapid gradients on high peak capacity packed emitter columns. By combining these short gradients with the recently implemented CCS enhanced library matching in MaxQuant[9] with stage-tip high pH reversed phase fractionation, we demonstrate very high identification rates for short gradients with complex peptide mixtures. Thus, the combination of high pH library generation, additional feature matching selectivity using CCS values and high peak capacity chromatography provides a deep and very high through-put label free quantification workflow. Interestingly, the level of proteins matched across the individual samples nears the level of proteins found across the 12 stage-tip fractions, suggesting that increasing the number of off-line high pH fractions could potentially improve the performance of shotgun matching depth even further. Further, the discrepancy of unique peptide counts between the individual technical replicates and the cumulative total count suggests that improvements to MS1 acquisition parameters and more advanced matching strategies that incorporate FDR based matched cut-offs may also further improve the data completeness. The rapid acquisition of 12 high pH fractions in 6 hours (for 50 samples per day method) enables routine high pH fractionation for each experiment. This depth of sequencing, speed and reported instrument robustness[7] now provides a very appealing setup for clinical analysis, where historically, numbers of patient samples have been limited due to the low through-put nature of long gradient LC-MS analysis, generally limiting the quality and reproducibility of the results[12]. Higher throughput and larger numbers of patient samples will undoubtedly result in greater robustness of potential identified markers and has the additional benefit of more reliable investigation of experimental and patient confounders[13].

The performance obtained here is accessible as it uses a defined series of available methods, analysis using freely available software and commercially available columns. The methods described here do not require a specialized skill set and are simple to implement making this accessible to a broad range of proteomics research laboratories. Challenges that remain include software analysis time (which needs to be improved by a factor of 2-3 to allow laboratories to operate inside of the dreaded time-debt scenario, where the time it takes to feature detect and search the data takes longer than the acquisition time). Additionally, we observed losses of peptides, presumably to the sidewalls of the MS vials in the autosampler within a day of sample acquisition, likely due to the relatively low concentration samples. This has been previously

reported and needs to be investigated further[14,15]. Finally, the ultra-fast acquisition of proteomes in the context of known physical limitations of MS dynamic range strongly suggests future strategies for deep proteome analysis will likely involve high pH reversed phase fractionation and even more rapid acquisition modes.

Acknowledgements
This work was supported by operational infrastructure grants through the Australian Government IRISS and the Victorian State Government OIS 9000220.

Conflict of Interest statement
JJS, GI and AIW state that they have potential conflicts of interest regarding this work as they are employees of Ion Opticks.

Figure legends

Figure 1.
Schematics of UHPLC methods. A) 90min sample gradient (11 samples per day). Total method time 128.8min. B) 16.8min sample gradient (50 samples per day). Total method time 28.5min. C) 5min sample gradient (180 samples per day). Total method time 8min.

Figure 2.
High numbers of peptide and protein identifications achieved from short gradients. A and B) Unique peptide sequence and unique protein group identifications from Hela tryptic digest using 50 (17min gradient; 200ng injection, n= 50) and 180 (5min gradient; 80ng injection, n=180) samples per day single shot runs. Single shot samples were matched to 12 high pH reversed phase fractions of the same samples. C and D) Cumulative total of 12 fractions used in A and B. 50 samples per day, n = 6; 180 samples per day n = 1. All dots represent individual replicate values, bars represent the median value across replicates and dashed line indicates the cumulative total identifications across all runs.

Figure 3.
Chromatographic peak characteristics from short gradient methods. A) Box plot of identified peptides full width at half maximum (FWHM) in seconds for column length and gradient duration combination. Three replicates of 200ng Hela tryptic digest injections shown. Line indicates median FWHM for each sample. Outliers were omitted from the plot. B and C) Comparison of three base peak chromatograms from a HeLa tryptic digest (200ng injection) for the 17 (B) and 5 (C) minute gradients.

Figure 4.
Short gradient methods allow reproducible analysis of samples. A and B) Retention time stability of selected peptides from 200ng (17min, 50 samples) or 80ng (5min, 180 samples) injections of a Hela tryptic digest that were identified across all samples.

Figure 5.
Short gradients facilitate analysis of Plasma samples. Unique protein group identifications from Plasma digest using 180 (5min gradient; 50ng injection, n= 180) samples per day single shot runs. Single shot samples were matched to 12 high pH reversed phase fractions of the same

sample. Dashed line indicates point in sample list where the sample vial in autosampler was replaced with a freshly thawed aliquot.

Supplementary Figure 1.
Schematic of a single 11 samples per day (90min gradient) method. A) Trace of Pump A pressure during run. B) Trace of percentage of buffer B during run. C) Total Ion Chromatogram plot across gradient. D) Base Peak Chromatogram across gradient.

Supplementary Figure 2.
Schematic of a single 50 samples per day (17min gradient) method. A) Trace of Pump A pressure during run. B) Trace of percentage of buffer B during run. C) Total Ion Chromatogram plot across gradient. D) Base Peak Chromatogram across gradient.

Supplementary Figure 3.
Schematic of 15 X 180 samples per day (5min gradient) method. A) Trace of Pump A pressure during runs. B) Trace of percentage of buffer B during runs. C) Total Ion Chromatogram plot across gradients. D) Base Peak Chromatogram across gradients.

Supplementary Figure 4.
Unique peptide sequence and unique protein group identifications from Hela tryptic digest using 50 (17min gradient; 200ng injection, n= 50) samples per day single shot runs. Single shot samples were matched to 12 high pH reversed phase fractions of the same samples.

Supplementary Figure 5.
Unique peptide sequence identifications from Hela tryptic digest using 50 (17min gradient; n= 6) samples per day analysis of 12 high pH reversed phase fractions. Far right bar indicates cumulative total of 12 fractions.

Supplementary Figure 6.
Unique protein group identifications from Hela tryptic digest using 50 (17min gradient; n= 6) samples per day analysis of 12 high pH reversed phase fractions. Far right bar indicates cumulative total of 12 fractions.

Supplementary Figure 7.
Unique peptide sequence and unique protein group identifications from Hela tryptic digest using 180 (5min gradient; 80ng injection, n= 180) samples per day single shot runs. Single shot samples were matched to 12 high pH reversed phase fractions of the same samples.

Supplementary Figure 8.
Unique peptide sequence identifications from Hela tryptic digest using 180 (5min gradient; n= 1) samples per day analysis of 12 high pH reversed phase fractions. Far right bar indicates cumulative total of 12 fractions.

Supplementary Figure 9.
Unique protein group sequences identifications from Hela tryptic digest using 180 (5min gradient; n= 1) samples per day analysis of 12 high pH reversed phase fractions. Far right bar indicates cumulative total of 12 fractions.

Supplementary Figure 10.

Unique peptide sequence identifications from Plasma digest using 180 (5min gradient; 50ng injection, n= 180) samples per day single shot runs. Single shot samples were matched to 12 high pH reversed phase fractions of the same samples. Dashed line indicates point in sample list where the sample vial in autosampler was replaced with a freshly thawed aliquot.

Materials and Methods

Hela tryptic digest. Hela cell tryptic peptides were obtained from the commercially prepared Pierce HeLa Protein Digest Standard (Thermo Fisher). Each vial was reconstituted in 2% acetonitrile (ACN)/1% formic acid (FA) in MilliQ water to a concentration of 200ng/µl before being aliquoted and frozen at -80°C prior to analysis.

Plasma collection and digestion. Blood was collected into tubes containing EDTA and centrifuged for 10 min at 1800 g, supernatant transferred to a new tube and centrifuged again for 15 min at 2000 g to harvest plasma. Blood was sampled from a healthy donor, who provided written informed consent, with prior approval of the ethics committee of the Walter and Eliza Hall Institute. The sample was digested using the SP3 protocol as described by Hughes et al.[16] with some modifications. A 1:1 combination mix of two of magnetic carboxylate beads was used (Sera-Mag Speed beads, #45152105050250, #65152105050250, GE Healthcare). Beads were prepared fresh by rinsing with water three times prior to use at a stock concentration of 20 µg/µL. The plasma sample (2.5µl) was simultaneously reduced and alkylated in a buffer containing 10 mM Tris HCl pH 7.4/10 mM Tris (2-Carboxyethyl) phosphine (TCEP)/5.5 mM 2-chloracetamide (2-CAA) by heating at 95°C for 10 mins. Carboxylate beads (4 µl) were added to the sample with ACN (70% final concentration v/v) and incubated at RT for 18 mins. Samples were placed on a magnetic rack (Ambion, Thermo Fisher Scientific), supernatant discarded, and the beads washed twice with 70% ethanol and once with neat ACN. ACN was completely evaporated from the tube using a CentriVap (Labconco) prior to the addition of 80µl digestion buffer (10% 2,2,2-Trifluoroethanol (TFE)/100 mM NH_4HCO_3) containing 4µg Trypsin-gold (Promega, V5280) and 4µg Lys-C (Wako, 129-02541). Enzymatic digestion proceeded for 1 hr at 37 °C using the ThermoMixer C (Eppendorf) shaking at 400 rpm. Following the digest, sample was placed on a magnetic rack and the supernatant containing peptides was collected and an additional elution (50µl) was performed using 2% dimethyl sulfoxide (DMSO, Sigma) prior sonication in a water bath for 1 min. Sample was lyophilised to dryness using a CentriVap (Labconco) before reconstitution in 250µl 2% ACN, 1% FA and frozen in 3 aliquots prior to analysis.

High-pH fractionation. 50µg of peptides from either digested plasma or 20µg of a Hela tryptic digest (Pierce, Thermo Fisher) were resuspended in 10mM Ammonium Formate pH 10. Peptides were separated into 12 fractions using a stage-tip containing 4 X C18 plugs. The stage-tips were activated using isopropanol, washed with 60% ACN in 10mM Ammonium Formate pH 10 and re-equilibrated using 10mM Ammonium Formate pH 10. Samples were then loaded onto the stage-tips, washed twice using 10mM Ammonium Formate pH 10 and eluted into fractions using an escalating concentration of ACN in 10mM Ammonium Formate pH 10 (2.75, 3.75, 5, 6, 7, 8, 9, 10, 13, 17.5, 25, 60% ACN). Fractions were lyophilised to dryness using a CentriVap (Labconco) before reconstitution in 2% ACN, 1% FA prior to analysis.

11 and 50 samples per day UHPLC settings. Samples were analyzed on a nanoElute (plug-in V1.1.0.27 ; Bruker, Germany) coupled to a timsTOF Pro (Bruker) equipped with a CaptiveSpray source. Peptides were separated on a 15cm X 75µm analytical column, 1.6µm C18 beads with a packed emitter tip (IonOpticks, Australia). The column temperature was

maintained at 50°C using an integrated column oven (Sonation GmbH, Germany). The column was equilibrated using 4 column volumes before loading sample in 100% buffer A (99.9% MilliQ water, 0.1% FA) (Both steps performed at 980bar). For the 11 samples per day method, samples were separated at 400nl/min using a linear gradient from 2% to 25% buffer B (99.9% ACN, 0.1% FA) over 90min before ramping to 37% buffer B (10min), ramp to 80% buffer B (10min) and sustained for 10min (total separation method time 120min). For the 50 samples per day method, samples were separated at 400nl/min using a linear gradient from 5% to 30% buffer B (99.9% ACN, 0.1% FA) over 16.8min before ramping to 95% buffer B (0.5min) and sustained for 2.4min (total separation method time 19.7min).

180 samples per day UHPLC settings. Samples were analyzed on a M-class (Waters, USA) coupled to a timsTOF Pro (Bruker) equipped with a CaptiveSpray source. Peptides were separated on a 5cm X 150µm analytical column, 1.6µm C18 beads with a packed emitter tip (IonOpticks, Australia) using a constant flow rate of 2µl/min. The column was maintained at room temperature. Sample was injected into a sample loop which takes approximately 0.5min. Mobile phase at 100% buffer A continues to flow over the analytical column during this period facilitating column equilibration. The sample loop was switched on-line for 1min at 100% buffer A. A linear gradient begins at 1.2min from 5% to 34% buffer B over 5min before ramping to 80% buffer B (0.5min) and sustained for 0.3min. Mobile phase is then ramped back to 100% buffer A (0.2min) and sustained for 0.3min (this period also contributes to column equilibration) (total method time 7.5min + 0.5min injection).

timsTOF Pro settings. The timsTOF Pro (Bruker) was operated in PASEF mode using Compass Hystar 5.0.36.0. Settings for the 11 samples per day method were as follows: Mass Range 100 to 1700m/z, 1/K0 Start 0.6 V·s/cm^2 End 1.6 V·s/cm^2, Ramp time 110.1ms, Lock Duty Cycle to 100%, Capillary Voltage 1600V, Dry Gas 3 l/min, Dry Temp 180°C, PASEF settings: 10 MS/MS scans (total cycle time 1.27sec), charge range 0-5, active exclusion for 0.4 min, Scheduling Target intensity 10000, Intensity threshold 2500, CID collision energy 42eV. Settings for the 50 and 180 samples per day method were as follows: Mass Range 100 to 1700m/z, 1/K0 Start 0.85 V·s/cm^2 End 1.3 V·s/cm^2, Ramp time 100ms, Lock Duty Cycle to 100%, Capillary Voltage 1600V, Dry Gas 3 l/min, Dry Temp 180°C, PASEF settings: 4 MS/MS scans (total cycle time 0.53sec), charge range 0-5, active exclusion for 0.4 min, Scheduling Target intensity 24000, Intensity threshold 2000, CID collision energy 42eV.

Raw data processing and analysis. All raw files were analyzed by MaxQuant v1.6.6.0 software using the integrated Andromeda search engine. Experiment type was set as TIMS-DDA with no modification to default settings. Data was searched against the human Uniprot Reference Proteome with isoforms (downloaded March 2019) and a separate reverse decoy database using a strict trypsin specificity allowing up to 2 missed cleavages. The minimum required peptide length was set to 7 amino acids. Modifications: Carbamidomethylation of Cys was set as a fixed modification, while N-acetylation of proteins and oxidation of Met were set as variable modifications. First search peptide tolerance was set at 70ppm and main search set at 30ppm (other settings left as default). Single shot samples and fractions were assigned as separate parameter groups and matching between runs was turned on and set as “from and to” for single shot samples and “from” for fractions. Maximum peptide mass [Da] was set at 8000. All other settings in group or global parameters were left as default for MaxQuant v1.6.6.0.

References

1 Michalski, A., Cox, J. & Mann, M. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *Journal of proteome research* **10**, 1785-1793, doi:10.1021/pr101060v (2011).

2 Beck, S. *et al.* The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics. *Molecular & cellular proteomics : MCP* **14**, 2014-2029, doi:10.1074/mcp.M114.047407 (2015).

3 Kanu, A. B., Dwivedi, P., Tam, M., Matz, L. & Hill, H. H., Jr. Ion mobility-mass spectrometry. *Journal of mass spectrometry : JMS* **43**, 1-22, doi:10.1002/jms.1383 (2008).

4 Cumeras, R., Figueras, E., Davis, C. E., Baumbach, J. I. & Gracia, I. Review on ion mobility spectrometry. Part 2: hyphenated methods and effects of experimental parameters. *The Analyst* **140**, 1391-1410, doi:10.1039/c4an01101e (2015).

5 Ridgeway, M. E. *et al.* Gated Trapped Ion Mobility Spectrometry Coupled to Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *International journal for ion mobility spectrometry : official publication of the International Society for Ion Mobility Spectrometry* **19**, 77-85, doi:10.1007/s12127-016-0197-0 (2016).

6 Meier, F. *et al.* Parallel Accumulation-Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device. *Journal of proteome research* **14**, 5378-5387, doi:10.1021/acs.jproteome.5b00932 (2015).

7 Meier, F. *et al.* Online Parallel Accumulation-Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. *Molecular & cellular proteomics : MCP* **17**, 2534-2545, doi:10.1074/mcp.TIR118.000900 (2018).

8 Bekker-Jensen, D. B. *et al.* An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes. *Cell systems*, doi:10.1016/j.cels.2017.05.009 (2017).

9 Prianichnikov, N. *et al.* MaxQuant software for ion mobility enhanced shotgun proteomics. *bioRxiv*, 651760, doi:10.1101/651760 (2019).

10 Zimmer, J. S., Monroe, M. E., Qian, W. J. & Smith, R. D. Advances in proteomics data analysis and display using an accurate mass and time tag approach. *Mass spectrometry reviews* **25**, 450-482, doi:10.1002/mas.20071 (2006).

11 Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & cellular proteomics : MCP* **13**, 2513-2526, doi:10.1074/mcp.M113.031591 (2014).

12 Skates, S. J. *et al.* Statistical design for biospecimen cohort size in proteomics-based biomarker discovery and verification studies. *Journal of proteome research* **12**, 5383-5394, doi:10.1021/pr400132j (2013).

13 Geyer, P. E., Holdt, L. M., Teupser, D. & Mann, M. Revisiting biomarker discovery by plasma proteomics. *Molecular systems biology* **13**, 942, doi:10.15252/msb.20156297 (2017).

14 Maes, K. *et al.* Improved sensitivity of the nano ultra-high performance liquid chromatography-tandem mass spectrometric analysis of low-concentrated neuropeptides by reducing aspecific adsorption and optimizing the injection solvent. *Journal of chromatography. A* **1360**, 217-228, doi:10.1016/j.chroma.2014.07.086 (2014).

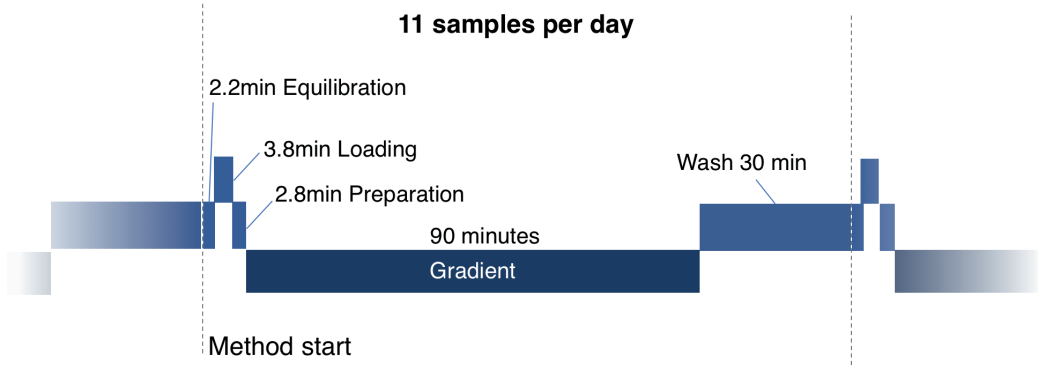
15 Warwood, S., Byron, A., Humphries, M. J. & Knight, D. The effect of peptide adsorption on signal linearity and a simple approach to improve reliability of

quantification. *Journal of proteomics* **85**, 160-164, doi:10.1016/j.jprot.2013.04.034 (2013).

16 Hughes, C. S. *et al.* Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nature protocols* **14**, 68-85, doi:10.1038/s41596-018-0082-x (2019).

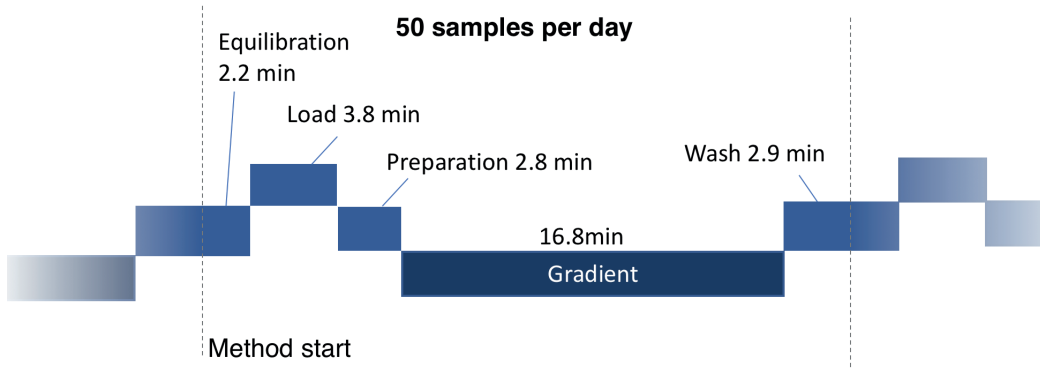
Figure 1

A.



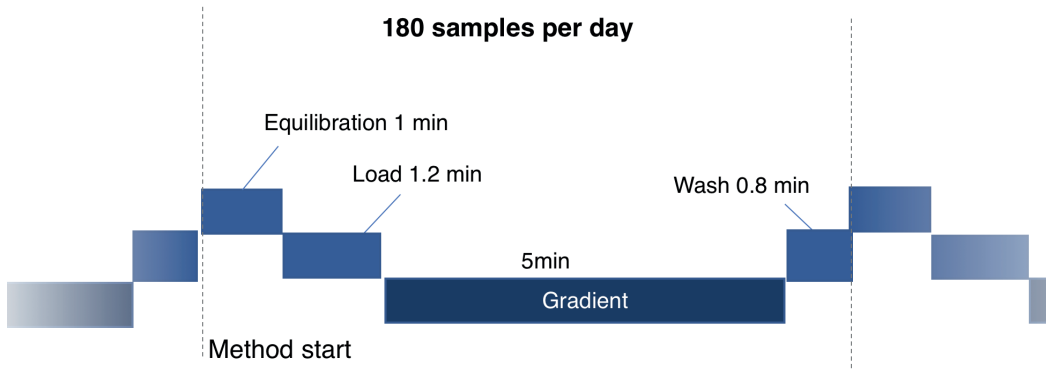
Column configuration: 15cm length X 75µm inner diameter, 1.6µm C18 resin.

B.



Column configuration: 15cm length X 75µm inner diameter, 1.6µm C18 resin.

C.



Column configuration: 5cm length X 150µm inner diameter, 1.6µm C18 resin.

Figure 2

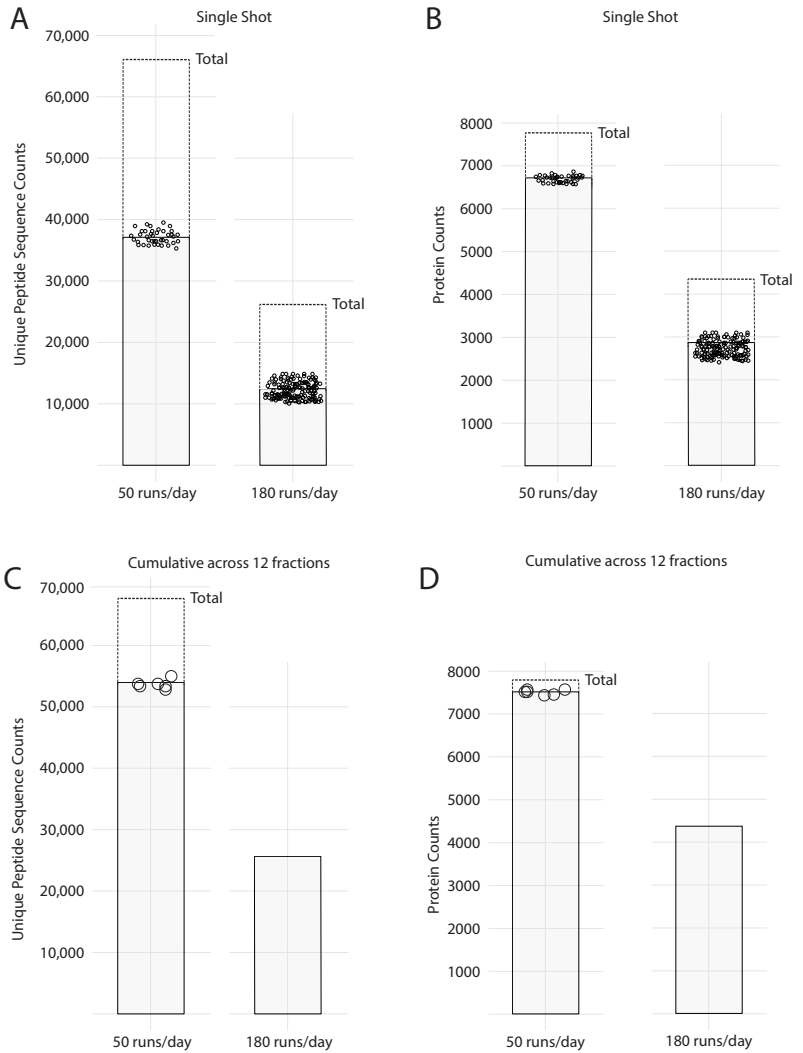
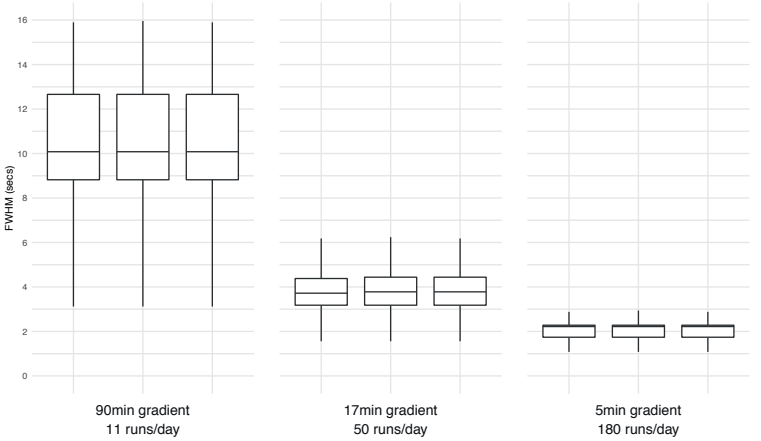
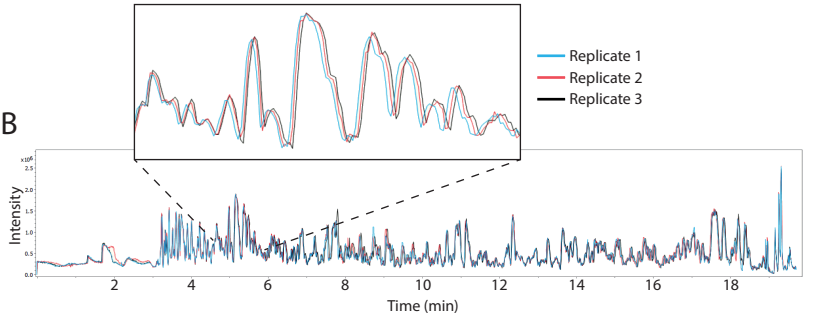

A
Single Shot
Unique Peptide Sequence Counts
70,000
60,000
50,000
40,000
30,000
20,000
10,000
Total
Total
50 runs/day
180 runs/day
B
Single Shot
Protein Counts
8000
7000
6000
5000
4000
3000
2000
1000
Total
Total
50 runs/day
180 runs/day
C
Cumulative across 12 fractions
Unique Peptide Sequence Counts
70,000
60,000
50,000
40,000
30,000
20,000
10,000
Total
50 runs/day
180 runs/day
D
Cumulative across 12 fractions
Protein Counts
8000
7000
6000
5000
4000
3000
2000
1000
Total
50 runs/day
180 runs/day

Figure 3

A


FWHM (secs)
16
14
12
10
8
6
4
2
0
90min gradient
11 runs/day
17min gradient
50 runs/day
5min gradient
180 runs/day

B


Replicate 1
Replicate 2
Replicate 3
Intensity
2
4
6
8
10
12
14
16
18
Time (min)

C

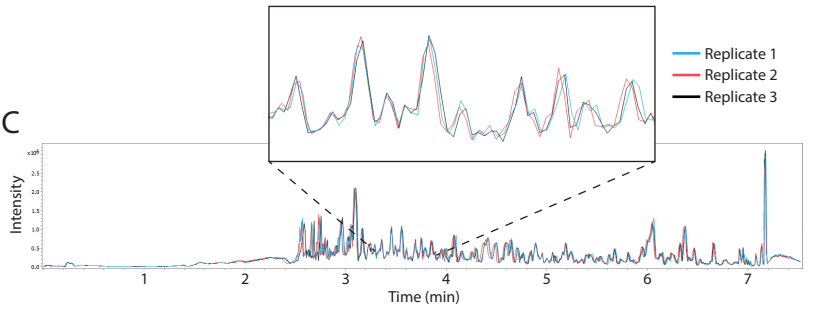
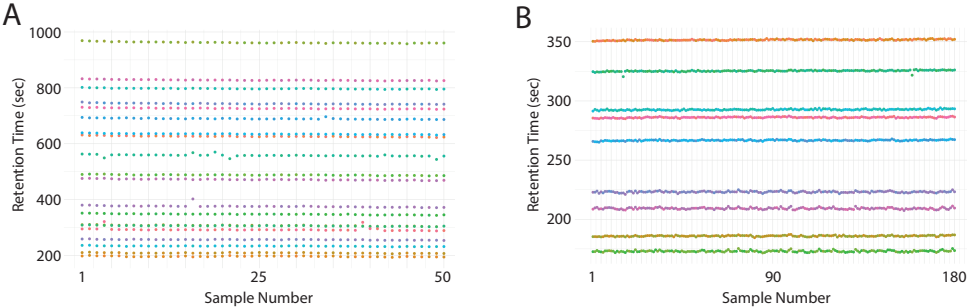

Replicate 1
Replicate 2
Replicate 3
Intensity
1
2
3
4
5
6
7
Time (min)

Figure 4


A
1000
800
600
400
200
Retention Time (sec)
1
25
50
Sample Number
B
350
300
250
200
Retention Time (sec)
1
90
180
Sample Number

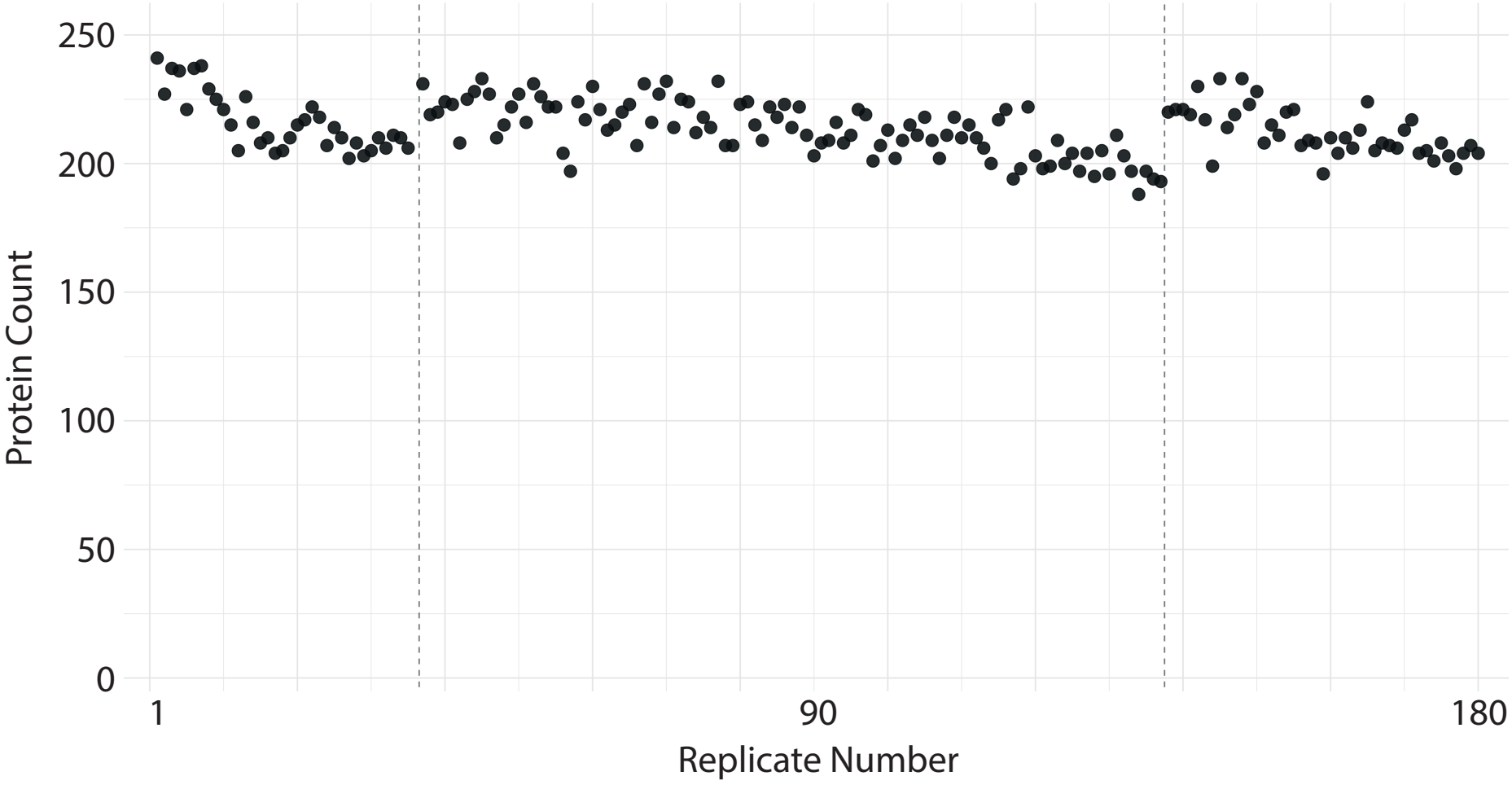



Figure 5