LFQ Benchmark Dataset - Generation Beta: Assessing Modern Proteomics Instruments and Acquisition Workflows with High-Throughput LC Gradients

Bart Van Puyvelde[1†], Robbe Devreese[2,3†], Cristina Chiva[4,5], Eduard Sabidó[4,5], Sibylle Pfammatter[6], Christian Panse[6,7], Jeewan Babu Rijal[8,9], Charline Keller[8,9], Ihor Batruch[10], Patrick Pribil[10], Jean-Baptiste Vincendet[10], Frédéric Fontaine[11], Lars Lefever[12,13], Pedro Magalhães[12,14], Dieter Deforce[1], Paolo Nanni[6], Bart Ghesquière[12,13,14], Yasset Perez-Riverol[15], Lennart Martens[2,3,8,9], Christine Carapito[8,9], Robbin Bouwmeester[2,3†],Maarten Dhaenens[1†,*]

1. ProGenTomics, Laboratory of Pharmaceutical Biotechnology, Ghent University, 9000 Ghent, Belgium
2. VIB-UGent Center for Medical Biotechnology, VIB, 9052 Ghent, Belgium
3. Department of Biomolecular Medicine, Ghent University, 9052 Ghent, Belgium
4. Centre for Genomic Regulation, The Barcelona Institute of Science and Technology (BIST), Dr Aiguader 88, 08003 Barcelona, Spain
5. Univeristat Pompeu Fabra, Dr. Aiguader 88, 08003 Barcelona, Spain.
6. Functional Genomics Center Zurich (FGCZ) - University of Zurich/ETH Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
7. Swiss Institute of Bioinformatics (SIB), Quartier Sorge - Batiment Amphipole, CH-1015 Lausanne, Switzerland
8. BioOrganic Mass Spectrometry Laboratory (LSMBO), IPHC UMR 7178, University of Strasbourg, CNRS, 67087 Strasbourg, France
9. Infrastructure Nationale de Protéomique ProFI – UAR 2048, Strasbourg, 67087, France
10. Sciex, Concord, Ontario, Canada
11. Thermo Fisher Scientific, Paris, France
12. Laboratory of Applied Mass Spectrometry, Department of Cellular and Molecular Medicine, KU Leuven, Leuven, 3000, Belgium.
13. Metabolomics Expertise Center, Center for Cancer Biology VIB, Leuven, 3000, Belgium
14. VIB-KU Leuven Center for Brain & Disease Research Technologies, Proteomics Unit, Leuven, Belgium
15. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, UK

[†] These authors contributed equally
*Corresponding author: Maarten Dhaenens (maarten.dhaenens@ugent.be)

Abstract

Recent advances in liquid chromatography–mass spectrometry (LC-MS) have accelerated the adoption of high-throughput workflows that deliver deep proteome coverage using minimal sample amounts. This trend is largely driven by clinical and single-cell proteomics, where sensitivity and reproducibility are essential. Here, we extend our previous benchmark dataset (PXD028735) using next-generation LC-MS platforms optimized for rapid proteome analysis. We generated an extensive DDA/DIA dataset using a human-yeast-*E. coli* hybrid proteome. The proteome sample was distributed across multiple laboratories together with standardized analytical protocols specifying two short LC gradients (5 and 15 min) and low sample input amounts. This dataset includes data acquired on four different platforms, and features new scanning quadrupole-based implementations, extending coverage across different instruments and acquisition strategies. Our comprehensive evaluation highlights how technological advances and reduced LC gradients may affect proteome depth, quantitative precision, and cross-instrument consistency. The release of this benchmark dataset via ProteomeXchange (PXD070049 and PXD071205), allows for the acceleration of cross-platform algorithm development, enhance data mining strategies, and supports standardization of short-gradient, high-throughput LC-MS-based proteomics.

Background & Summary

Proteomics has rapidly evolved into a cornerstone of molecular biology, providing unparalleled insight into protein function, interactions, and regulation. While DNA and RNA measurements reveal what a cell is capable of, proteomics captures what it is actually doing—reflecting the dynamic protein complement and post-translational modifications (PTMs) that define cellular function and phenotype.

The complexity of the proteome, shaped by alternative splicing, variable translation rates, and diverse PTMs, necessitates analytical approaches capable of resolving this intricate molecular landscape with both depth and precision. Liquid Chromatography coupled to Mass spectrometry (LC-MS)-based proteomics has emerged as the leading technology for large-scale protein analysis, offering an unbiased approach to protein identification and quantification [1]. Advancements in mass spectrometry instrumentation, including Zeno trapping, next-generation time-of-flight (TOF) analyzers such as the multi-reflectron TOF, and emerging hybrid platforms like the Orbitrap Astral, have significantly enhanced sensitivity, acquisition speed, and dynamic range [2–6]. These advancements have propelled proteomics from bulk tissue analysis toward increasingly challenging applications, such as clinical proteomics, single-cell proteomics and spatially resolved protein mapping [7].

While data-dependent acquisition (DDA) has been the cornerstone for many years, it has now been taken over by data-independent acquisition (DIA) due to its increased reproducibility and sensitivity [8,9]. DDA relies on sequential selection of precursors for fragmentation and identification, but this approach is inherently limited by sampling bias. In contrast, DIA captures all fragment ions across a defined mass range, overcoming selection bias and enabling more comprehensive proteome coverage. These gains come with a trade-off, as DIA generates complex data that is difficult to interpret. In response, the field has rapidly developed advanced computational workflows capable of handling these complexities [10]. Machine learning has rapidly emerged as a powerful method to predict the behavior of peptides, including their fragmentation spectra, retention times and collisional cross-sections [11–18]. These predicted properties can then be incorporated into spectral libraries to expand coverage and improve the accuracy of peptide identification and quantification [19,20].

Additional separation dimensions, such as ion mobility, like TIMS and FAIMS, have been introduced to enhance resolution and reduce spectral overlap [21–24]. Alternatively, some have implemented scanning quadrupole based acquisition, i.e. SONAR [25], Scanning SWATH [26] or ZT Scan DIA [27] and diagonalPASEF [28] (midia-PASEF [29] and SynchroPASEF [30]), providing an extra dimension to alleviate the spectral complexity by sampling a sliding quadrupole selection window with a very fast TOF detector. Others have implemented the usage of narrow-window DIA (nDIA) to increase the specificity [6].

A key driver behind these innovations is the growing field of single-cell proteomics, wherein the objective is to profile the proteome at the resolution of a single cell. This field specifically must address the technical challenges associated with analyzing ultra-low amounts of material [31]. The demand for measuring thousands of proteins from a single cell has pushed instruments toward higher speed and sensitivity [32]. To meet the high-throughput requirements of single-cell studies, workflows increasingly use extremely short LC gradients. By minimizing sample carryover and leveraging low flow rates, these ultra-fast chromatographic workflows allow simultaneous detection of a large number of proteins from sub-nanogram sample inputs. During method development, many groups work with “single-cell equivalent” inputs, commonly ~250 pg, to emulate the average protein content of a single cell [33]. These developments made to meet the demands of single-cell proteomics underscore the importance of combining accelerated LC separations, minimal sample requirements, and advanced DIA [34,35].

The shift from DDA to DIA, together with increasingly complex acquisition strategies, has expanded the analytical possibilities of proteomics, yet equally magnifies the need for standardized benchmarks. In particular, the rise of single-cell proteomics has reinforced the importance of shared datasets that enable consistent evaluation of workflows across diverse experimental contexts. Our previous LFQ benchmark study (PXD028735) has become a widely used community resource, with 145,147 PRIDE downloads since 2022, placing it within the top 1% of PRIDE datasets by downloads. Therefore, we present a second-generation LFQ benchmark study that builds upon our previous dataset by evaluating the performance of the latest generation MS instruments, i.e. SCIEX ZenoTOF 7600+ system & ZenoTOF

8600 system, Bruker Ultra 2, and Thermo Orbitrap Astral, on recent DIA, narrow-window DIA, and scanning quadrupole DIA methodologies (diagonal-PASEF, ZT Scan DIA) under conditions relevant to bulk as well as single-cell proteomics [36]. We adapt the widely used benchmark experimental design employing a triple proteome-mixture [37], and include adequate technical replicates in both nano and capillary flow LC setups. In line with our previous benchmark, our aim is not to maximize peptide and protein identifications but to create a comprehensive, community-driven resource for the proteomics field. This dataset is intended to facilitate transparent comparisons of algorithms and workflows and already contributes to the ongoing development of community initiatives such as ProteoBench [38], helping to shift the focus beyond simple metrics of identification counts toward more rigorous assessment of quantitative accuracy and precision.

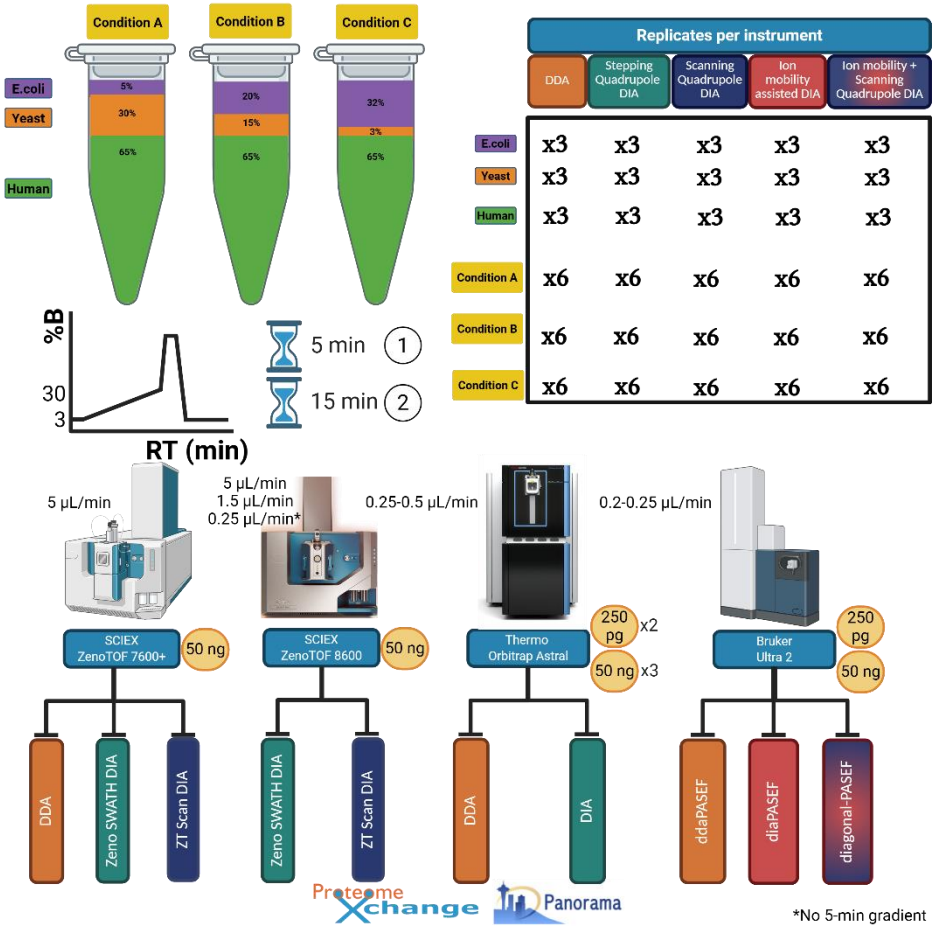


Figure 1. Schematic overview of Acquisition Strategies and Instrumentation. A comprehensive LC-MS dataset was generated using commercial Human K562, Yeast, and *Escherichia coli* (*E. coli*) proteome digests. To simulate experimental ratio differences, three hybrid proteome samples (A, B, and C) containing defined proportions of tryptic peptides from Human, Yeast, and *E. coli* were prepared. Both the individual proteome digests and the hybrid mixtures were analyzed using two short LC gradient lengths (5- and 15-minutes active separation) to assess the impact of chromatographic speed on data quality and proteome coverage. Data acquisition was performed in both Data-Dependent Acquisition (DDA) and Data-Independent Acquisition (DIA) modes across four advanced LC-MS platforms: SCIEX ZenoTOF 7600+ and 8600, Thermo Scientific Orbitrap Astral (2 sites), and Bruker timsTOF Ultra 2. Each commercial lysate was measured individually in triplicate, while the hybrid mixtures were analyzed in six replicates per gradient and acquisition method, ensuring robust statistical comparisons. The complete dataset is publicly available via ProteomeXchange under the identifiers PXD070049 and PXD071205, providing a valuable resource for the proteomics community. Additionally, system suitability monitoring was integrated into the workflow using commercial *E. coli* lysate digest. These QC samples were acquired in DDA at multiple time points throughout the batch on each instrument. The AutoQC workflow automatically imported QC data into Skyline and uploaded it to the Panorama AutoQC server via AutoQC Loader, enabling real-time system suitability assessment across all LC-MS platforms used in the study. Figure created in BioRender.

Methods

Sample preparation

Mass spectrometry-compatible protein digest extracts from Human K562 (P/N: V6951, Lot: 0000608712) and Yeast (P/N: V7461, Lot: 0000601330) were obtained from Promega (Madison, Wisconsin, USA), while the lyophilized MassPrep *Escherichia coli* protein digest extract (P/N: 186003196, Lot: W01111913) was sourced from Waters Corporation (Milford, Massachusetts, USA). Each extract underwent reduction with dithiothreitol (DTT), alkylation with iodoacetamide (IAA), and digestion with sequencing grade Trypsin (-Lys C) by the respective manufacturers. The digested protein extracts were reconstituted in a solution of 0.1% formic acid (FA) in 5% acetonitrile (Biosolve B.V, Valkenswaard, The Netherlands). Following the approach of Navarro et al.[37], two master samples, A and B, were prepared as illustrated in Figure 1. Sample A consisted of Human, Yeast, and *E. coli* digests mixed at 65%, 30%, and 5% weight/weight (w/w), respectively. Sample B was formulated by combining Human, Yeast, and *E. coli* protein digests at 65%, 15%, and 20% w/w, respectively. Additionally, a third mixture, Sample C, was created with Human, Yeast, and *E. coli* protein digests at 65%, 3%, and 32% w/w, respectively, to assess the impact of larger ratio differences on analytical performance. Sample injection order was performed using a standardized scheme (see Supplementary Table 1), in which replicates of the same sample were intentionally interspersed rather than measured consecutively to minimize batch effects.

LC-MS

In this section, a detailed description of the different LC-MS/MS parameters is given for each instrumental setup applied to generate this comprehensive dataset. Most of the instruments were operated according to the lab's best practice, i.e. not necessarily the best attainable, but rather most realistic data quality. A second Orbitrap Astral dataset was generated under more refined conditions. The 1.5 µL/min and nano-flow dataset on the ZenoTOF 8600 was generated by the SCIEX demo lab. The sample load was kept consistent across each lab at 50 ng per injection, which is 100-fold lower compared to our previous LFQ benchmark dataset [36]. Additionally, some laboratories measured 250 pg injections to mimic single-cell conditions. Samples were analyzed using two different standard gradients which included a 5-min and a 15-min active separation ramp of 3-30% acetonitrile content as a set criteria followed to the best efforts, with the rest of the method adjusted by the operators to accommodate LCs and workflows variety. Also, acquisitions and results were not meant to either reflect or conflict with the literature. Differences in the absolute number of identified peptides and proteins can be attributed to factors such as LC conditions (flow rate, column), MS instrumentation, operator choices (instrument settings e.g. window size), and search algorithm compatibility; therefore, direct conclusions on MS instrument performance cannot be drawn from this dataset.

SCIEX ZenoTOF 7600+ system (Capillary flow LC – 5 µL/min) – 50 ng

An ACQUITY™ M-Class™ UPLC System (Waters Corporation, Milford, MA) was used and operated in direct injection mode with the SCIEX ZenoTOF 7600+ system (SCIEX, Concord, Ontario, Canada). Mobile phases consisted of 0.1% FA in water (A) and 0.1% FA in acetonitrile (B). Microflow separations were done using a Kinetex XB-C18 (15 cm x 300 µm, 2.6 µm particle size)(Phenomenex, California, USA) column at a flow rate of 5 µL/min and column temperature of 30 °C using the following gradient: i) after an initial hold at 3% B for 1 min, a gradient from 3% to 30% mobile phase B in 5 or 15 min, respectively, was applied ii) ramp to 80% mobile phase B in 1 min. The washing step at 80% mobile phase B lasted 2 min and was followed by an equilibration step at 3% mobile phase B (starting conditions) for 7 min.

An OptiFlow Turbo V ion source was used with a microflow probe (1-10 µL/min electrode) for the microflow experiments, with source parameters as following: GS1: 20 psi, GS2: 60 psi, Curtain Gas: 35 psi, CAD gas: 7 psi, source temperature: 200 °C and Ionspray voltage: 5000V.

Data-Dependent Acquisition

For the 5-min gradient (a cycle time of 0.573 s), MS1 spectra were collected between 400-1200 m/z for 100 ms. The 30 most intense precursors ions with charge states 2-4 that exceeded 300 counts per second were selected for fragmentation, and the corresponding fragmentation MS2 spectra were collected between 140-1750 m/z for 10 ms. After the fragmentation event, the precursor ions were dynamically excluded from reselection for 3 s.

For the 15-min gradient (a cycle time of 0.795 s), MS1 spectra were collected between 400-900 m/z for 100 ms. The 45 most intense precursors ions with charge states 2-4 that exceeded 300 counts per second were selected for fragmentation, and the corresponding fragmentation MS2 spectra were collected between 140-1800 m/z for 10 ms. After the fragmentation event, the precursor ions were dynamically excluded from reselection for 4 s.

Zeno SWATH DIA

The 5 min-Zeno SWATH DIA experiments used 60 variable windows (Supplementary Table 2) spanning the TOF MS mass range 400-900 Da and MS/MS mass range 140-1800 Da, with Zeno trap pulsing turned on, with MS/MS accumulation times of 7 msec. Before each Zeno SWATH DIA MS cycle an additional MS1 survey scan from 400-900 Da was recorded for 50 ms.

The 15min-Zeno SWATH DIA experiments used 65 variable windows (Supplementary Table 3) spanning the TOF MS mass range 400-900 Da and MS/MS mass range 140-1800 Da, with Zeno trap pulsing turned on, with MS/MS accumulation times of 13 msec. Before each Zeno SWATH DIA MS cycle an additional MS1 survey scan from 400-900 Da was recorded for 50 ms.

ZT Scan DIA

For the 5-min ZT Scan DIA experiments, 5 Da Q1 isolation windows were scanned at 750 Da/s to accommodate PWHH values ≤1 s, resulting in an MS2 accumulation time of 6.7 ms. For the 15-min ZT Scan DIA experiments, 10 Da Q1 isolation windows were scanned at 750 Da/s to accommodate PWHH values between 1.1 and 4.5 s, resulting in an MS2 accumulation time of 13.3 ms. For both gradients, the TOF MS and MS/MS mass range was set to 400-900 Da and 140-1750 Da, respectively.

In Sciex OS 3.4, the acquired ZT Scan DIA raw data is calibrated using residual precursor signals and converted to SCIEX's proprietary .wiff and .wiff.scan file formats.

SCIEX ZenoTOF 8600 system – 50 ng

An ACQUITY™ M-Class™ UPLC System (Waters Corporation, Milford, MA) was used and operated in direct injection mode with the SCIEX ZenoTOF 8600 system (SCIEX, Concord, Ontario, Canada). Mobile phases consisted of 0.1% FA in water (A) and 0.1% FA in acetonitrile (B).

5 µL/min

The 5 µL/min separations were done using a Kinetex XB-C18 (15 cm x 300 µm, 2.6 µm particle size)(Phenomenex, California, USA) at a column temperature of 35 °C using the following gradient: i) after an initial hold at 3% B for 1 min, a gradient from 3% to 30% mobile phase B in 5 or 15 min, respectively, was applied ii) ramp to 80% mobile phase B in 1 min. The washing step at 80% mobile

phase B lasted 2 min and was followed by an equilibration step at 3% mobile phase B (starting conditions) for 7 min.

An OptiFlow Pro ion source was used with a microflow probe (1-10 µL/min electrode) for the microflow experiments, with source parameters as following: GS1: 20 psi, GS2: 60 psi, Curtain Gas: 35 psi, CAD gas: 7 psi, source temperature: 200 °C and Ionspray voltage: 5000V.

1.5 µL/min

The 1.5 µL/min separations were done using an Evosep EV1109 (8 cm x 150 µm, 1.5 µm particle size)(Evosep, Odense, Denmark) column at a column temperature of 40 °C using the following gradient: i) after an initial hold at 1% B for 4 min, a gradient from 1% to 30% mobile phase B in 11.5 min, respectively, was applied ii) ramp to 80% mobile phase B in 0.5 min. The washing step at 80% mobile phase B lasted 2 min and was followed by an equilibration step at 1% mobile phase B (starting conditions) for 2 min.

An OptiFlow Pro ion source was used with a microflow probe (1-10 µL/min electrode), with source parameters as following: GS1: 10 psi, GS2: 20 psi, Curtain Gas: 35 psi, CAD gas: 7 psi, source temperature: 100 °C and Ionspray voltage: 5000V.

0.250 µL/min – Refined conditions

The 0.250 µL/min separations were done using an Aurora® XS Ultimate column (25 cm x 75 µm, 1.7 µm particle size)(IonOpticks, Victoria, Australia) at a column temperature of 50 °C using a 15- and 30-min gradient: i) after an initial hold at 3% B for 18 min, a gradient from 3% to 35% mobile phase B in 15 or 30 min, respectively, was applied ii) ramp to 65% mobile phase B in 2 min, followed by ramping to 80% in 1 min. The washing step at 80% mobile phase B lasted 4 min and was followed by an equilibration step at 1% mobile phase B (starting conditions) for 15 min.

An OptiFlow Pro Nano ion source was used with a NanoCal probe (<1 µL electrode), with source parameters as following: GS1: 10 psi, Nano cel temperature: 250 °C and Nano spray voltage: 2100V.

Zeno SWATH DIA

The Zeno SWATH DIA experiments for the 5-min (5 µL/min) and 11.5-min (1.5 µL/min) gradient used 60 variable windows (Supplementary Table 2) spanning the TOF MS mass range 400-900 Da and MS/MS mass range 140-1750 Da, with Zeno trap pulsing turned on, with MS/MS accumulation times of 13 msec. Before each Zeno SWATH DIA MS cycle an additional MS1 survey scan from 400-1500 Da was recorded for 50 ms.

The 15-min (5 µL/min)-Zeno SWATH DIA experiments used 65 variable windows (Supplementary Table 3) spanning the TOF MS mass range 400-900 Da and MS/MS mass range 140-1750 Da, with Zeno trap pulsing turned on, with MS/MS accumulation times of 13 msec. Before each Zeno SWATH DIA MS cycle an additional MS1 survey scan from 400-1500 Da was recorded for 50 ms.

The 15- and 30-min (0.250 µL/min)-Zeno SWATH DIA experiments used 85 variable windows (Supplementary Table 4) spanning the TOF MS mass range 400-900 Da and MS/MS mass range 140-1750 Da, with Zeno trap pulsing turned on, with MS/MS accumulation times of 16 ms (15-min) and 18 ms (30-min). Before each Zeno SWATH DIA MS cycle an additional MS1 survey scan from 400-1500 Da was recorded for 50 ms.

ZT Scan DIA

For the 5-min (5 µL/min) and 11.5-min (1.5 µL/min) ZT Scan DIA experiments, 5.1 Da Q1 isolation windows were scanned at 509 Da/s, resulting in an MS2 accumulation time of 10 ms. For the 15-min (5 µL/min) ZT Scan DIA experiments, 10.3 Da Q1 isolation windows were scanned at 514 Da/s, resulting in an MS2 accumulation time of 20 ms. For both gradients, the TOF MS and MS/MS mass range was set to 400-900 Da and 140-1750 Da, respectively. For the 15- and 30-min ZT Scan DIA experiments at 0.250 µL/min, 5 and 5.3 Da Q1 isolation windows were scanned at 298 and 267 Da/s, respectively, resulting in an MS2 accumulation time of 16.8 and 19.9 ms. For both gradients, the TOF MS and MS/MS mass range was set to 400-1500 Da and 140-1750 Da, respectively.

In Sciex OS 4.0, the acquired ZT Scan DIA raw data was calibrated using residual precursor signals and converted to SCIEX's proprietary .wiff and .wiff.scan file formats.

Thermo Orbitrap Astral (Nano flow LC)

1) 50 ng

A Thermo Orbitrap Astral (Thermo Fisher Scientific, Waltham, Massachusetts, United States) was coupled to a Vanquish Neo LC instrument (Thermo Fisher Scientific). Peptides were loaded directly onto the analytical column and were separated by reversed-phase chromatography using a 50 cm µPAC™ column (Thermo Fischer Scientific), featuring a structured pillar array bed with a 180 µm bed width. Buffer A consisted of 0.1% FA in water, while Buffer B contained 0.1% FA in 80% acetonitrile.

15-min gradient

Chromatographic gradient was initiated at 96% buffer A and 4% buffer B at a flow rate of 750 nL/min during 1 minute. The flow rate was then reduced to 250 nL/min, and the percentage of buffer B was further increased to 40% over 15 minutes.

5-min gradient

Chromatographic gradient was initiated with 96% buffer A and 4% buffer B at a flow rate of 750 nL/min during 1 minute. The flow rate was then reduced to 250 nL/min, and the percentage of buffer B was further increased to 40% over 5 minutes.

DDA

The mass spectrometer was operated in positive ionization mode with data-dependent acquisition, with a full MS scans over a mass range of m/z 380-980 with detection in the Orbitrap at a resolution of 180,000 and a fixed cycle time of 0.5 s. Precursor ion selection width was kept at 2-Th and a normalized collision energy of 30% was used for higher-energy collisional dissociation (HCD) fragmentation. MS2 scan range was set from 145 to 1450 m/z with detection in the Astral and maximum fill time of 2.5 ms. Dynamic exclusion was enabled and set to 10 s.

DIA

The mass spectrometer was operated in positive ionization mode with data-independent acquisition, with a full MS scans over a mass range of m/z 380-980 with detection in the Orbitrap at a resolution of 240,000. In each cycle of data-independent acquisition, 300 windows of 2 Th were used to isolate and fragment all precursor ions from 380 to 980 m/z. A normalized collision energy of 25% was used for HCD fragmentation. MS2 scan range was set from 150 to 2000 m/z with detection in the Astral with a maximum injection time of 3 ms.

2) 50 ng – refined conditions

284 Ion Optics Direct Injection

285 The Orbitrap Astral (Thermo Scientific) was coupled to a Vanquish Neo UHPLC instrument (Thermo
286 Scientific). Peptides were separated by reversed-phase chromatography using an Aurora® Elite XT (15
287 cm x 75 µm) (IonOpticks) column to accommodate both 5 and 15mins separations. Buffer A consisted
288 of 0.1% FA in water, while Buffer B contained 0.1% FA in 80% acetonitrile.

289 Peptides were separated using a 5 and 15-min linear gradient from 3-38%B at a flow rate of 0.55 and
290 0.45 µL/min respectively.

291 DIA

292 The mass spectrometer was operated in positive ionization mode with narrow-window data-
293 independent acquisition, with a full MS scan over a mass range of 400-800 (5-min) and 400-900 m/z
294 (15-min) in the Orbitrap, at a resolution of 180,000 (5-min)| 240,000 (15-min). For the 5-min gradient,
295 133 windows of 3 Th were used to isolate and fragment all precursor ions from m/z 400 to 800, while
296 for the 15-min gradient, 166 windows of 3 Th were used to isolate and fragment all precursor ions from
297 m/z 400 to 900. A normalized collision energy of 25% was used for HCD fragmentation. MS2 scan range
298 was set from 150 to 1700 m/z with detection in the Astral with a maximum injection time of 5 and 3 ms
299 for the 5- and 15-min gradient, respectively. AGC targets were set to 500% (for MS1 scans) and 100%
300 (for MS2 scans) for both methods, the RF lens was set to 45%, and a loop control of 0.4s (for 180,000
301 MS1 res) or 0.6s (for 240,000 MS1 res) ruled the cycle time.

302 *uPAC Neo plus*

303 An Orbitrap Astral (Thermo Scientific) was coupled to a Vanquish Neo UHPLC (Thermo Scientific).
304 Peptides were loaded onto a µPAC™ Neo Trapping Column (*COL-TRPNANO16G1B2*, Thermo Scientific),
305 and separated by reversed-phase chromatography on the latest generation µPAC™ column, a 50 cm
306 µPAC™ Neo Plus HPLC column(*COL-UPAC050NAN*, Thermo Scientific) featuring a structured pillar array
307 bed with a 180 µm bed width. To minimize post-column dispersion, the column was hosted in an
308 external column oven (Sonation, PRSO-V2-PF) directly mounted onto an EASY-Spray source (Thermo
309 Scientific) and connected via a µPAC™ Neo Plus voltage spacer (ACC-VSP1006, Thermo Scientific) to an
310 EASY-Spray™ Nano Emitter (ES993, Thermo Scientific).

311 Buffer A consisted of 0.1% FA in water, while Buffer B contained 0.1% FA in 80% acetonitrile.

312 *15-min gradient*

313 The chromatographic gradient was initiated at 4% buffer B at a flow rate of 750 nL/min, ramping to
314 12.5% over 3.1 minutes. The flow rate was then reduced to 450 nL/min while buffer B increased to
315 22.5% over 6 minutes, followed by 40% over 3.75 minutes. Finally, the column was regenerated using
316 90%B for 5.5 minutes.

317 *5-min gradient*

318 The chromatographic gradient was initiated at 1 to 8% buffer B in 0.1 minutes at a flow rate of 750
319 nL/min, ramping to 22.4% over 3.7 minutes. The flow rate was then reduced to 650 nL/min, and buffer
320 B was further increased to 40% over 1.25 minutes. Finally, the column was regenerated using 90%B for
321 4 minutes.

322 DIA

323 The mass spectrometer was operated in positive ionization mode with narrow-window data-
324 independent acquisition, with a full MS scan over a mass range of 400-900 in the Orbitrap, at a

resolution of 180,000 (5-min) and 240,000 (15-min), respectively. For the 5-min gradient, 166 windows of 3 Th were used to isolate and fragment all precursor ions from m/z 400 to 900, while for the 15-min gradient, 250 windows of 2 Th were used to isolate and fragment all precursor ions from m/z 400 to 900. A normalized collision energy of 25% was used for HCD fragmentation. MS2 scan range was set from 150 to 1700 m/z with detection in the Astral and a maximum injection time of 5 ms (5 min) or 3.5 ms (15 min). AGC targets were set to 500% (for MS1 scans) and 100% (for MS2 scans) for both methods, the RF lens was set to 45%, and a loop control of 0.4s (for 180,000 MS1 res) or 0.6s (for 240,000 MS1 res) ruled the cycle time.

3) 250 pg

An Orbitrap Astral (Thermo Scientific) was coupled to a Vanquish Neo UHPLC instrument (Thermo Scientific). Peptides were loaded onto a µPAC™ Trapping Column (Thermo Fisher Scientific) and were separated by reversed-phase chromatography using a 50 cm µPAC™ column (Thermo Scientific), featuring a structured pillar array bed with a 180 µm bed width. Buffer A consisted of 0.1% FA in water, while Buffer B contained 0.1% FA in 80% acetonitrile.

15-min gradient

Chromatographic gradient was initiated at 96.3% buffer A and 3.7% buffer B at a flow rate of 600 nL/min for 1 minute. The flow rate was then reduced to 200 nL/min, and the percentage of buffer B was further increased to 37.5% over 15 minutes.

5-min gradient

Chromatographic gradient was initiated with 96.3% buffer A and 3.7% buffer B at a flow rate of 600 nL/min for 1 minute. The flow rate was then reduced to 200 nL/min, and the percentage of buffer B was further increased to 37.5% over 5 minutes.

For MS measuring, the Orbitrap Astral mass spectrometer equipped with a FAIMS Pro Duo interface (Thermo Scientific) and an EASY-Spray source was coupled to the LC. A compensation voltage of –45 V was used.

DDA

The mass spectrometer was operated in positive ionization mode with data-dependent acquisition, with a full MS scans over a mass range of 350-1400 m/z with detection in the Orbitrap at a resolution of 240,000 and a fixed cycle time of 1 s. Precursor ion selection width was kept at 2 Th and a normalized collision energy of 30% was used for HCD fragmentation. MS2 scan range was set from 100 to 2500 m/z with detection in the Astral and a maximum fill time of 100 ms. Dynamic exclusion was enabled and set to 20 s.

DIA

The mass spectrometer was operated in positive ionization mode with data-independent acquisition, with a full MS scan over a mass range of m/z 400-800 with detection in the Orbitrap at a resolution of 240,000. In each cycle of data-independent acquisition, 20 windows of 20 Th were used to isolate and fragment all precursor ions from 400 to 800 m/z. A normalized collision energy of 25% was used for HCD fragmentation. MS2 scan range was set from 150 to 2000 m/z with detection in the Astral with a maximum injection time of 100 ms.

4) 250 pg – refined conditions

Ion Optics Direct injection

366 The Orbitrap Astral (Thermo Scientific) was coupled to a Vanquish Neo UHPLC instrument (Thermo
367 Scientific). Peptides were separated by reversed-phase chromatography using an Aurora® Elite XT (15
368 cm x 75 µm) (IonOpticks) column to accommodate both 5 and 15 mins separations. Buffer A consisted
369 of 0.1% FA in water, while Buffer B contained 0.1% FA in 80% acetonitrile.

370 Peptides were separated using a 5 and 15-min linear gradient from 3-30%B at a flow rate of 0.55 or 0.3
371 µL/min respectively.

372 DIA

373 The parameters from this section were applied for both gradients. The mass spectrometer was operated
374 in positive ionization mode using FAIMS at CV-50 and a total carrier gas flow at minimum, or 3.5L/min.
375 The full MS scan was covered over a mass range of 400-800 in the Orbitrap, at a resolution of 240,000.
376 Windows of 20 Th were used to isolate and fragment all precursor ions from m/z 400 to 800, and a
377 normalized collision energy of 25% was used for HCD fragmentation. MS2 scan range was set from 150
378 to 1700 m/z with detection in the Astral with a maximum injection time of 40ms. AGC targets were set
379 to 500% (for MS1 scans) and 800% (for MS2 scans), the RF lens was set to 45%, and a loop control or
380 0.6s ruled the cycle time.

381 *Bruker TimsTOF Ultra 2*

382 *1) 250 pg*

383 A Bruker timsTOF Ultra 2 (Bruker Daltonics, Bremen, Germany) was coupled to a nanoElute 2 LC
384 instrument (Bruker Daltonics). Peptides were separated by reversed-phase chromatography using a 5
385 cm Aurora Rapid CSI column (IonOpticks) featuring a 75 µm inner diameter, 1.7 µm particle size, and
386 120 Å pore size. Buffer A consisted of 0.1% FA in water, while Buffer B contained 0.1% FA in acetonitrile.

387 *15-min gradient*

388 Chromatographic gradient was initiated at 97% buffer A and 3% buffer B at a flow rate of 250 nL/min
389 for 1 minute. The percentage of buffer B was further increased to 30% over 15 minutes.

390 *5-min gradient*

391 Chromatographic gradient was initiated at 97% buffer A and 3% buffer B at a flow rate of 250 nL/min
392 for 1 minute. The percentage of buffer B was further increased to 30% over 5 minutes.

393 For MS measuring, the timsTOF Ultra mass spectrometer (Bruker Daltonics) equipped with a
394 CaptiveSpray nanoelectrospray source was coupled to the LC. An electrospray voltage of 1.3 kV was
395 applied for ionization.

396 Data-Dependent Acquisition Parallel Accumulation Serial Fragmentation (dda-PASEF)

397 The mass spectrometer was operated in positive ionization mode using dda-PASEF with 10 PASEF
398 MS/MS ramps. A full MS scan over a mass range of 100-1700 m/z and an ion mobility range from 0.64
399 to 1.45 1/K0 (V.s/cm2) were set, resulting in a cycle time of 1.18 s. Precursor ion selection width was
400 kept at 2-Th@700 m/z and 3-Th @800 m/z. Ion accumulation and ramp time were set to 100 ms, and
401 the collision energy was applied as a linear ramp from 20 eV at 1/K0 = 0.60 V.s/cm2 to 59 eV at 1/K0 =
402 1.6 V.s/cm2. MS2 scan range was set from 100 to 1700 m/z, the target intensity value was set to 20,000
403 with an intensity threshold of 1500.

404 Data- Independent Acquisition Parallel Accumulation Serial Fragmentation (dia-PASEF)

405 The mass spectrometer was operated in positive ionization mode using dia-PASEF, with a full MS scan
406 over a mass range of m/z 100-1700 and an ion mobility range from 0.64 to 1.45 1/K0 (V.s/cm2) resulting
407 in a cycle time of 0.96 s. In each DIA cycle, MS2 scan windows ranged from 400 to 1000 m/z using 25
408 m/z isolation windows (Supplementary Table 5). Ion accumulation and ramp time were set to 100 ms,
409 and the collision energy was applied as a linear ramp from 20 eV at 1/K0 = 0.60 V.s/cm2 to 59 eV at 1/K0
410 = 1.6 V.s/cm2.

411 Diagonal Parallel Accumulation Serial Fragmentation (diagonal-PASEF)

412 The mass spectrometer was operated in positive ionization mode using diagonal-PASEF, with a full MS
413 scan over a mass range of 50-3000 m/z and an ion mobility range from 0.64 to 1.45 1/K0 (V.s/cm2),
414 resulting in a cycle time of 0.749 s.

415 In each cycle MS2 scan range was set from a m/z start of 400 to 850 and a m/z end from 550 to 1000,
416 and an ion mobility range from 0.84 to 1.15 1/K0 (V.s/cm2). The number of MS2 scans was set to 6 using
417 25 Th isolation windows. Ion accumulation and ramp time were set to 100 ms, and the collision energy
418 was applied as a linear ramp from 20 eV at 1/K0 = 0.60 V.s/cm2 to 59 eV at 1/K0 = 1.6 V.s/cm2.

419 *2) 50 ng*

420 A Bruker timsTOF Ultra 2 (Bruker Daltonics, Bremen, Germany) was coupled to a nanoElute LC
421 instrument (Bruker Daltonics). Peptides were separated by reversed-phase chromatography using a 5
422 cm Aurora Rapid column (IonOpticks) featuring a 150 µm inner diameter, 1.7 µm particle size, and 120
423 Å pore size. Buffer A consisted of 0.1% FA in water, while Buffer B contained 0.1% FA in acetonitrile.

424 *15-min gradient*

425 Chromatographic gradient was initiated at 97% buffer A and 3% buffer B at a flow rate of 200 nL/min
426 for 1 minute. The percentage of buffer B was further increased to 30% over 15 minutes. The column
427 was washed with 80% B for 2 min and afterwards the column was re-equilibrated to starting conditions
428 for an additional 6 min.

429 *5-min gradient*

430 Chromatographic gradient was initiated at 97% buffer A and 3% buffer B at a flow rate of 200 nL/min
431 for 1 minute. The percentage of buffer B was further increased to 30% over 5 minutes. The column was
432 washed with 80% B for 2 min and afterwards the column was re-equilibrated to starting conditions for
433 an additional 6 min.

434 For MS measuring, the timsTOF Ultra 2 mass spectrometer (Bruker Daltonics) equipped with a
435 CaptiveSpray nanoelectrospray source was coupled to the LC. An electrospray voltage of 1.6 kV was
436 applied for ionization.

437 In all methods, inverse mobilities [1/K0] from 0.64 Vs/cm^2 to 1.45 Vs/cm^2 with an ion accumulation and
438 ramp time of 100 ms, respectively.

439 ddaPASEF

440 For ddaPASEF, one survey TIMS-MS scan was followed by ten PASEF ramps for MS/MS
441 acquisition,wherein precursors with charge state 0 to 5 were selected for fragmentation using an
442 intensity threshold of 500 and a target intensity of 20,000, resulting in a 1.18 s cycle time. MS/MS
443 isolation windows were set to m/z 2.0 for precursor ions below m/z 700 and m/z 3.0 for those above.
444 MS/MS spectra were scanned from *m/z* 100 to *m/z* 1700. Active Exclusion was enabled and set to 0.40
445 min.

446 diaPASEF

447 The mass spectrometer was operated in positive ionization mode using dia-PASEF, with a full MS scan
448 over a mass range of m/z 100-1700 and an ion mobility range from 0.64 to 1.37 1/K0 (V.s/cm2) resulting
449 in a cycle time of 0.96 s. In each DIA cycle, MS2 scan windows ranged from 400 to 1000 m/z using 25
450 m/z isolation windows (Supplementary Table 5). The collision energy was applied as a linear ramp from
451 20 eV at 1/K0 = 0.60 V.s/cm2 to 63 eV at 1/K0 = 1.6 V.s/cm2.

452 diagonal-PASEF

453 The mass spectrometer was operated in positive ionization mode using diagonal-PASEF, with a full MS
454 scan over a mass range of 100-1700 m/z and an ion mobility range from 0.64 to 1.45 1/K0 (V.s/cm2),
455 resulting in a cycle time of 0.64 s.

456 In each cycle MS2 scan range was set from a m/z start of 400 to 980 and a m/z end from 550 to 1130,
457 and an ion mobility range from 0.82 to 1.25 1/K0 (V.s/cm2). The number of MS2 scans was set to 6 using
458 30 Th isolation windows. The collision energy was applied as a linear ramp from 20 eV at 1/K0 = 0.60
459 V.s/cm2 to 59 eV at 1/K0 = 1.6 V.s/cm2.

460

461 ## Data Records

462 Data record 1. The mass spectrometry proteomics data have been deposited to the ProteomeXchange
463 Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository [39] with the
464 dataset identifiers, PXD070049 [40] and PXD071205 [41]. A Sample and Data Relationship File (SDRF) and an
465 Investigation Description File (IDF) have been uploaded to ProteomeXchange [42,43]. These files are used
466 to annotate the sample metadata and link the metadata to the corresponding data file(s) and thus will
467 improve the reproducibility and reanalysis of this comprehensive benchmark dataset. Log in to the
468 PRIDE website using the following project accession: PXD070049 (token: hcXc5dENPPrf). Alternatively,
469 reviewer(s) can access the dataset by logging in to the PRIDE website using the following username:
470 reviewer_pxd070049@ebi.ac.uk (Password: LF0Bpyt0gt1g). The refined Astral datasets can be accessed
471 through PRIDE using the following project accession: PXD071205 (token: 5Er73LxcO8VH). Alternatively,
472 reviewer(s) can access the dataset by logging in to the PRIDE website using the following username:
473 reviewer_pxd071205@ebi.ac.uk (Password: KXlk44y48XYF)

474 Data record 2. The AutoQC data analysed in Skyline is available from Panorama Public with the link
475 https://panoramaweb.org/LFQBenchmark_Generation_Beta.url [44].

476

477 ## Technical Validation

478 Similar to the first generation LFQ benchmark dataset, a continuous system suitability
479 monitoring protocol was implemented to assess LC-MS/MS performance over time. A
480 commercial *E. coli* protein digest extract was used as automatic quality control (AutoQC)
481 sample across most of the instruments. These samples were acquired in DDA mode using a 15-
482 minute LC gradient, with the exception of the 50 ng timsTOF Ultra 2, where 5- and 15-minute
483 gradients were alternated intermittently.

484 System suitability was evaluated using peptide-identification-free metrics—such as retention
485 time, peak area, and mass accuracy—extracted via the vendor-neutral Panorama AutoQC
486 framework [45]. To define suitable peptides, a subset of *E. coli* samples per instrument was peak

487 picked using MSConvert (v3.0.23229) and searched with Mascot Daemon (v2.8.0) against an *E.*
488 *coli* FASTA database using the following parameters: (i) 20 ppm precursor mass tolerance, (ii)
489 50 ppm fragment mass tolerance, and (iii) up to two missed cleavages.

490 Mascot identification results were exported as .dat files and imported into Skyline Daily (v24.1).
491 Four highly abundant proteins—P0A6F3, P0A6F5, P0A6Y8, and P0A853—were retained. Low-
492 intensity and co-eluting precursors were excluded. AutoQC Loader (v24.1.0.443) was
493 configured per instrument to automatically import raw files containing "AutoQC" in the
494 filename pattern. Processed results and reports are available in the PanoramaWeb folder: U of
495 Ghent Pharma Biotech Lab – LFQBenchmark_Generation_Beta.

496 Notably, far fewer outliers were observed compared to the previous generation benchmark
497 dataset when evaluating TIC and precursor area, retention time, and mass accuracy
498 (Supplementary Figure 1). The improvement in performance may be attributed to the use of
499 recent workflows, much shorter LC gradients, as well as the enhanced robustness of the newer
500 generation LC and MS instruments. As inherent variations in chromatographic setups would
501 directly influence identifications, raw files were also visually inspected to ensure data quality.
502 A set of chromatograms representative of the pool of separations is available in Supplementary
503 Figure 2 and 3.

504 For the timsTOF Ultra 2 – 250 pg dataset, a decrease in precursor area was seen over the time
505 course of the sample batch. Whether the observed effect stems from a true loss in instrument
506 sensitivity or from sample-related issues cannot be definitively confirmed. However, since the
507 overall TIC area remained stable, instrument degradation is likely not the primary cause.

508 The 250 pg samples on the Orbitrap Astral were measured using a stock concentration of 5
509 ng/µL, with injection volumes of 0.05 µL, which is well established and supported by the
510 Vanquish Neo LC system. Nonetheless, a general decrease in instrument sensitivity was
511 observed on the Orbitrap Astral system used for the 250 pg batch, as evidenced by lower
512 precursor intensities and reduced total ion chromatogram (TIC) areas from overlapped
513 chromatograms, compared to earlier runs. This could suggest a partial workflow performance
514 decline during the analysis, or more likely, an effect caused by the ultra-low load nature of the
515 acquisitions which may have contributed to the reduced identification depth in this specific
516 dataset.

517 For the ZenoTOF 8600 (5 µL/min) and Orbitrap Astral 50 ng (refined conditions) dataset, a
518 different quality control procedure was applied. More specifically, for the ZenoTOF 8600 the *E.*
519 *coli* protein digest was measured in the respective LC gradient and DIA methodology employed
520 during the acquisition of the different sample batches. Therefore, we were no longer able to
521 use Mascot as search engine and opted for DIA-NN (v2.1.0) for identification purposes. As
522 expected, the reproducibility in number of precursor and protein group identifications within
523 batch was extremely high which corroborates with the AutoQC DDA results obtained for the
524 ZenoTOF 7600+. For the refined 50 ng Orbitrap Astral dataset, we selected a more complex
525 proteome, namely a HeLa protein digest standard (Thermo Fischer Scientific, Waltham,
526 Massachusetts, United States), and analyzed it using DIA to expand the number of target

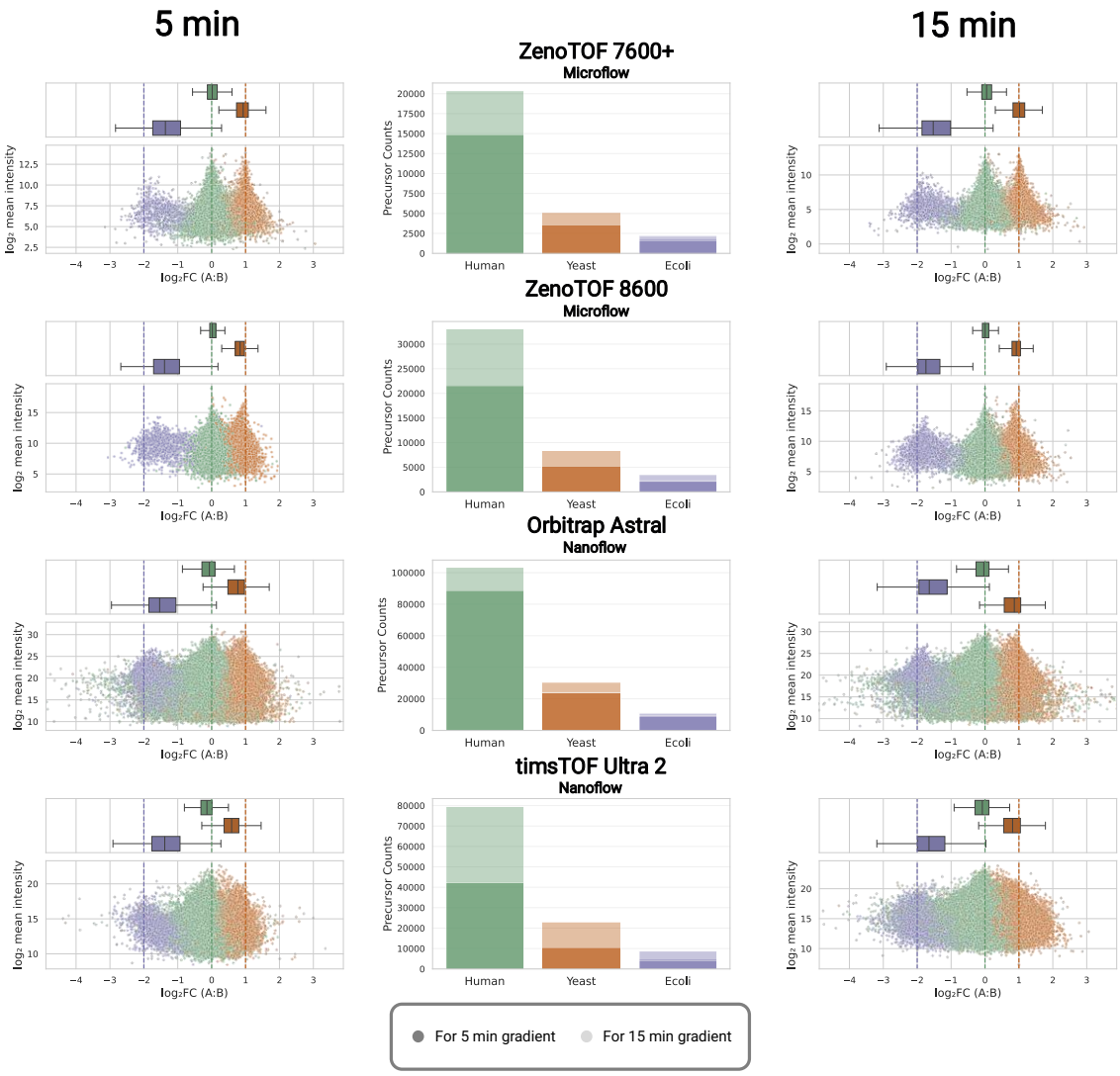
peptides available for QC monitoring and again achieved extremely high reproducibility, with 98.5% data completeness when performing directDIA+ analysis in Spectronaut.

In summary, the implemented system suitability monitoring confirmed that the dataset is of high quality and representative of realistic performance achieved in modern proteomics core facilities.

Usage Notes

In the following section, we illustrate the potential of this curated dataset through selected use cases, without positioning it as a benchmarking exercise of instrument performance. The core dataset was generated using unified operating conditions, rather than instrument-specific performance tuning. Although additional analyses were performed under refined configurations, these serve as complementary examples. Therefore, the dataset should not be used to draw conclusions about which instrument performs "best." Instead, it enables controlled, transparent, and methodologically fair evaluations of tools, workflows, and acquisition strategies.

MA plots are well-established diagnostic plots that evaluate quantification performance using hybrid-proteome benchmarks such as triple proteome mixtures. They offer a direct view of quantification accuracy, dynamic range and systematic biases in quantification by displaying the relationship between the mean precursor intensity and the $\log_2$ fold change (FC) between conditions (e.g. A/B). Across both 5- and 15-min gradients, MA plots generated for each instrument highlight expected abundance-dependent error patterns and illustrate how the dataset supports cross-platform evaluation of quantitative behavior (Figure 2).



548

549 **Figure 2. Assessment of dataset usability and quality.** Eight datasets, acquired with 5- and 15-minute LC gradients on four
550 different MS platforms using a standardized analytical setup, were evaluated. The left and right columns show MA plots for the
551 5-minute and 15-minute gradients on 50 ng input, respectively, depicting the $\log_2$FC of precursor ion intensities on the x-axis
552 as a function of their mean intensity on the y-axis, for the A and B conditions. For each MA plot, the adjacent boxplot
553 summarizes the same distribution of $\log_2$FCs and indicates the observed median $\log_2$FC; whiskers extend to 1.5 * the
554 interquartile range. Outliers are not shown in the boxplots but are visible in the corresponding MA plots. Colored vertical lines
555 denote the expected $\log_2$FC for each species. The middle column shows stacked barplots representing, for each MS platform
556 and species, the number of unique precursors quantified in the A and B conditions. The dark bars correspond to precursors
557 quantified in the 5-minute gradient datasets, and the lighter segments stacked on top indicate the additional precursors
558 quantified in the 15-minute gradient datasets. Precursor ion identification and quantification were performed with DIA-NN
559 (see Supplementary Methods).

560 A central use case for this dataset is its integration into ProteoBench, a community-driven
561 benchmarking platform that inspired the development of this second-generation LFQ reference
562 dataset. ProteoBench provides a transparent, reproducible environment for evaluating
563 proteomics data analysis software. It also aims to facilitate open discussions on evaluation
564 principles, for example on how benchmarking datasets, should be composed. ProteoBench
565 enables users and software developers to perform targeted, side-by-side comparisons [38]. The
566 diversity of instruments and acquisition strategies represented in this dataset makes it well-
567 suited as a reasonable resource for assessing identification overlap, quantification error, and

568 workflow robustness across proteomics data analysis tools. For example, precursor and protein
569 identifications and their quantification accuracy can be compared across DIA pipelines such as
570 AlphaDIA [46], DIA-NN [47], FragPipe [48], PEAKS [49] and Spectronaut (Supplementary Figure 4 and 5).

571 Since this dataset includes data acquired across multiple MS instruments, LC setups and
572 acquisition methodologies, it also provides a unique resource to evaluate how technological
573 innovations and parameter tweaking affect proteomics performance. Firstly, the use of a
574 scanning quadrupole for DIA has gained momentum across multiple vendors: SCIEX introduced
575 ZT Scan DIA[27] with the ZenoTOF 7600+, while Bruker implemented a similar concept into the
576 timsTOF platform as miDIA-PASEF[29] and synchro-PASEF[30] (collectively referred to as diagonal-
577 PASEF[28]). These approaches have been shown to improve precursor selectivity by incorporating
578 movement of the Q1 isolation window over time, effectively adding an additional separation
579 dimension to DIA acquisition. The advantage of these approaches have been especially
580 important in short-gradient, low-input experiments [28]. Here, we evaluated performance trends
581 previously reported in the literature and found them reflected in our dataset. In the low-input
582 250 pg Ultra 2 dataset, we observe a clear improvement in quantification accuracy when using
583 diagonal-PASEF compared to classical dia-PASEF (Supplementary Figure 6), reflected in the
584 lower median deviation from the expected $\log_2$FCs. This effect is consistent across both the 5-
585 and 15- minute gradients, although it is accompanied by a reduction in the number of
586 quantified precursors. It remains unclear whether this reduction stems from suboptimal DIA-
587 NN parameter settings or if it is consistent accross the latest generation of TIMS instrument, as
588 we were unable to evaluate a 250 pg dataset on the timsUltra AIP in this study. Similarly, for
589 the ZenoTOF 8600 datasets, ZT Scan DIA acquisition improves quantification relative to Zeno
590 SWATH DIA for the 5- and 11- minute gradients using microflow LC, whereas this benefit is not
591 observed for the longer 15-minute gradient (Supplementary Figure 7). Using nanoflow LC, on
592 the other hand, ZT Scan DIA acquisition provides better quantification for 15-minute gradients
593 as well.

594 Secondly, and also shown by this last observation, this dataset allows for the investigation of
595 how chromatographic setups impact identification performance. For many years, SCIEX users
596 have favored microflow LC (~5 µL/min) as a robust and reproducible alternative to nanoflow
597 separations, despite its known sensitivity trade-offs [50]. Given the rising importance of sensitivity
598 in low-input samples, especially single-cell proteomics, we systematically compared three flow
599 rates—5.0, 1.5, and 0.25 µL/min—on the ZenoTOF 8600 at 50 ng (Supplementary Figure 8).
600 Identification rates scaled inversely with flow rate, when moving from microflow to low-
601 nanoflow, with intermediate results at 1.5 µL/min. Gains were most pronounced at the
602 precursor level, consistent with improved ionization efficiency and reduced ion suppression at
603 lower flow rates [51]. At the protein group level, increases were more modest, reflecting
604 redundancy in peptide-to-protein mapping.

605 It should be emphasized that the main aim of this dataset is not to maximize performance on
606 each MS platform, but rather to generate a representative dataset that reflects average
607 instrument usage conditions. Nonetheless, recognizing that acquisition parameters can be
608 further optimized and that doing so is expected to enhance performance, we acquired a set of

609 two additional 50 ng and one 250 pg datasets on the Orbitrap Astral under more refined
610 analytical conditions.

611 Across the two refined 50 ng Orbitrap Astral datasets, we observed clear performance
612 improvements compared to the standardized acquisitions (Supplementary Figure 9). Notably,
613 the nature of these improvements differed depending on the refinement strategy. In one of
614 these, the most vital change concerned the standardized linear LC gradient from 3–30%, which
615 was optimized to achieve a more uniform distribution of peptide separation across the
616 chromatographic run. This workflow was also operated in a trap-and-elute setup that minimizes
617 post-column dispersion, a feature enabled by the latest generation µPAC Neo Plus column,
618 which is particularly advantageous for continuous measurements and for meeting laboratory
619 robustness requirements. For the 15-min gradient, this configuration substantially increased
620 the number of quantified precursors, albeit at the cost of higher quantification error. By
621 contrast, a refined direct-injection configuration yielded a more moderate gain in sensitivity
622 but did not introduce additional quantification error. For the 5-minute gradient, both refined
623 methods improved quantification accuracy, while the increase in quantified precursors was
624 more modest, likely limited by the narrower m/z scan range (400-800 vs. 380-980). A MS2
625 quantitative strategy was opted for this study, and further quantitative improvements might
626 have been observed on the short gradient while using an MS1 strategy. Importantly, the
627 improvements presented here are not exhaustive. Further optimization, such as fine-tuning LC
628 gradients and workflow, or operating FAIMS at lower resolution for low-input samples may
629 provide additional gains in sensitivity, sampling efficiency, and quantitative performance [52].
630 Collectively, these results highlight that substantial performance benefits can be unlocked
631 when LC-MS parameters are carefully tailored to the analytical objectives and sample amount.

632 It should be noted that all LC-MS platforms included in this study are capable of such improved
633 performance through similar fine-tuning. Again, the data presented here were intentionally
634 generated under standardized instructions and to the best efforts of the operators.
635 Consequently, no direct conclusions can be drawn regarding which platform is "best" in terms
636 of identification depth or quantitative performance.

637 By spanning different instruments, acquisition methods, and LC gradients, this dataset provides
638 a controlled yet representative resource for fair and transparent evaluations across tools and
639 platforms. This enables assessment of performance metrics, whether in the context of software
640 benchmarking through initiatives such as ProteoBench, the evaluation of novel acquisition
641 approaches like narrow-window DIA, diagonal-PASEF and ZT Scan DIA, or the study of
642 methodological variables such as workflows, flow rate and MS parameter optimization. In the
643 long term, such community-driven benchmarking efforts will not only drive continuous
644 improvements in both bioinformatics pipelines and experimental practices but also empower
645 end-users, including biologists, to select the strategies and tools that best match their research
646 needs and expectations.

647 ## Code Availability

648 The Jupyter notebooks used to generate Figure 2 and the Supplementary Figures 4, 5, 6, 7, 8 and 9 are
649 available through Zenodo, under 10.5281/zenodo.1793665.

Acknowledgements

This research was funded by grants from the Research Foundation Flanders (FWO) and the Special University Research Fund of Ghent University awarded to B.VP (grant numbers: 1278023N and BOF/PDO/2025/049). R.D., L.M. and R.B. acknowledge funding from the Research Foundation Flanders (FWO) [1SH9O24N, G010023N, 12A6L24N]. L.M. acknowledges funding from the Horizon Europe Projects BAXERNA 2.0 [101080544] and COMBINE [101191739], and from the Ghent University Concerted Research Action [BOF21/GOA/033]. L.M. and C.Car are further supported by the CHIST-ERA project ODEEP-EU [G0GDV23N]. We thank ProGenTomics for providing access to the ZenoTOF 7600+ system used in this study. YPR thanks European Molecular Biology Laboratory core funding, Wellcome grants (208391/Z/17/Z and 223745/Z/21/Z) and the BBSRC grant 'DIA-Exchange' (BB/X001911/1). This work was supported by the Agence Nationale de la Recherche via the French Proteomic Infrastructure (ProFI UAR2048; ANR-10-INBS08-03 & ANR-24-INBS-0015), by the Region Grand-Est (SC-Proteomics project) and by the ITMO Cancer of Aviesan within the framework of the 2021-2030 Cancer Control Strategy, on funds administered by Inserm (ProteomiSC project), for equipment funds at C.Car. C. Car and C.K. acknowledge the Interdisciplinary Thematic Institute IMS, the drug discovery and development institute, as part of the ITI 2021-2028 program of the University of Strasbourg, CNRS and Inserm, supported by IdEx Unistra (ANR-10-IDEX-0002), and by SFRI-STRAT'US project (ANR-20-SFRI-0012). We would like to thank BSI for their support in setting up the analysis in PEAKS.

Author contributions

B.VP, R.D., R.B., and M.D. conceived the study and wrote the draft manuscript with contributions from all authors. B.VP prepared the samples and performed the ZenoTOF 7600+ data acquisition with the help from I.B. and JB.V.; P.P. performed the ZenoTOF 8600 data acquisition; C.C., E.S., J.B.R., C.K., C.Car, F.F., L.L., P.M. and B.G. performed the Orbitrap Astral data acquisition; J.B.R., C.K., C.Car performed the timsTOF Ultra 2 – 250 pg data acquisition; S.P., C.P. and P.N. performed the timsTOF Ultra 2 – 50 ng data acquisition and R.D. and R.B. wrote the custom scripts for data analysis. YPR organised the ProteomeXchange submission. DD and LM co-supervised the experiment.

Competing interests

Frédéric Fontaine is employed by Thermo Fisher Scientific. Ihor Batruch, Patrick Pribil and Jean-Baptiste Vincendet are employed by SCIEX. Bart Van Puyvelde joined SCIEX after the completion of this work.

References

1. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).

2. Zubarev, R. A. & Makarov, A. Orbitrap mass spectrometry. *Anal. Chem.* **85**, 5288–5296 (2013).

3. Stewart, H. I. *et al.* Parallelized Acquisition of Orbitrap and Astral Analyzers Enables High-Throughput Quantitative Analysis. *Anal. Chem.* **95**, 15656–15664 (2023).

4. Verenchikov, A. N., Makarov, V. V., Vorobyev, A. V. & Kirillov, S. N. A Perspective of Multi-Reflecting TOF MS. *Mass Spectrom. Rev.* https://doi.org/10.1002/mas.21915 (2024) doi:10.1002/mas.21915.

5. Loboda, A. V. & Chernushevich, I. V. A Novel Ion Trap That Enables High Duty Cycle and Wide m/z Range on an Orthogonal Injection TOF Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **20**, 1342–1348 (2009).

690 6. Guzman, U. H. *et al.* Ultra-fast label-free quantification and comprehensive proteome coverage
691 with narrow-window data-independent acquisition. *Nature Biotechnology 2024 42:12* **42**, 1855–
692 1866 (2024).

693 7. Guo, T., Steen, J. A. & Mann, M. Mass-spectrometry-based proteomics: from single cells to
694 clinical applications. *Nature 2025 638:8052* **638**, 901–911 (2025).

695 8. Barkovits, K. *et al.* Reproducibility, Specificity and Accuracy of Relative Quantification Using
696 Spectral Library-based Data-independent Acquisition. *Mol. Cell. Proteomics* **19**, 181 (2019).

697 9. Fröhlich, K. *et al.* Data-Independent Acquisition: A Milestone and Prospect in Clinical Mass
698 Spectrometry–Based Proteomics. *Molecular & Cellular Proteomics* **23**, 100800 (2024).

699 10. Ludwig, C. *et al.* Data-independent acquisition-based SWATH-MS for quantitative proteomics: a
700 tutorial. *Mol. Syst. Biol.* **14**, (2018).

701 11. Bouwmeester, R., Gabriels, R., Van Den Bossche, T., Martens, L. & Degroeve, S. The Age of Data-
702 Driven Proteomics: How Machine Learning Enables Novel Workflows. *Proteomics* **20**, 1900351
703 (2020).

704 12. Bouwmeester, R., Gabriels, R., Hulstaert, N., Martens, L. & Degroeve, S. DeepLC can predict
705 retention times for peptides that carry as-yet unseen modifications. *Nat. Methods* **18**, 1363–
706 1369 (2021).

707 13. Gessulat, S. *et al.* Prosit: proteome-wide prediction of peptide tandem mass spectra by deep
708 learning. *Nat. Methods* **16**, 509–518 (2019).

709 14. Degroeve, S. & Martens, L. MS2PIP: a tool for MS/MS peak intensity prediction. *Bioinformatics*
710 **29**, 3199–3203 (2013).

711 15. Gabriels, R., Martens, L. & Degroeve, S. Updated MS^2PIP web server delivers fast and accurate
712 MS^2 peak intensity prediction for multiple fragmentation methods, instruments and labeling
713 techniques. *Nucleic Acids Res.* **47**, 295–299 (2019).

714 16. Devreese, R. *et al.* Collisional Cross-Section Prediction for Multiconformational Peptide Ions with
715 IM2Deep. *Anal. Chem.* **97**, 15113–15121 (2025).

716 17. Meier, F. *et al.* Deep learning the collisional cross sections of the peptide universe from a million
717 experimental values. *Nat. Commun.* **12**, 1185 (2021).

718 18. Zeng, W.-F. *et al.* AlphaPeptDeep: a modular deep learning framework to predict peptide
719 properties for proteomics. *Nat. Commun.* **13**, 7238 (2022).

720 19. Van Puyvelde, B. *et al.* Removing the Hidden Data Dependency of DIA with Predicted Spectral
721 Libraries. *Proteomics* **20**, 1900306 (2020).

722 20. Searle, B. C. *et al.* Generating high quality libraries for DIA MS with empirically corrected peptide
723 predictions. *Nat. Commun.* **11**, 1548 (2020).

724 21. Meier, F., Park, M. A. & Mann, M. Trapped Ion Mobility Spectrometry and Parallel Accumulation–
725 Serial Fragmentation in Proteomics. *Molecular & Cellular Proteomics* **20**, 100–138 (2021).

726 22. Shliaha, P. V., Bond, N. J., Gatto, L. & Lilley, K. S. Effects of traveling wave ion mobility separation
727 on data independent acquisition in proteomics studies. *J. Proteome Res.* **12**, 2323–2339 (2013).

23. Helm, D. *et al.* Ion mobility tandem mass spectrometry enhances performance of bottom-up proteomics. *Mol. Cell. Proteomics* **13**, 3709–3715 (2014).

24. Swearingen, K. E. & Moritz, R. L. High Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) for Mass Spectrometry-Based Proteomics. *Expert Rev. Proteomics* **9**, 505 (2012).

25. Moseley, M. A. *et al.* Scanning Quadrupole Data-Independent Acquisition, Part A: Qualitative and Quantitative Characterization. *J. Proteome Res.* **17**, 770–779 (2018).

26. Messner, C. B. *et al.* Ultra-fast proteomics with Scanning SWATH. *Nat. Biotechnol.* **39**, 846–854 (2021).

27. Sinn, L. R. *et al.* Performance Characteristics of Zeno Trap Scanning DIA for Sensitive and Quantitative Proteomics at High Throughput. *Proteomics* **26**, 68–81 (2026).

28. Below, C. R. *et al.* Enhanced Identifications and Quantification Through Retention Time Down-Sampling in Fast-Cycling Diagonal-PASEF Methods. *Molecular & Cellular Proteomics* **25**, 101480 (2026).

29. Distler, U. *et al.* midiaPASEF maximizes information content in data-independent acquisition proteomics. *bioRxiv* 2023.01.30.526204 Preprint at https://doi.org/10.1101/2023.01.30.526204 (2023).

30. Skowronek, P. *et al.* Synchro-PASEF Allows Precursor-Specific Fragment Ion Extraction and Interference Removal in Data-Independent Acquisition. *Mol. Cell. Proteomics* **22**, (2023).

31. Sinn, L. R. & Demichev, V. Entering the era of deep single-cell proteomics. *Nat. Methods* **22**, 459–460 (2025).

32. Momenzadeh, A. & Meyer, J. G. Single-cell proteomics using mass spectrometry. *Cell Genomics* **5**, 100973 (2025).

33. Bubis, J. A. *et al.* Challenging the Astral mass analyzer to quantify up to 5,300 proteins per single cell at unseen accuracy to uncover cellular heterogeneity. *Nat. Methods* **22**, 510 (2025).

34. Slavov, N. Scaling Up Single-Cell Proteomics. *Mol. Cell. Proteomics* **21**, 100179 (2021).

35. Ghosh, G., Shannon, A. E. & Searle, B. C. Data acquisition approaches for single cell proteomics. *Proteomics* **25**, (2025).

36. Van Puyvelde, B. *et al.* A comprehensive LFQ benchmark dataset on modern day acquisition strategies in proteomics. *Scientific Data 2022 9:1* **9**, 1–12 (2022).

37. Navarro, P. *et al.* A multicenter study benchmarks software tools for label-free proteome quantification. *Nat. Biotechnol.* **34**, 1130–1136 (2016).

38. Devreese, R. *et al.* ProteoBench: the community-curated platform for comparing proteomics data analysis workflows. Preprint at https://doi.org/10.64898/2025.12.09.692895 (2025).

39. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

40. Van Puyvelde, B. *et al.* ProteomeXchange Dataset PXD070049. *LFQ Benchmark Dataset - Generation Beta: Assessing Modern Proteomics Instruments and Acquisition Workflows with High-Throughput LC Gradients* https://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD070049 (2025).

41. Van Puyvelde, B. *et al.* ProteomeXchange Dataset PXD071205. *LFQ Benchmark Dataset - Generation Beta: Assessing Modern Proteomics Instruments and Acquisition Workflows with High-Throughput LC Gradients - Astral Optimized* https://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD071205 (2025).

42. Dai, C. *et al.* A proteomics sample metadata representation for multiomics integration and big data analysis. *Nat. Commun.* **12**, 5854 (2021).

43. Bittremieux, W. *et al.* The European Bioinformatics Community for Mass Spectrometry (EuBIC-MS): an open community for bioinformatics training and research. *Rapid Communications in Mass Spectrometry* e9087 (2021).

44. Van Puyvelde, B. *et al.* Panorama Public: U of Ghent - LFQ Benchmark Dataset - Generation Beta: Assessing Modern Proteomics Instruments and Acquisition Workflows with High-Throughput LC Gradients. https://panoramaweb.org/LFQBenchmark_Generation_Beta.url (2025).

45. Bereman, M. S. *et al.* An Automated Pipeline to Monitor System Performance in Liquid Chromatography–Tandem Mass Spectrometry Proteomic Experiments. *J. Proteome Res.* **15**, 4763–4769 (2016).

46. Wallmann, G. *et al.* AlphaDIA enables DIA transfer learning for feature-free proteomics. *Nat. Biotechnol.* 1–10 (2025) doi:10.1038/s41587-025-02791-w.

47. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2019).

48. Yu, F. *et al.* Analysis of DIA proteomics data using MSFragger-DIA and FragPipe computational platform. *Nat. Commun.* **14**, 1–14 (2023).

49. Krieger, J., Xin, L. & Shan, B. (Paul). Multi-User, High-Throughput PEAKS Online Workflow, for large scale DIA and DDA Proteomic Analysis. *J. Biomol. Tech.* **31**, S22 (2020).

50. Vowinckel, J. *et al.* Cost-effective generation of precise label-free quantitative proteomes in high-throughput by microLC and data-independent acquisition. *Sci. Rep.* **8**, 1–10 (2018).

51. Kelly, R. T. Single-cell Proteomics: Progress and Prospects. *Molecular & Cellular Proteomics* **19**, 1739–1748 (2020).

52. Hoch, D. G., Matzinger, M. & Mechtler, K. Low-resolution FAIMS for increased peptide coverage in low-load and single-cell proteomics. *bioRxiv* 2025.08.22.671812 Preprint at https://doi.org/10.1101/2025.08.22.671812 (2025).