ISSN 0006-2979, Biochemistry (Moscow), 2024, Vol. 89, No. 8, pp. 1349-1361 © The Author(s) 2024. This article is an open access publication.

MINI-REVIEW

Ultrafast Proteomics

Ivan I. Fedorov[1,2], Sergey A. Protasov[1,2], Irina A. Tarasova[2], and Mikhail V. Gorshkov[2,a]*

[1]*Moscow Institute of Physics and Technology (National University), 141700 Dolgoprudny, Moscow Region, Russia*
[2]*V. L. Talrose Institute for Energy Problems of Chemical Physics, N. N. Semenov Federal Research Center for Chemical Physics, Russian Academy of Sciences, 119334 Moscow, Russia*
[a]*e-mail: mike.gorshkov@gmail.com*

Received May 23, 2024
Revised June 21, 2024
Accepted June 24, 2024

Abstract—Current stage of proteomic research in the field of biology, medicine, development of new drugs, population screening, or personalized approaches to therapy dictates the need to analyze large sets of samples within the reasonable experimental time. Until recently, mass spectrometry measurements in proteomics were characterized as unique in identifying and quantifying cellular protein composition, but low throughput, requiring many hours to analyze a single sample. This was in conflict with the dynamics of changes in biological systems at the whole cellular proteome level upon the influence of external and internal factors. Thus, low speed of the whole proteome analysis has become the main factor limiting developments in functional proteomics, where it is necessary to annotate intracellular processes not only in a wide range of conditions, but also over a long period of time. Enormous level of heterogeneity of tissue cells or tumors, even of the same type, dictates the need to analyze biological systems at the level of individual cells. These studies involve obtaining molecular characteristics for tens, if not hundreds of thousands of individual cells, including their whole proteome profiles. Development of mass spectrometry technologies providing high resolution and mass measurement accuracy, predictive chromatography, new methods for peptide separation by ion mobility and processing of proteomic data based on artificial intelligence algorithms have opened a way for significant, if not radical, increase in the throughput of whole proteome analysis and led to implementation of the novel concept of ultrafast proteomics. Work done just in the last few years has demonstrated the proteome-wide analysis throughput of several hundred samples per day at a depth of several thousand proteins, levels unimaginable three or four years ago. The review examines background of these developments, as well as modern methods and approaches that implement ultrafast analysis of the entire proteome.

DOI: 10.1134/S0006297924080017

Keywords: proteomics, mass spectrometry, peptides, proteins, ultrafast analysis, quantitative proteomics

INTRODUCTION

Currently, whole proteome analysis is widely used in many areas of biological and medical research [1, 2]. The primary method for such analysis is mass spectrometry, which provides quantitative information on the changes in cell proteomes under different conditions. One of the fundamental initial stages in the development of quantitative proteomics are implementations of the concepts of database and/or spectral library search [3] and identification of proteins using unique set of masses of their proteolytic

Abbreviations: AMT tags, accurate mass and time tags; DDA, data-dependent acquisition method for whole proteome analysis; DIA, data-independent acquisition method for whole proteome analysis; DISPA, direct infusion shotgun proteome analysis without the use of online liquid chromatography separation; DirectMS1, method of direct protein identification using liquid chromatography and mass spectrometry; FDR, false discovery rate; FT-ICR, Fourier-transform ion cyclotron resonance mass spectrometry;HPLC, high performance liquid chromatography;MS1, precursor ion mass spectra; MS/MS, tandem mass spectrometry;PMF, peptide mass fingerprint.

* To whom correspondence should be addressed.

(typically tryptic) peptides, the latter known as Peptide Mass Fingerprint (PMF) [4]. Regarding the types of mass analyzers employed, the first proteomes of model organisms were identified using radio-frequency quadrupole ion traps and time-of-flight mass spectrometers [5]. First results were obtained without controlling the false-discovery rate (FDR) [6, 7], a concept introduced in proteomics in 2007 with the target-decoy approach [7]. Early on, starting with the works by Smith et al., necessity of high-resolution mass spectrometry for comprehensive proteome analysis was recognized, initially represented exclusively by the Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR) in combination with nanoflow peptide separation and ionization [8, 9]. Emergence of more efficient high resolution mass analyzers compared to FT-ICR, such as the Orbitrap, enabled identification of up to 80% of the yeast proteome within the one-hour single HPLC-MS/MS run by 2015 [10]. Currently, 5000 to 6000 proteins can be identified for human cell proteomes in a two-hour single-shot experiment [11]. Further increase in the depth of proteome-wide analysis is achieved through additional sample fractionation at the protein or peptide levels [12-14], by extending the HPLC gradient duration to several hours, as well as by employing long chromatographic columns [12, 15]. For instance, combining proteolytic mixture fractionation, long-hour LC gradients, and more than 40-cm columns enabled identification of more than half of the human proteome [16]. While achieving such proteome coverage is of great interest, total instrumental time in the cited work amounted to 288 hours, making such analysis unique but impractical for many applications involving routine characterization of hundreds of samples per day, such as in chemical and population proteomics, or clinical studies. Sample multiplexing by labeling techniques [17, 18], currently implemented using tandem mass tags (TMT) [19], partially addresses the issue of instrumental costs of quantitative proteome-wide analysis of a single sample. However, problems associated with the increased analytical complexity of the samples and the need for fractionation do not position the TMT approach as a method of ultrafast proteomics, which can be defined as the analysis of more than 200 samples per day. Indeed, recent studies on glioblastoma cell lines treated with interferon demonstrated that the 40-minute quantitative analysis using 10-plex TMT (equivalent to about 200 proteome analyses per day) provides a rather poor picture of interferon-regulated proteins [20].

After a significantly long time since the initial demonstrations of quantitative proteome-wide analysis based on Accurate Mass and Time tags (AMT) within a minute range of gradient separations, interest has recently renewed in this area, which could be tentatively called “ultrafast proteomics”. A number of methods were developed for its implementation based on the novel high-resolution mass spectrometry instruments combined with the ultra-short separations of peptide mixtures (including peptide ion separations in the gas phase), such as Data Independent Acquisition (DIA) [13] and DirectMS1 [20]. These methods allow semi-quantitative proteome analysis with a throughput of over 200 samples per day.

This review discusses new approaches in ultrafast proteomics developed in recent years and briefly explores their future prospects.

EARLY METHODS OF ULTRAFAST PROTEOMICS

One of the first implementations of the idea of ultrafast proteome analysis was the PMF approach [4, 21, 22]. This approach involves preliminary protein separation using gel electrophoresis or liquid chromatography, digestion of the protein fractions into proteolytic peptides (typically, using trypsin), forming a set of peptides with masses specific for each protein, and measuring mass spectra of these peptide ions in the fractions. For protein identification, the experimentally obtained peptide ion masses are compared with the theoretical ones derived from the available protein sequence databases of the organism’s proteome under study, as shown schematically in Fig. 1 [23].

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) is the most commonly used method for PMF implementation [4, 24]. Generally speaking, this approach is not ultrafast, as the proteome is divided into many fractions, usually using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) [21], each subjected to digestion and analysis. It quickly became apparent that the PMF method is ineffective for analyzing complex mixtures [25], involving digests of dozens or even thousands of proteins of a proteome. Currently, it is used almost exclusively for the analysis and confirmation of individual, usually pre-purified proteins.

With advancements in high resolution mass analyzers, such as ion cyclotron resonance mass spectrometry, the idea of identifying proteins based on measuring peptide ion mass spectra as a way of rapid proteome-wide analysis was realized in the Accurate Mass Tags (AMT) approach [26]. The method involves generation of a list of peptides potentially present in the analyzed samples based on preliminary MS/MS analysis of a pool of samples under study. This is followed by obtaining peptide ion mass spectra in the individual samples of the pool and protein identification based on matching the experimental masses with this project-specific list of unique peptide masses linked to the particular proteins. The basic idea of the approach is that if the molecular mass of a peptide

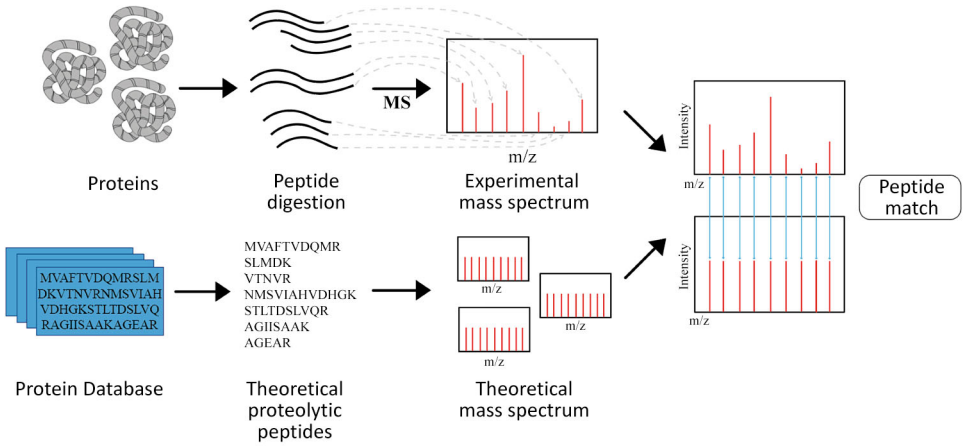

Fig. 1. Diagram of the PMF method. Since peptide masses are not specific to amino acid sequence, a single measured mass or even a set of masses could correspond to multiple, and in some cases, dozens of possible proteins, complicating their identification in complex mixtures.

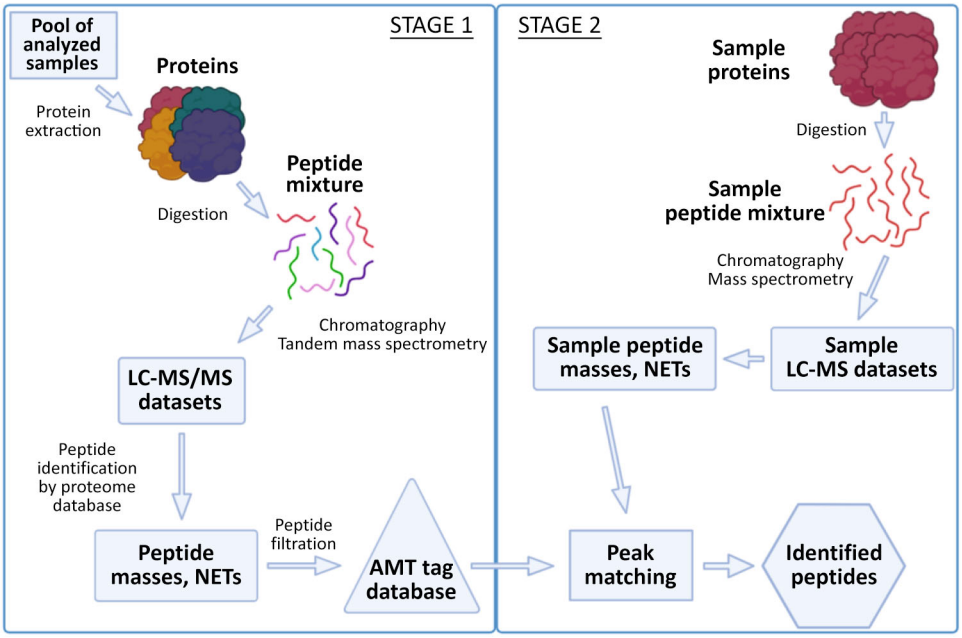

Fig. 2. Diagram of the AMT method based on accurate peptide ion masses and normalized elution times (NET). Normalization is performed into the range [0,1]. Predicted NETs are calculated for the employed separation conditions using either simple linear conversion, or neural networks [27].

could be measured with high enough accuracy so that its mass was unique among all possible peptides predicted from the genome, it could then be used as an "accurate mass tag" for protein identification. Accordingly, generating a list of such AMTs allows analyzing products of the whole proteome digest (e.g., obtained by trypsin cleavage) at greater speed and sensitivity. Moreover, the subsequent analysis of the individual samples can be conducted without peptide fragmentation stage, making the approach potentially MS/MS-free and, thus, compatible with separations of proteolytic mixtures using short gradients.

It quickly became evident that using additional complementary data to the accurately measured masses, such as peptide elution times, makes such combinations unique for peptide amino acid sequences. Therefore, the subsequent development of this approach involved addition of normalized peptide elution times (NETs), further transforming it into the Accurate Mass and Time tags (AMT-tags) method [26]. Early applications demonstrated feasibility of using the AMT tags method for proteome-wide analysis of relatively small proteomes, particularly, *Deinococcus radiodurans* [27]. Moreover, since the method does not require peptide

fragmentation for identification (except at the stage of the AMT tag list generation), such proteome-wide proteome analysis was performed for the first time within a minute-range time frame. Standard implementation of the AMT method consists of two main stages (Fig. 2): (i) generation of peptide AMT tags for pooled control and test sample groups using deep (typically, employing fractionation) proteome-wide LC-MS/MS analysis; and (ii) rapid HPLC-MS1-based analysis with protein identification based on the AMT tags database created in the first stage. In the first stage, each identified peptide is assigned its mass within measurement error and NET. The next stage involves analysis of a large cohort of unfractionated samples under study using HPLC-MS1, resulting in a list of experimental peptide masses and charge states, as well as elution times. The latter are converted to a normalized time scale, most simply by linear function. Peptide identification for each analyzed sample is based on matching the experimental data with the AMT tags, followed by protein identification and quantitation. The latter is performed for the identified proteins using the peptide ion intensities in the mass spectra.

While the AMT method demonstrated the possibility of ultrafast quantitative analysis of proteomes of various organisms [28, 29], its broader acceptance in proteomics is hindered by the lack of FDR control. Also, there are issues with aligning peptide elution times for calculating NETs across different experiments and separation conditions, especially between the ones used for generating the AMT tag database and those for subsequent rapid proteome analysis [30]. One of the possible solutions of the latter problem was the use of various peptide retention time prediction models [31, 32] and generation of the standardized and/or universal peptide elution time databases for the AMT tags based on them [33].

As the proteome-wide analysis using the so-called Data-Dependent Acquisition (DDA) approach with hybrid mass spectrometers featuring high-resolution Orbitrap ion traps [34-36] became routine laboratory practice, the AMT method ceased to be widely used. However, the DDA approach itself, where the peptide ions detected in MS1 spectra are, next, sequentially isolated in the radio-frequency ion trap of a hybrid mass spectrometer and accumulated to quantities sufficient for obtaining high quality fragmentation spectra, inherently involves using long HPLC gradients. Even in the case of multi-hour separations of the proteome digests, extending up to 10 h in some extreme cases [37], only a small fraction of peptides detectable in the MS1 spectra are identified [38-40]. Nevertheless, DDA has become the method of choice for quantitative proteome-wide analysis in recent years with an achievable depth of 10,000 or more protein identifications in some studies [16, 37, 41, 42]. Despite the obvious importance of achieving as large as possible depth of proteome analysis, there is also an evident issue: the enormously high instrumental time for analyzing a single sample, especially when extensive pre-fractionation of the analyzed mixtures is employed [41, 43-47].

DATA INDEPENDENT ACQUISITION FOR ULTRAFAST PROTEOMICS

One of the obvious methods for implementing ultrafast proteomics is DIA [13]. Unlike DDA, this method does not rely on sequential selection of the precursor ions based on their accurately measured mass in MS1 spectra for subsequent isolation, accumulation, and fragmentation, which is a primary reason for using long separation gradients. Instead, in DIA, ion accumulation and fragmentation occur within a broad mass window, with the sequential change in the operating parameters of the accumulation device to an adjacent window, and so on (Fig. 3). As a result, nearly all precursor ions present in the MS1 mass spectra are fragmented in a series of such windows (typically of 20-25 Th size), covering the entire *m/z* range of peptide ions. It is clear that the fragmentation spectra in such windows are mixed (or, are said to be highly multiplexed) and contain fragments from dozens of peptide ions simultaneously, which present another challenge of their further interpretation (deconvolution). Each series of such windows corresponds to the preliminary measured MS1 spectrum and elution time, the latter being a key parameter for subsequent deconvolution of the fragmentation spectra and peptide identification. Size of the windows and, thus, efficiency of the fragmentation spectrum deconvolution are determined by characteristics of the mass analyzer. For instance, combination of the Orbitrap mass analyzer and the Astral (ASymmetric TRAck Lossless) analyzer allowed reducing the fragmentation windows to 2 Th, effectively erasing the boundary between the DIA and DDA methods in proteome-wide analysis [48, 49].

The above-described DIA scheme corresponds to its most widely used implementation, called SWATH-MS (Sequential Window Acquisition of All THeoretical Mass Spectra) [50]. The main advantage of this method is overcoming the data stochasticity problem of the DDA approach, related to choosing a limited number of the most intense precursor ions in a given mass spectrum for selective accumulation and fragmentation. The result is a significantly lower level of missing values, making DIA an alternative to DDA in quantitative proteomics [51]. Simultaneously, since in DIA all precursor ions are fragmented in a limited number of windows, this method allows working with shorter HPLC gradients [52]. Further optimization of the isolation windows increased the depth of proteome analysis

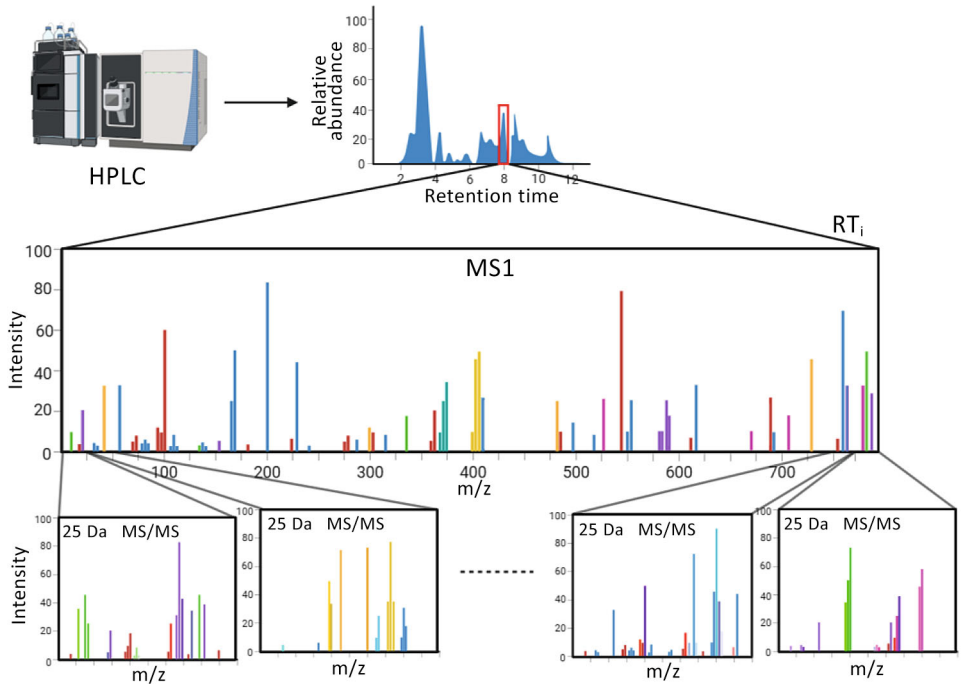

Fig. 3. Scheme of DIA method. Instead of isolation of individual precursor ions detected in MS1 mass spectra for subsequent fragmentation, the whole mass range is divided into a number of windows, in which all present ions are accumulated and fragmented. Thus, DIA allows obtaining tandem mass spectra for all precursor ions potentially present in a sample.

using ultra-short gradients [53, 54]. Additional separation of peptide ions by ion mobility demonstrated the possibility of identifying over 1000 proteins from 5 ng of HeLa digest in the DIA mode with 5-minute LC gradients [55].

One of the limiting factors in developing DIA as a routinely used method for ultrafast proteome-wide analysis was high level of the fragmentation mass spectra multiplexing, requiring complex data processing algorithms for deconvolution. Standard solution to this problem was using spectral libraries for the pool of analyzed samples. These libraries were generated using deep proteome-wide analysis by standard DDA, making DIA not entirely data-independent. Besides the obvious instrumental time costs for obtaining such libraries that makes DIA a conditionally fast proteome analysis method, the use of experimental libraries significantly limited application of DIA in inter-laboratory and clinical studies. Moreover, a fundamental issue remains with this approach: inability to identify peptides with fragmentation spectra not present in the library. Progress in the development of machine learning algorithms for predicting fragmentation spectra and peptide retention times *in silico* has solved the latter problem [56-58]. However, extremely high level of multiplexing, resulting from the interference of fragmentation spectra originating from different simultaneously eluting precursor ions, is significantly exacerbated when short separation gradients are used. Until recently, this made it impossible to extract any meaningful number of identifications from such spectra and limited the use of DIA in applications requiring large number of analyses. These issues were addressed in the recent development of the DIA-NN algorithm, based on the use of neural networks to distinguish signals of fragment ion from noise in the mass spectra and employing new strategies for extracting quantitative information and chromatogram aligning based on the identified peptides [59]. In the DIA-NN algorithm, elution peak of each precursor ion is described by a set of scores, and the best candidate for the elution peak of a particular precursor ion is determined through an iteration procedure based on the linear classifier. A key step in the algorithm's operation is using deep neural networks for assigning statistical significance (q-value) to the identified precursors, calculated for target and false candidates based on characteristics of the corresponding elution peaks. Capabilities of the DIA-NN algorithm for ultrafast proteome-wide analysis were fully demonstrated in implementation of the Scanning SWATH method [60, 61]. In this method the sequential selection of peptide isolation windows, in which fragmentation occurs, is replaced with continuous scanning of the first ion isolating RF quadrupole of the mass spectrometer by a broad *m/z* window across the entire mass range, simultaneously fragmenting incoming precursor ions in the collisional RF quadrupole. This creates an additional dimension for matching fragmentation spectra to precursor candidates during the subsequent deconvolution of highly

interfering MS/MS spectra and peptide identification by the DIA-NN algorithm. In the recent collaborative study by the developers of the Scanning SWATH method and DIA-NN algorithm, the previously unattainable performance with the depth reaching several thousand proteins was demonstrated in the analysis of human cell proteomes using ultrafast HPLC gradients of 0.5 to 5 min [61]. It is important to note that one of the conditions for the method's operation in ultrafast gradient mode required to maintain the resolving power of chromatographic separation of complex mixtures, is using high HPLC flow rates, around several hundred μl/min, which, in turn, leads to significant sample consumption (up to several μg of the human cell line digest).

Implementation of the DIA method in ultra-short separation mode looks straightforward for the time-of-flight mass analyzers, currently capable of acquiring mass spectra over a broad *m/z* range with a scanning rate of about 100 Hz and peak resolution in the spectra of 50,000 or higher. One example of applying high resolution time-of-flight mass analyzer for ultrafast proteome-wide analysis is combination of additional precursor separation by ion mobility using ion holding in an electric field gradient against the moving gas column TIMS (Trapped Ion Mobility Spectrometry) [62, 63] with parallel accumulation and sequential fragmentation of peptide ions PASEF (Parallel Accumulation SErial Fragmentation [64]). In TIMS, peptide ions eluting from the HPLC column and ionized in the ionization source enter a drift chamber where they are held in the radial direction by a constant electric field, compensating their drift in collisions with carrier gas molecules, and are, thus, separated by ion mobility. Instead of selecting one precursor ion for fragmentation, changes in the RF quadrupole parameters are synchronized with ion mobility chamber operation to isolate and fragment ions within the designated *m/z* range. One 50-ms step of changing the trapping electric field in the TIMS chamber allows obtaining fragmentation spectra of several peptide ions. PASEF significantly increases the rate of fragmentation spectrum acquisition without losing analysis sensitivity [65]. Implementing the TIMS-TOP/PASEF combination in DIA mode (dia-PASEF) and using the DIA-NN algorithm for data processing demonstrated the possibility of proteome-wide analysis with several thousand proteins identified for human cell line at a rate of up to 400 samples per day (3-minute HPLC gradient) [66].

DIRECT INFUSION METHOD DISPA

A logical step in the development of ultrafast proteomic methods and simplification of the instrumental component of analysis is elimination of the on-line chromatographic separation of proteolytic mixtures. This approach is not unique and was used in proteome analysis more than fifteen years ago [67]. However, its early implementations were based on mass analyzers of low resolution and mass measurement accuracy, there were no advanced search engines for identification existing at the time, and no any capabilities for additional ion separation, e.g., by ion mobility. Several years ago, the concept of direct injection of proteolytic mixture into the ionization source without on-line HPLC separation was renewed in the Direct Infusion Shotgun Proteomic Analysis approach (DISPA) owing to advancements in the mass spectrometry technologies, emergence of high resolution mass analyzers and fast ion mobility separation methods [68]. Technically, implementation of the DISPA method is quite simple: the proteolytic mixture is injected in a nanoflow mode directly from a syringe filled with a sample into the mass spectrometer's ionization source. Ion mobility separation is used as an additional dimension. Analysis is performed using DIA. It is clear that the multiplexing level of fragmentation spectra in this case is more than an order of magnitude higher than in the case of HPLC-based analysis, which, accordingly, limits the achievable depth of proteome coverage. In the cited work, a depth of about 500 proteins was demonstrated for the human cell line proteome. However, this depth was achieved within a few minutes of experimental time, allowing analysis of 132 samples in 4.5 h (3 min per sample) with quantitative identification of over 300 proteins. Inability to "link" fragment peaks to chromatographic times for effective deconvolution of highly multiplexed tandem spectra significantly limited capabilities of the DISPA method when using the standard HPLC-based DIA data processing algorithms. To overcome these limitations, software based on the CsoDIAq algorithm (Cosine similarity optimization for DIA qualitative and quantitative analysis [69]) was developed. Using this algorithm for the DISPA data processing demonstrated a depth of human cell line proteome analysis (HeLa and 293T cell lines) of about 2000 proteins in a single experimental run with total analysis time of a few minutes [70]. This work also demonstrated capabilities of the method for quantitative analysis of large cohorts of samples. In particular, 96 human cell line samples treated with a drug were analyzed within 8 h with a depth of about 1000 quantitatively identified proteins. It is worth noting that the DISPA method is an interesting alternative to the standard HPLC-based ultrafast proteome-wide analysis approaches, yet, it is at its early development stage. Limitations of the method stem from the extreme complexity of proteolytic mixtures, such as the whole proteome digests, which can contain millions of individual peptide sequences in a dynamic concentration range reaching several orders of magnitude

(over 10 orders in the case of human blood plasma proteome, for example). These limitations affect both the number of identifications and accuracy of quantitative analysis. High dynamic range leads to the strong suppression of the low-concentration peptides ions in both ionization source and accumulation ion trap of a hybrid mass spectrometer. Limitations of the method also include low sequence coverage of the identified proteins due to the low efficiency of deconvolution of the highly multiplexed tandem mass spectra, and difficulties in controlling the false discovery rate. Nevertheless, DISPA continues to demonstrate its potentially high efficiency as an ultrafast proteome analysis method. In the recent study, a throughput of 1000 samples per day (1.4 min per sample) was demonstrated, profiling blood plasma proteins in the corona of nanoparticles with identification of 280 proteins, 44 of which were confirmed biomarkers of various pathologies [71].

DIRECT PROTEIN IDENTIFICATION METHOD

As mentioned above, one of the reasons for long duration of the standard proteome-wide analysis using DDA is the need to obtain fragmentation spectra from as many peptide ions eluting from the chromatographic column as possible. The dominant fragmentation method is peptide backbone dissociation in collisions with carrier gas molecules at -CO-NH- bonds, mostly forming y- and b-series fragments. The peptide collision-activated dissociation process is ergodic, requiring time for breaking bonds. Additionally, to enhance information content of the fragmentation spectra and, hence, increase accuracy of the corresponding peptide sequence identification, it is desirable to obtain as many fragment ions as possible, which also requires time for accumulation of the isolated precursor ions for this purpose. Thus, during the analysis of a proteolytic mixture by DDA, for each MS1 spectrum acquired in a wide *m/z* range and registering all peptide ions eluting at a given time, the sequential selection of a limited number of the most intense precursor ions, their accumulation, and fragmentation is performed. In the case of complex organism proteomes, the analyzed proteolytic mixture can contain millions of individual peptide sequences. Accordingly, to increase the number of peptide ions selected for fragmentation and, consequently, depth of the proteome coverage by the analysis, and considering time constraints imposed by the ion accumulation rates and tandem mass spectrum acquisition, the multi-hour HPLC gradients are used. Thus, in order to increase efficiency of the proteome-wide characterization, the concept of direct mass spectrometric protein identification was suggested, in which the peptide ion fragmentation step of the analysis is omitted. Proteins are identified directly from the MS1 spectra based on the accurately measurable masses of all ions in the spectra and taking into account specificity of the protein digestion. At the same time, intensity distribution of the registered peptide ion peaks in the mass spectrum, corresponding to different ^{13}C isotope content in the sequence (peptide ion spectral profile), provides information specific to this peptide's elemental composition [72]. Simultaneously, retention times are specific to the peptide amino acid sequences [73-75], including peptides with residue modifications [76, 77]. Thus, MS1 spectra allow extracting a set of complementary data about the peptide sequence and elemental composition. Obviously, omission of peptide fragmentation reduces capabilities of their identification due to the significantly lower specificity of chromatographic times to amino acid sequences compared to the tandem mass spectra. However, omission of the fragmentation stage allows significant shortening the time of analysis by using short HPLC gradients.

The concept of direct mass spectrometric protein identification was implemented in the DirectMS1 method, which in early works demonstrated depth of the proteome-wide analysis of over 1000 proteins using 5-minute separation gradients [78]. Schematic representation of the method implementation is presented in Fig. 4. Mass analyzer operates in the mode of continuous MS1 spectra acquisition throughout the entire gradient elution time. Speed of the MS1 spectra acquisition depends on the mass analyzer type and requirements for the mass resolution and measurement accuracy. These requirements are high: at least 100,000 mass resolving power and measurement accuracy of less than 1 ppm. Another key factor affecting efficiency of the method is scanning speed of the mass analyzer and accuracy of the prediction of peptide elution times, which are used in the DirectMS1 data processing algorithm to distinguish correct and false identifications. Until recently, several models and algorithms for the peptide retention time prediction existed with prediction accuracy (correlation between experimental and predicted times) of $R^2 \sim 0.96$ for the Pearson coefficient [32]. In recent years, with the development of machine learning algorithms, peptide retention time prediction models of significantly higher accuracy have emerged. Specifically, the DirectMS1 search algorithm uses the DeepLC prediction model, which has substantially increased the proteome coverage depth to over 2000 identified proteins using 5-minute HPLC gradients and 7.5 min total time per experimental run [79]. Further increase in the number of identifiable proteins is achieved by adding peptide ion mobility separation. The DirectMS1 method implementation does not require significant changes in the instrument, except for the need for higher

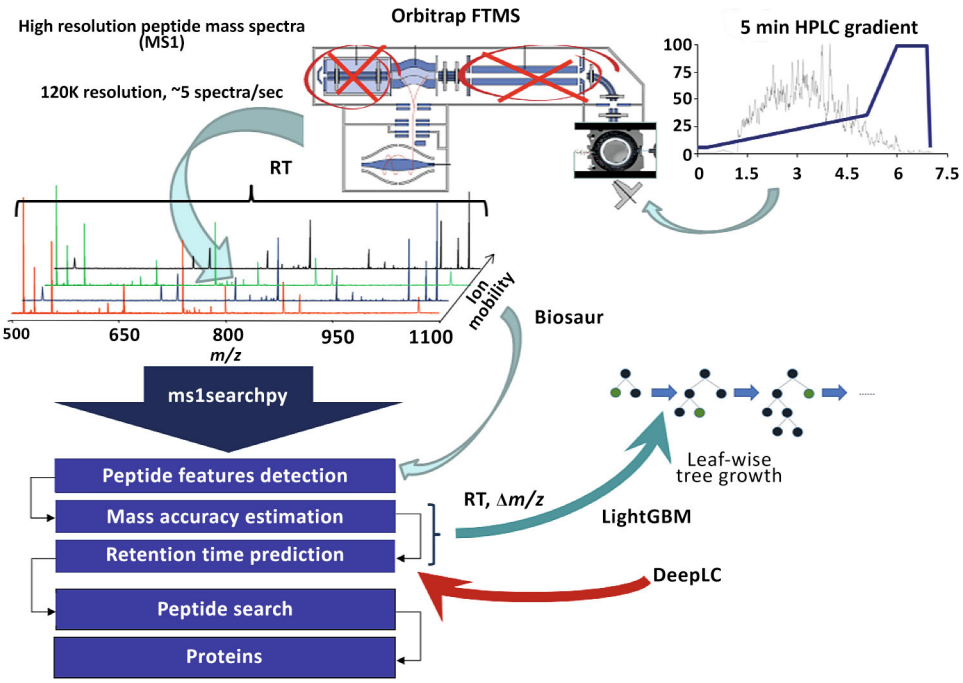

Fig. 4. Diagram of the DirectMS1 experiment. Key factor in the method's efficiency is the use of machine learning algorithms for predicting peptide retention times and classifying correct and false identifications based on combination of the complementary data, such as the peptide ion mass spectral ^{13}C profile, retention times, ion mobility (when ion mobility separation is added to the workflow), and measured accurate peptide masses.

HPLC flow rates, up to 1 µl/min or more, in order to maintain chromatographic resolution under ultrashort gradient conditions. Software for processing of the peptide ion mass spectra, ranking identifications, correlating them with proteins in the corresponding databases, and determining confidence levels is the key part of the method. This task is performed using the tools for determination of the peptide ion spectral profiles in the MS1 spectra, such as Biosaur [80], and protein identification *ms1searchpy* [81]; the latter is based on the machine learning algorithms and integrated with the peptide retention time prediction models. It should be noted that the drawback of the DirectMS1 method is lack of the FDR control at the peptide level. According to the authors of the method, the level of false positive peptide identifications can reach 30% [78]. Importantly, in the ultra-short gradient mode of separation, unlike in the MS/MS-based approaches, the DirectMS1 method allows identifying proteins with significantly (almost an order of magnitude) greater sequence coverage. This, in turn, provides more accurate measurements of the changes in protein concentrations. Notably, despite the lower proteome coverage depth, the DirectMS1 method allows protein quantitation in ultrafast analysis with efficiency comparable to the long HPLC gradient DIA and DDA methods [20], and it was successfully applied for identifying differentially expressed proteins in the cellular response to drug treatment [82].

CONCLUSIONS

Currently, we observe an active development of the technologies for proteome-wide analysis based on mass spectrometry and their application in various fields of post-genomic research. However, throughput of this analysis, which typically takes hours of experimental time for quantitative profiling of a single proteolytic mixture, is one of the main factors behind the limited use of proteomics in many areas of biomedical studies. These areas include drug development and repurposing, personalized medicine, population and clinical proteomics, single-cell proteomics, and more. Advancement of the high throughput high resolution mass spectrometry technologies, as well as new data processing methods based on machine learning algorithms, increased the throughput to several hundred whole proteome analyses per day. These capabilities have been realized in proteomics in just the last few years, and ultrafast proteomics methods are now rapidly evolving to become dominant approaches in addressing many of the above-mentioned problems and emerging areas of post-genomic research. Methods such as DIA, DirectMS1, and DISPA not only reduce analysis time by more than an order of magnitude but also increase its depth to the levels of 2000 to 5000 proteins identified within 3 to 5 min of total experimental time, which was unimaginable a decade ago. Further development of the technologies and methods

for ultrafast proteome-wide analysis will allow large-scale studies on large sample cohorts with less time, enabling more efficient determination of protein interaction mechanisms and cellular changes at the proteome level resulting from pathological processes, or under the influence of chemotherapeutic and external factors.

Contributions. I.I.F. and S.A.P. literature review and analysis, manuscript writing; I.A.T. discussion of the content and structure of the review; M.V.G. supervision of the work, manuscript writing and editing.

Funding. This work was financially supported by the Russian Science Foundation, grant no. 20-14-00229.

Ethics declarations. This work does not contain any studies involving human and animal subjects performed by any of the authors. The authors of this work declare that they have no conflicts of interest.

REFERENCES

1. Wang, H., Dey, K. K., Chen, P. C., Li, Y., Niu, M., Cho, J. H., Wang, X., Bai, B., Jiao, Y., Chepyala, S. R., Haroutunian, V., Zhang, B., Beach, T. G., and Peng, J. (2020) Integrated analysis of ultra-deep proteomes in cortex, cerebrospinal fluid and serum reveals a mitochondrial signature in Alzheimer's disease, *Mol. Neurodegener.*, **15**, 43, https://doi.org/10.1186/s13024-020-00384-6.
2. Radovanovic, M., Day, R. O., Jones, G. D. R., Galettis, P., and Norris, R. L. G. (2022) LC-MS/MS method for simultaneous quantification of ten antibiotics in human plasma for routine therapeutic drug monitoring, *J. Mass Spectrom. Adv. Clin. Lab.*, **26**, 48-59, https://doi.org/10.1016/j.jmsacl.2022.11.001.
3. Desiere, F., Deutsch, E. W., King, N. L., Nesvizhskii, A. I., Mallick, P., Eng, J., Chen, S., Eddes, J., Loevenich, S. N., and Aebersold, R. (2006) The PeptideAtlas project, *Nucleic Acids Res.*, **34**, D655-D658, https://doi.org/10.1093/nar/gkj040.
4. Henzel, W. J., Watanabe, C., and Stults, J. T. (2003) Protein identification: the origins of peptide mass fingerprinting, *J. Am. Soc. Mass Spectrom.*, **14**, 931-942, https://doi.org/10.1016/s1044-0305(03)00214-9.
5. Lay, J. O., Jr. (2001) MALDI-TOF mass spectrometry of bacteria, *Mass Spectrom. Rev.*, **20**, 172-194, https://doi.org/10.1002/mas.10003.
6. Aggarwal, S., and Yadav, A. K. (2016) False Discovery Rate Estimation in Proteomics, in *Statistical Analysis in Proteomics* (Jung, K., ed.) Springer, New York, NY, pp. 119-128.
7. Elias, J. E., and Gygi, S. P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry, *Nat. Methods*, **4**, 207-214, https://doi.org/10.1038/nmeth1019.
8. Shen, Y., Tolić, N., Masselon, C., Pasa-Tolić, L., Camp, D. G., 2nd, Hixson, K. K., Zhao, R., Anderson, G. A., and Smith, R. D. (2004) Ultrasensitive proteomics using high-efficiency on-line micro-SPE-nanoLC-nanoESI MS and MS/MS, *Anal. Chem.*, **76**, 144-154, https://doi.org/10.1021/ac030096q.
9. Shen, Y., Zhao, R., Berger, S. J., Anderson, G. A., Rodriguez, N., and Smith, R. D. (2002) High-efficiency nanoscale liquid chromatography coupled on-line with mass spectrometry using nanoelectron spray ionization for proteomics, *Anal. Chem.*, **74**, 4235-4249, https://doi.org/10.1021/ac0202280.
10. Richards, A. L., Hebert, A. S., Ulbrich, A., Bailey, D. J., Coughlin, E. E., Westphall, M. S., and Coon, J. J. (2015) One-hour proteome analysis in yeast, *Nat. Protoc.*, **10**, 701-714, https://doi.org/10.1038/nprot.2015.040.
11. Hebert, A. S., Prasad, S., Belford, M. W., Bailey, D. J., McAlister, G. C., Abbatiello, S. E., Huguet, R., Wouters, E. R., Dunyach, J. J., Brademan, D. R., Westphall, M. S., and Coon, J. J. (2018) Comprehensive single-shot proteomics with FAIMS on a hybrid orbitrap mass spectrometer, *Anal. Chem.*, **90**, 9529-9537, https://doi.org/10.1021/acs.analchem.8b02233.
12. Zubarev, R. A. (2013) The challenge of the proteome dynamic range and its implications for in-depth proteomics, *Proteomics*, **13**, 723-726, https://doi.org/10.1002/pmic.201200451.
13. Doerr, A. (2015) DIA mass spectrometry, *Nat. Methods*, **12**, 35-35, https://doi.org/10.1038/nmeth.3234.
14. Geyer, P. E., Kulak, N. A., Pichler, G., Holdt, L. M., Teupser, D., and Mann, M. (2016) Plasma proteome profiling to assess human health and disease, *Cell Syst.*, **2**, 185-195, https://doi.org/10.1016/j.cels.2016.02.015.
15. Pirmoradian, M., Budamgunta, H., Chingin, K., Zhang, B., Astorga-Wells, J., and Zubarev, R. A. (2013) Rapid and deep human proteome analysis by single-dimension shotgun proteomics, *Mol. Cell Proteomics*, **12**, 3330-3338, https://doi.org/10.1074/mcp.o113.028787.
16. Nagaraj, N., Wisniewski, J. R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Pääbo, S., and Mann, M. (2011)

Deep proteome and transcriptome mapping of a human cancer cell line, *Mol. Syst. Biol.*, **7**, 548, https://doi.org/10.1038/msb.2011.81.

17. Masselon, C., Anderson, G. A., Harkewicz, R., Bruce, J. E., Pasa-Tolic, L., and Smith, R. D. (2000) Accurate mass multiplexed tandem mass spectrometry for high-throughput polypeptide identification from mixtures, *Anal. Chem.*, **72**, 1918-1924, https://doi.org/10.1021/ac991133+.
18. Chapman, J. D., Goodlett, D. R., and Masselon, C. D. (2014) Multiplexed and data-independent tandem mass spectrometry for global proteome profiling, *Mass Spectrom. Rev.*, **33**, 452-470, https://doi.org/10.1002/mas.21400.
19. Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A. K., and Hamon, C. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS, *Anal. Chem.*, **75**, 1895-1904, https://doi.org/10.1021/ac0262560.
20. Ivanov, M. V., Bubis, J. A., Gorshkov, V., Tarasova, I. A., Levitsky, L. I., Solovyeva, E. M., Lipatova, A. V., Kjeldsen, F., and Gorshkov, M. V. (2022) DirectMS1Quant: ultrafast quantitative proteomics with MS/MS-free mass spectrometry, *Anal. Chem.*, **94**, 13068-13075, https://doi.org/10.1021/acs.analchem.2c02255.
21. Aebersold, R., and Goodlett, D. R. (2001) Mass spectrometry in proteomics, *Chem. Rev.*, **101**, 269-295, https://doi.org/10.1021/cr990076h.
22. Pappin, D. J., Hojrup, P., and Bleasby, A. J. (1993) Rapid identification of proteins by peptide-mass fingerprinting, *Curr. Biol.*, **3**, 327-332, https://doi.org/10.1016/0960-9822(93)90195-t.
23. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics, *Nature*, **422**, 198-207, https://doi.org/10.1038/nature01511.
24. Gevaert, K., and Vandekerckhove, J. (2000) Protein identification methods in proteomics, *Electrophoresis*, **21**, 1145-1154, https://doi.org/10.1002/(sici)1522-2683(20000401)21:6<1145::aid-elps1145>3.0.co;2-z.
25. He, Z., Yang, C., Yang, C., Qi, R. Z., Tam, J. P., and Yu, W. (2010) Optimization-based peptide mass fingerprinting for protein mixture identification, *J. Comput. Biol.*, **17**, 221-235, https://doi.org/10.1089/cmb.2009.0160.
26. Conrads, T. P., Anderson, G. A., Veenstra, T. D., Pasa-Tolić, L., and Smith, R. D. (2000) Utility of accurate mass tags for proteome-wide protein identification, *Anal. Chem.*, **72**, 3349-3354, https://doi.org/10.1021/ac0002386.
27. Norbeck, A. D., Monroe, M. E., Adkins, J. N., Anderson, K. K., Daly, D. S., and Smith, R. D. (2005) The utility of accurate mass and LC elution time information in the analysis of complex proteomes, *J. Am. Soc. Mass Spectrom.*, **16**, 1239-1249, https://doi.org/10.1016/j.jasms.2005.05.009.
28. Shen, Y., Strittmatter, E. F., Zhang, R., Metz, T. O., Moore, R. J., Li, F., Udseth, H. R., Smith, R. D., Unger, K. K., Kumar, D., and Lubda, D. (2005) Making broad proteome protein measurements in 1-5 min using high-speed RPLC separations and high-accuracy mass measurements, *Anal. Chem.*, **77**, 7763-7773, https://doi.org/10.1021/ac051257o.
29. Shen, Y., Smith, R. D., Unger, K. K., Kumar, D., and Lubda, D. (2005) Ultrahigh-throughput proteomics using fast RPLC separations with ESI-MS/MS, *Anal. Chem.*, **77**, 6692-6701, https://doi.org/10.1021/ac050876u.
30. Lange, E., Tautenhahn, R., Neumann, S., and Gröpl, C. (2008) Critical assessment of alignment procedures for LC-MS proteomics and metabolomics measurements, *BMC Bioinformatics*, **9**, 375, https://doi.org/10.1186/1471-2105-9-375.
31. Petritis, K., Kangas, L. J., Yan, B., Monroe, M. E., Strittmatter, E. F., Qian, W. J., Adkins, J. N., Moore, R. J., Xu, Y., Lipton, M. S., Camp, D. G., 2nd, and Smith, R. D. (2006) Improved peptide elution time prediction for reversed-phase liquid chromatography-MS by incorporating peptide sequence information, *Anal. Chem.*, **78**, 5026-5039, https://doi.org/10.1021/ac060143p.
32. Tarasova, I. A., Masselon, C. D., Gorshkov, A. V., and Gorshkov, M. V. (2016) Predictive chromatography of peptides and proteins as a complementary tool for proteomics, *Analyst*, **141**, 4816-4832, https://doi.org/10.1039/c6an00919k.
33. Tarasova, I. A., Guryca, V., Pridatchenko, M. L., Gorshkov, A. V., Kieffer-Jaquinod, S., Evreinov, V. V., Masselon, C. D., and Gorshkov, M. V. (2009) Standardization of retention time data for AMT tag proteomics database generation, *J. Chromatogr. B*, **877**, 433-440, https://doi.org/10.1016/j.jchromb.2008.12.047.
34. Han, X., Aslanian, A., and Yates, J. R. (2008) Mass spectrometry for proteomics, *Curr. Opin. Chem. Biol.*, **12**, 483-490, https://doi.org/10.1016/j.cbpa.2008.07.024.
35. Zhang, Y., Fonslow, B. R., Shan, B., Baek, M. C., and Yates, J. R., 3rd (2013) Protein analysis by shotgun/bottom-up proteomics, *Chem. Rev.*, **113**, 2343-2394, https://doi.org/10.1021/cr3003533.
36. Miller, R. M., and Smith, L. M. (2023) Overview and considerations in bottom-up proteomics, *Analyst*, **148**, 475-486, https://doi.org/10.1039/d2an01246d.
37. Wang, H., Yang, Y., Li, Y., Bai, B., Wang, X., Tan, H., Liu, T., Beach, T. G., Peng, J., and Wu, Z. (2015) Systematic optimization of long gradient chromatography mass spectrometry for deep analysis of brain proteome, *J. Proteome Res.*, **14**, 829-838, https://doi.org/10.1021/pr500882h.
38. Michalski, A., Cox, J., and Mann, M. (2011) More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS, *J. Proteome Res.*, **10**, 1785-1793, https://doi.org/10.1021/pr101060v.

39. Kreimer, S., Belov, M. E., Danielson, W. F., Levitsky, L. I., Gorshkov, M. V., Karger, B. L., and Ivanov, A. R. (2016) Advanced precursor ion selection algorithms for increased depth of bottom-up proteomic profiling, *J. Proteome Res.*, **15**, 3563-3573, https://doi.org/10.1021/acs.jproteome.6b00312.
40. Hebert, A. S., Thöing, C., Riley, N. M., Kwiecien, N. W., Shiskova, E., Huguet, R., Cardasis, H. L., Kuehn, A., Eliuk, S., Zabrouskov, V., Westphall, M. S., McAlister, G. C., and Coon, J. J. (2018) Improved precursor characterization for data-dependent mass spectrometry, *Anal. Chem.*, **90**, 2333-2340, https://doi.org/10.1021/acs.analchem.7b04808.
41. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014) Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells, *Nat. Methods*, **11**, 319-324, https://doi.org/10.1038/nmeth.2834.
42. Zhang, B., Pirmoradian, M., Chernobrovkin, A., and Zubarev, R. A. (2014) DeMix workflow for efficient identification of cofragmented peptides in high resolution data-dependent tandem mass spectrometry, *Mol. Cell Proteomics*, **13**, 3211-3223, https://doi.org/10.1074/mcp.o114.038877.
43. Angel, T. E., Aryal, U. K., Hengel, S. M., Baker, E. S., Kelly, R. T., Robinson, E. W., and Smith, R. D. (2012) Mass spectrometry-based proteomics: existing capabilities and future directions, *Chem. Soc. Rev.*, **41**, 3912-3928, https://doi.org/10.1039/c2cs15331a.
44. Stanley, J. R., Adkins, J. N., Slysz, G. W., Monroe, M. E., Purvine, S. O., Karpievitch, Y. V., Anderson, G. A., Smith, R. D., and Dabney, A. R. (2011) A statistical method for assessing peptide identification confidence in accurate mass and time tag proteomics, *Anal. Chem.*, **83**, 6135-6140, https://doi.org/10.1021/ac2009806.
45. Zimmer, J. S. D., Monroe, M. E., Qian, W.-J., and Smith, R. D. (2006) Advances in proteomics data analysis and display using an accurate mass and time tag approach, *Mass Spectrom. Rev.*, **25**, 450-482, https://doi.org/10.1002/mas.20071.
46. Dos Santos, A., Court, M., Thiers, V., Sar, S., Guettier, C., Samuel, D., Bréchot, C., Garin, J., Demaugre, F., and Masselon, C. D. (2010) Identification of cellular targets in human intrahepatic cholangiocarcinoma using laser microdissection and accurate mass and time tag proteomics, *Mol. Cell Proteomics*, **9**, 1991-2004, https://doi.org/10.1074/mcp.m110.000026.
47. Agron, I. A., Avtonomov, D. M., Kononikhin, A. S., Popov, I. A., Moshkovskii, S. A., and Nikolaev, E. N. (2010) Accurate mass tag retention time database for urine proteome analysis by chromatography-mass spectrometry, *Biochemistry (Moscow)*, **75**, 636-641, https://doi.org/10.1134/s0006297910050147.
48. Heil, L. R., Damoc, E., Arrey, T. N., Pashkova, A., Denisov, E., Petzoldt, J., Peterson, A. C., Hsu, C., Searle, B. C., Shulman, N., Riffle, M., Connolly, B., MacLean, B. X., Remes, P. M., Senko, M. W., Stewart, H. I., Hock, C., Makarov, A. A., Hermanson, D., Zabrouskov, V., Wu, C. C., and MacCoss, M. J. (2023) Evaluating the performance of the astral mass analyzer for quantitative proteomics using data-independent acquisition, *J. Proteome Res.*, **22**, 3290-3300, https://doi.org/10.1021/acs.jproteome.3c00357.
49. Guzman, U. H., Martinez-Val, A., Ye, Z., Damoc, E., Arrey, T. N., Pashkova, A., Renuse, S., Denisov, E., Petzoldt, J., Peterson, A. C., Harking, F., Østergaard, O., Rydbirk, R., Aznar, S., Stewart, H., Xuan, Y., Hermanson, D., Horning, S., Hock, C., Makarov, A., Zabrouskov, V., and Olsen, J. V. (2024) Ultra-fast label-free quantification and comprehensive proteome coverage with narrow-window data-independent acquisition, *Nat. Biotechnol.*, https://doi.org/10.1038/s41587-023-02099-7.
50. Gillet, L. C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., Bonner, R., and Aebersold, R. (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis, *Mol. Cell Proteomics*, **11**, O111.016717, https://doi.org/10.1074/mcp.o111.016717.
51. Bruderer, R., Bernhardt, O. M., Gandhi, T., Miladinović, S. M., Cheng, L. Y., Messner, S., Ehrenberger, T., Zanotelli, V., Butscheid, Y., Escher, C., Vitek, O., Rinner, O., and Reiter, L. (2015) Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues, *Mol. Cell Proteomics*, **14**, 1400-1410, https://doi.org/10.1074/mcp.m114.044305.
52. Kelstrup, C. D., Bekker-Jensen, D. B., Arrey, T. N., Hogrebe, A., Harder, A., and Olsen, J. V. (2018) Performance evaluation of the Q Exactive HF-X for shotgun proteomics, *J. Proteome Res.*, **17**, 727-738, https://doi.org/10.1021/acs.jproteome.7b00602.
53. Doellinger, J., Blumenscheit, C., Schneider, A., and Lasch, P. (2020) Isolation window optimization of data-independent acquisition using predicted libraries for deep and accurate proteome profiling, *Anal. Chem.*, **92**, 12185-12192, https://doi.org/10.1021/acs.analchem.0c00994.
54. Doellinger, J., Blumenscheit, C., Schneider, A., and Lasch, P. (2023) Increasing proteome depth while maintaining quantitative precision in short-gradient data-independent acquisition proteomics, *J. Proteome Res.*, **22**, 2131-2140, https://doi.org/10.1021/acs.jproteome.3c00078.
55. Bekker-Jensen, D. B., Martínez-Val, A., Steigerwald, S., Rüther, P., Fort, K. L., Arrey, T. N., Harder, A., Makarov, A., and Olsen, J. V. (2020) A compact quadrupole-orbitrap mass spectrometer with FAIMS interface improves proteome coverage in short LC

gradients, *Mol. Cell Proteomics*, **19**, 716-729, https://doi.org/10.1074/mcp.tir119.001906.

56. Yang, Y., Liu, X., Shen, C., Lin, Y., Yang, P., and Qiao, L. (2020) In silico spectral libraries by deep learning facilitate data-independent acquisition proteomics, *Nat. Commun.*, **11**, 146, https://doi.org/10.1038/s41467-019-13866-z.
57. Lou, R., Liu, W., Li, R., Li, S., He, X., and Shui, W. (2021) DeepPhospho accelerates DIA phosphoproteome profiling through *in silico* library generation, *Nat. Commun.*, **12**, 6685, https://doi.org/10.1038/s41467-021-26979-1.
58. Van Puyvelde, B., Willems, S., Gabriels, R., Daled, S., De Clerck, L., Vande Casteele, S., Staes, A., Impens, F., Deforce, D., Martens, L., Degroeve, S., and Dhaenens, M. (2020) Removing the hidden data dependency of DIA with predicted spectral libraries, *Proteomics*, **20**, e1900306, https://doi.org/10.1002/pmic.201900306.
59. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S., and Ralser, M. (2020) DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput, *Nat. Methods*, **17**, 41-44, https://doi.org/10.1038/s41592-019-0638-x.
60. Moseley, M. A., Hughes, C. J., Juvvadi, P. R., Soderblom, E. J., Lennon, S., Perkins, S. R., Thompson, J. W., Steinbach, W. J., Geromanos, S. J., Wildgoose, J., Langridge, J. I., Richardson, K., and Vissers, J. P. C. (2018) Scanning quadrupole data-independent acquisition, part A: qualitative and quantitative characterization, *J. Proteome Res.*, **17**, 770-779, https://doi.org/10.1021/acs.jproteome.7b00464.
61. Messner, C. B., Demichev, V., Bloomfield, N., Yu, J. S. L., White, M., Kreidl, M., Egger, A. S., Freiwald, A., Ivosev, G., Wasim, F., Zelezniak, A., Jürgens, L., Suttorp, N., Sander, L. E., Kurth, F., Lilley, K. S., Mülleder, M., Tate, S., and Ralser, M. (2021) Ultra-fast proteomics with Scanning SWATH, *Nat. Biotechnol.*, **39**, 846-854, https://doi.org/10.1038/s41587-021-00860-4.
62. Fernandez-Lima, F., Kaplan, D. A., Suetering, J., and Park, M. A. (2011) Gas-phase separation using a trapped ion mobility spectrometer, *Int. J. Ion Mobil. Spectrom.*, **14**, 93-98, https://doi.org/10.1007/s12127-011-0067-8.
63. Ridgeway, M. E., Lubeck, M., Jordens, J., Mattias, M., and Melvin, A. P. (2018) Trapped ion mobility spectrometry: a short review, *Int. J. Mass Spectrom.*, **425**, 22-35, https://doi.org/10.1016/j.ijms.2018.01.006.
64. Meier, F., Beck, S., Grassl, N., Lubeck, M., Park, M. A., Raether, O., and Mann, M. (2015) Parallel Accumulation-Serial Fragmentation (PASEF): multiplying sequencing speed and sensitivity by synchronized scans in a trapped ion mobility device, *J. Proteome Res.*, **14**, 5378-5387, https://doi.org/10.1021/acs.jproteome.5b00932.
65. Meier, F., Brunner, A. D., Koch, S., Koch, H., Lubeck, M., Krause, M., Goedecke, N., Decker, J., Kosinski, T., Park, M. A., Bache, N., Hoerning, O., Cox, J., Räther, O., and Mann, M. (2018) Online parallel accumulation-serial fragmentation (PASEF) with a novel trapped ion mobility mass spectrometer, *Mol. Cell Proteomics*, **17**, 2534-2545, https://doi.org/10.1074/mcp.tir118.000900.
66. Szyrwiel, L., Gille, C., Mülleder, M., Demichev, V., and Ralser, M. (2024) Fast proteomics with dia-PASEF and analytical flow-rate chromatography, *Proteomics*, **24**, e2300100, https://doi.org/10.1002/pmic.202300100.
67. Chen, S. (2006) Rapid protein identification using direct infusion nanoelectrospray ionization mass spectrometry, *Proteomics*, **6**, 16-25, https://doi.org/10.1002/pmic.200500043.
68. Meyer, J. G., Niemi, N. M., Pagliarini, D. J., and Coon, J. J. (2020) Quantitative shotgun proteome analysis by direct infusion, *Nat. Methods*, **17**, 1222-1228, https://doi.org/10.1038/s41592-020-00999-z.
69. Cranney, C. W., and Meyer, J. G. (2021) CsoDIAq software for direct infusion shotgun proteome analysis, *Anal. Chem.*, **93**, 12312-12319, https://doi.org/10.1021/acs.analchem.1c02021.
70. Jiang, Y., Hutton, A., Cranney, C. W., and Meyer, J. G. (2023) Label-free quantification from Direct Infusion Shotgun Proteome Analysis (DISPA-LFQ) with CsoDIAq software, *Anal. Chem.*, **95**, 677-685, https://doi.org/10.1021/acs.analchem.2c02249.
71. Jiang, Y., and Meyer, J. G. (2024) 1.4 min plasma proteome profiling via nanoparticle protein corona and direct infusion mass spectrometry, *bioRxiv*, https://doi.org/10.1101/2024.02.06.579213.
72. Miladinović, S. M., Kozhinov, A. N., Gorshkov, M. V., and Tsybin, Y. O. (2012) On the utility of isotopic fine structure mass spectrometry in protein identification, *Anal. Chem.*, **84**, 4042-4051, https://doi.org/10.1021/ac2034584.
73. Krokhin, O. V. (2006) Sequence-specific retention calculator. Algorithm for peptide retention prediction in ion-pair RP-HPLC: application to 300- and 100-A pore size C18 sorbents, *Anal. Chem.*, **78**, 7785-7795, https://doi.org/10.1021/ac060777w.
74. Gorshkov, A. V., Tarasova, I. A., Evreinov, V. V., Savitski, M. M., Nielsen, M. L., Zubarev, R. A., and Gorshkov, M. V. (2006) Liquid chromatography at critical conditions: comprehensive approach to sequence-dependent retention time prediction, *Anal. Chem.*, **78**, 7770-7777, https://doi.org/10.1021/ac060913x.
75. Pridatchenko, M. L., Perlova, T. Y., Ben Hamidane, H., Goloborodko, A. A., Tarasova, I. A., Gorshkov, A. V., Evreinov, V. V., Tsybin, Y. O., and Gorshkov, M. V. (2012) On the utility of predictive chromatography to complement mass spectrometry based intact protein identification, *Anal. Bioanal. Chem.*, **402**, 2521-2529, https://doi.org/10.1007/s00216-011-5350-3.
76. Perlova, T. Y., Goloborodko, A. A., Margolin, Y., Pridatchenko, M. L., Tarasova, I. A., Gorshkov, A. V., Moskovets, E., Ivanov, A. R., and Gorshkov, M. V. (2010)

Retention time prediction using the model of liquid chromatography of biomacromolecules at critical conditions in LC-MS phosphopeptide analysis, *Proteomics*, **10**, 3458-3468, https://doi.org/10.1002/pmic.200900837.

77. Sargaeva, N. P., Goloborodko, A. A., O'Connor, P. B., Moskovets, E., and Gorshkov, M. V. (2011) Sequence-specific predictive chromatography to assist mass spectrometric analysis of asparagine deamidation and aspartate isomerization in peptides, *Electrophoresis*, **32**, 1962-1969, https://doi.org/10.1002/elps.201000507.
78. Ivanov, M. V., Bubis, J. A., Gorshkov, V., Tarasova, I. A., Levitsky, L. I., Lobas, A. A., Solovyeva, E. M., Pridatchenko, M. L., Kjeldsen, F., and Gorshkov, M. V. (2020) DirectMS1: MS/MS-Free identification of 1000 proteins of cellular proteomes in 5 minutes, *Anal. Chem.*, **92**, 4326-4333, https://doi.org/10.1021/acs.analchem.9b05095.
79. Ivanov, M. V., Bubis, J. A., Gorshkov, V., Abdrakhimov, D. A., Kjeldsen, F., and Gorshkov, M. V. (2021) Boosting MS1-only proteomics with machine learning allows 2000 protein identifications in single-shot human proteome analysis using 5 min HPLC gradient, *J. Proteome Res.*, **20**, 1864-1873, https://doi.org/10.1021/acs.jproteome.0c00863.
80. Abdrakhimov, D. A., Bubis, J. A., Gorshkov, V., Kjeldsen, F., Gorshkov, M. V., and Ivanov, M. V. (2021) Biosaur: an open-source Python software for liquid chromatography-mass spectrometry peptide feature detection with ion mobility support, *Rapid Commun. Mass Spectrom.*, https://doi.org/10.1002/rcm.9045.
81. Ivanov, M. V., Tarasova, I. A., Levitsky, L. I., Solovyeva, E. M., Pridatchenko, M. L., Lobas, A. A., Bubis, J. A., and Gorshkov, M. V. (2017) MS/MS-free protein identification in complex mixtures using multiple enzymes with complementary specificity, *J. Proteome Res.*, **16**, 3989-3999, https://doi.org/10.1021/acs.jproteome.7b00365.
82. Solovyeva, E. M., Bubis, J. A., Tarasova, I. A., Lobas, A. A., Ivanov, M. V., Nazarov, A. A., Shutkov, I. A., and Gorshkov, M. V. (2022) On the feasibility of using an ultra-fast DirectMS1 method of proteome-wide analysis for searching drug targets in chemical proteomics, *Biochemistry (Moscow)*, **87**, 1342-1353, https://doi.org/10.1134/s000629792211013x.