Published in final edited form as:

Expert Rev Proteomics. 2017 May ; 14(5): 419–429. doi:10.1080/14789450.2017.1322904.

Clinical applications of quantitative proteomics using targeted and untargeted data-independent acquisition techniques

Jesse G. Meyer and Birgit Schilling

Mass Spectrometry Core, Buck Institute for Research on Aging, Novato, CA, USA

Abstract

Introduction—While selected/multiple-reaction monitoring (SRM or MRM) is considered the gold standard for quantitative protein measurement, emerging data-independent acquisition (DIA) using high-resolution scans have opened a new dimension of high-throughput, comprehensive quantitative proteomics. These newer methodologies are particularly well suited for discovery of biomarker candidates from human disease samples, and for investigating and understanding human disease pathways.

Areas covered—This article reviews the current state of targeted and untargeted DIA mass spectrometry-based proteomic workflows, including SRM, parallel-reaction monitoring (PRM) and untargeted DIA (e.g., SWATH). Corresponding bioinformatics strategies, as well as application in biological and clinical studies are presented.

Expert commentary—Nascent application of highly-multiplexed untargeted DIA, such as SWATH, for accurate protein quantification from clinically relevant and disease-related samples shows great potential to comprehensively investigate biomarker candidates and understand disease.

Keywords

Data-independent acquisition; SWATH; mass spectrometry; quantification; proteomics; clinic; biomarker; SRM; PRM

1. Introduction – data independent acquisition for quantitative proteomics

Liquid chromatography coupled to mass spectrometry (LC-MS) is well established as an extremely valuable research tool to quantify proteins in biological samples. Protein quantification can be done with a diverse collection of instruments that utilize fundamentally unique data acquisition strategies. These acquisition strategies mainly divide into two groups: data-dependent and data-independent. Data-dependent acquisition (DDA) uses knowledge obtained during the acquisition to decide what MS1 peptide precursors to subject

CONTACT: Birgit Schilling, bschilling@buckinstitute.org, Buck Institute for Research on Aging, 8001 Redwood Blvd Novato, CA 94945, USA.

Declaration of interest

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties

for fragmentation (MS/MS) in the collision cell. Data-independent acquisition (DIA), in contrast, performs predefined MS/MS fragmentation and data collection regardless of sample content, which allows for more sensitive and accurate protein quantification compared to DDA. DIA strategies can be further segregated into targeted or untargeted acquisitions. Targeted DIA methods fragment predefined precursor ions that correspond to the peptide analytes, usually at known (measured or predicted) retention times. Targeted DIA has become widely used in academic, pharmaceutical, and biotechnology research for quantification of small molecules (metabolites), peptides, and post-translational modifications (PTMs). For example, selected-reaction monitoring (SRM), a type of targeted DIA, is currently considered the gold standard method for mass spectrometric quantification due to its high accuracy and precision. SRM was declared 'Method of the Year' in 2012 by *Nature Methods* [1]. Biomedical, systems biology, and translational medicine applications of SRM have been recently reviewed [2–5]. In other recent reviews, Percy et al. discussed the challenges, requirements, and strategies for clinical translation of plasma protein biomarkers [4], while Parker et al. reviewed the SRM-based biomarker pipeline as it pertains to discovery, verification, and validation studies [5]. Multiplexed SRM technologies have been thoroughly tested in large multi-laboratory studies, such as in a study reported by Abbatiello et al. [6] that quantified cancer-relevant proteins across 11 different laboratories and 14 LC-MS systems. Best practices and guidelines for SRM studies were recently outlined by Carr et al. [7].

Many useful resources for developing sensitive and specific SRM assays can be found, for example, in SRM Atlas [8] or at the CPTAC Portal [9,10] as recently reviewed [11]. Kusebauch et al. have synthesized and measured 166,174 proteotypic peptides, and this data is now available in SRM Atlas to aid in development of targeted SRM assays for quantification of 99.7% of the 20,277 annotated human proteins [12]. This large resource includes proteins with splice variants, non-synonymous mutations, and PTMs.

Even though SRM assays can be highly multiplexed to quantify >100 proteins (typically <200) per acquisition using accurate retention time scheduling and high-speed triple-quadrupole instrumentation [13,14], additional quantitative mass spectrometric scan types have recently attracted great interest. Those include targeted DIA using high-resolution accurate-mass techniques, such as parallel-reaction monitoring (PRM) [15–18], and untargeted DIA (which we refer to as uDIA), such as SWATH [19]. These different quantitative assay types for proteomics research have been reviewed by Aebersold [20] and others [21]. Similar to public SRM assay libraries, SWATH Atlas contains a growing number of DIA spectral libraries [22].

A comparison of the most common DIA methods and their hierarchical relations is pictured in Figure 1. SRM and PRM are both types of targeted DIA, where the instrument isolates predefined precursor masses for fragmentation. In SRM, the third quadrupole monitors a single fragment with high sensitivity, whereas in PRM, the instrument records all fragments with high mass resolution. In contrast, with SWATH, or other uDIA workflows, the mass spectrometer measures all intact precursor ions followed by isolation of a much wider precursor window (between 2 and 200 m/z) for fragmentation that usually contains several peptide precursors. The instrument then measures all fragment ions obtained from the

precursor isolation window, before moving to the next precursor isolation window and measuring MS/MS fragmentation thereof, until the entire MS1 mass range is covered. These consecutive isolations are performed such that the instrument measures and fragments all precursor ions, all of which is completed in a cycle of ~2–3 s. This cycle of data-independent MS1 precursor and MS2 fragment scans is repeated throughout the entire LC-MS acquisition, similar to SRM and PRM experiments.

Predefined uDIA precursor isolation windows or segments were statically set to 25 m/z in first reports by Gillet et al., who used 32 windows per cycle to fragment all precursors between 400 and 1200 m/z [19]; however, more recently, most research groups have implemented methods using a greater number of smaller m/z windows per cycle for various instrument platforms, because this clearly reduces interferences and improves quantitative quality (for more detailed discussions and a MS parameter overview, see Table 1, or a previous summary by Sajic et al. [23]). The best quantification is achieved when variable window width is chosen so that the density of precursor ions in each of the isolation windows is equalized across the m/z range [24,25].

Because all possible precursor ions are fragmented at any time, uDIA allows for comprehensive and accurate quantification of any observable analyte in the sample; however, these complex, chimeric spectra make interpretation challenging. Chimeric MS/MS spectra produced from uDIA are often processed for peptide identification and quantification by comparisons to reference spectra or spectral libraries as originally described by Gillet et al. [19]; however, other strategies are also discussed in Section 2. Kockmann et al. [31] have made an effort to directly compare SRM, PRM, and uDIA across platforms within their own lab; however, the technological advances of uDIA methods improve rapidly, and newer high-resolution instrument platforms capable of uDIA have since emerged.

Using uDIA, quantification of thousands of proteins within complex matrices, such as plasma, biofluids, or tissue lysates, has become feasible. The introduction of uDIA has been enabled by technological innovations in MS hardware, such as efficient high-speed scanning at high resolution. uDIA also relies on a wealth of new software development in order to realize its full potential. One type of uDIA referred to as SWATH is performed using orthogonal quadrupole time-of-flight mass spectrometers, also known as SCIEX TripleTOF 5600 or 6600 [19,32–34]. uDIA workflows, however, are also possible on other instrument platforms, such as on Thermo Scientific Q-Exactive-type mass spectrometers [28], which are hybrid quadrupole-Orbitrap mass spectrometers that combine quadrupole precursor selection with high-resolution, accurate-mass (HR/AM) Orbitrap detection. Some of the most popular instrument platforms and their instrument settings and parameters for uDIA are summarized in Table 1. A novel uDIA workflow (WiSIM-DIA) has also been suggested for an Orbitrap Fusion Tribrid mass spectrometer in which three high-resolution/accurate mass SIM scans with wide isolation windows (200 m/z) were acquired for precursor ions (400–1000 m/z). In parallel with each SIM scan, 17 ion trap MS/MS with 12 m/z isolation windows were acquired (communication Thermo Scientific application note).

Several additional uDIA strategies have been introduced, and examples of uDIA were demonstrated as early as 2009, such as PAcIFIC [35] and XDIA [36]. These early methods, however, only provide strategies for unbiased identification and require external quantification methods. Additionally, Waters provides a unique uDIA strategy, MS^E which collects high-resolution spectra of all ions first at low and then at high collision energy [37]. Waters has also introduced a new DIA method named SONAR, which utilizes a wide, continuously sliding precursor window for fragmentation, thereby allowing easy correlation of precursor and fragment ions.

Given the growing application of uDIA to clinical proteomics, an overview of analysis options and exemplary clinical studies is useful. In the following sections, we will introduce several data analysis options that are available for uDIA data. Next, we describe several recent clinically relevant studies that use DIA quantification. Finally, we offer our synthesis of the most recent results in the field and give our projection with regard to the research priorities and anticipated developments that will continue to drive quantitative clinical proteomics.

2. Bioinformatic resources for analysis of data from uDIA

After uDIA workflows were established, a number of software tools have emerged for application-specific data processing (Table 2). Data analysis of uDIA can be analyzed in one of two general ways, targeted or untargeted. Targeted data analysis relies on pre-assembled MS/MS spectral libraries containing normalized retention times [38], and is done using tools such as OpenSWATH [39], Skyline [28,40], Spectronaut [38], SWATH 2.0 [41], and MSPLIT-DIA [42]. Using these tools, chimeric MS/MS spectra from uDIA are compared to spectral libraries either generated in-house or by other laboratories. Alternatively, untargeted data analysis approaches, such as DIA-Umpire [43,44], and Group-DIA [45], find co-eluting precursor and fragment ions to generate pseudo-tandem MS/MS spectra, which then can be searched by any database search engine. Since these untargeted data analysis approaches abrogate the need for a spectral library in certain cases and experimental designs, they allow the possibility of combined identification and quantification all from one uDIA. Large multi-laboratory studies have since emerged and used these new bioinformatics tools [25,26,46].

Due to the great interest in uDIA by the proteomics field, new ideas and alternative approaches for uDIA data processing and data mining of complex MS/MS spectra with co-fragmenting precursor ions have been discussed. For example, Ting et al. presented a perspective on spectrum-centric analysis versus peptide-centric analysis, the latter of which assesses the probability that a peptide is present without requiring a spectral library [47]. This ‘peptide-centric analysis’ has been implemented in the software PECAN, which tests directly for statistical probability that the query peptides are present [47]. Another tool recently introduced by our group makes use of uDIA and isotopic labeling to assess absolute stoichiometry of lysine acetylation and succinylation [53]. The use of uDIA to assess PTM stoichiometry allows the computation of occupancy for peptides with several lysines, which is not possible using previously reported precursor quantification methods.

Several post-quantification tools have also been introduced to assess statistical significance of abundance changes, such as MSstats [50]. MSstats, which has been implemented into Skyline as external tool [54], is useful for data visualization, normalization, and statistical assessment using linear models to infer changes. mapDIA [48] and noninterfering fragment ion (NOFI) ranking [51] are also able to test for statistical significance of abundance changes, but have the added ability to detect and remove fragment-level signals containing interferences. mapDIA removes interfered fragments using outlier detection and also by selecting the most abundant correlated fragments. NOFI ranking finds interferences using a multidimensional filter that includes observed/predicted ion rank and chromatographic features. Recently, another type of post-processing method for nonlinear cross-acquisition feature alignment called TRIC was introduced that improves false discovery rates when used with OpenSWATH [49].

Quality control (QC) and system suitability metrics of uDIA workflows are very similar to those used for SRM [55]. Sophisticated QC tools have been published and already found extensive applications in the field, such as Statistical Process Control in Proteomics (SProCoP) [56], that can be installed as external tool in Skyline; or Panorama AutoQC [57], which can automatically import raw data and system performance of any targeted (SRM/PRM) or uDIA.

Additional innovative aspects of DIA workflows continue to drive widespread adoption. Reusable spectral libraries for analysis of uDIA data have led to much greater depth and speed of data processing [58]. This is particularly true for clinical and preclinical human samples, which can make use of a repository of assays to quantify 10,000 human proteins using uDIA [58]. The discovery of endogenous standard peptides produced from digests of eukaryote proteins from yeast to human that can be used as internal retention time standards has also been useful to bridge data sets and spectral libraries between different labs [59]. In addition, large-scale initiatives are taking on tasks to expand shared community resources, for example, Zolg et al. recently published a study that synthesized and analyzed (LC-MS/MS) >330,000 synthetic tryptic peptides representing essentially all canonical human gene products [60,61]. Overall, uDIA has driven a lot of innovations not only in instrumentation but also in the bioinformatics tools; consequently, this mass spectrometric method now can be routinely applied to challenging biological questions. After uDIA workflows were established, a number of software tools have emerged for application-specific data processing (Table 2).

3. Application of DIA/SWATH workflows to clinically relevant projects

The versatility of the DIA workflows allows for investigating projects from very different human matrices, including plasma, tissue, bronchoalveolar lavage fluid (BALF), urine, vesicles containing secreted proteins, and others. Clinical mass spectrometry-based proteomic projects aim to identify and verify protein biomarkers that can differentiate a cohort of disease patients from a group of control individuals. This is done with a combination of candidate protein biomarker identification followed by quantitative verification from clinical samples, traditionally using SRM, or more recently using uDIA.

Several examples that use DIA quantification for biologically interesting projects are presented in the following.

Many studies focus on human plasma due to the ease of collection and the presence of many different secreted proteins contained in plasma that may be specific for diseases. Recently, a large-scale study was performed by Liu et al. who investigated the quantitative variability of 342 plasma proteins in a human twin population [33]. Using the highly accurate and reproducible SWATH mass spectrometry technique, 1904 peptides defining 342 unique plasma proteins in 232 plasma samples could be monitored in samples collected longitudinally from pairs of monozygotic and dizygotic twins at intervals of 2–7 years. In addition to the reproducibility assessments, 13 single nucleotide polymorphisms (SNPs) were identified and several of the corresponding proteins could be quantified in human plasma determining plasma protein levels. The same group also developed some SWATH (and SRM) assays for quantification of N-linked glycoproteins in human plasma [62]. Monitoring quantitative levels of glycopeptides has long been of particular interest in cancer biomarker studies in plasma, mainly as many cell surface proteins or secreted proteins from cells are glycosylated and potentially qualify as biomarkers.

uDIA has been presented using a range of sample types to search for biomarkers of several different cancer types. For example, a recent study used SWATH quantification of glycoproteins to differentiate normal, non-aggressive, aggressive, and metastatic cancer types by analyzing samples from human prostate tissue from biopsies [63]. Over 1400 glycosylation sites were identified and quantified, and 220 glycoproteins showed significant changes. In particular, N-acylethanolamine acid amidase and protein tyrosine kinase 7 were found to indicate distinct signatures for tumor aggressiveness [63], as a combined area under the ROC curve (AUROC = 0.801) score indicated from immunostained tissues. Biopsies from another type of cancer, esophageal squamous cell carcinoma, were analyzed with SWATH quantification to discover new biomarker candidates [64]. In addition to investigating biomarkers from biopsies, proteins from BALF hold great potential as lung cancer biomarkers due to new mildly invasive endoscopy-based collection techniques. Ortea et al. analyzed BALF from patients with lung adenocarcinoma using SWATH [65], and 44 proteins showed a higher than 3.75 fold-change relative to controls. In another study, the rapid label-free quantitative capabilities of uDIA were demonstrated for investigations of oral cancer biomarker proteins using conditioned media from cultured squamous cancer cell lines. Specifically, interleukin-8 and vascular endothelial growth factor A were highlighted as potential markers [66]. According to the authors, biomarker candidates will be followed up in serum/plasma of oral cancer patients.

Several clinical assays use urine samples to assess prognostic or disease markers, which exhibits interesting options and opportunities to apply data-independent workflows. Muntel et al. developed a fast uDIA method for comprehensive mapping of the urinary proteome that enabled large-scale urine proteomics studies. They used a short uDIA method with 30-min gradient to increase throughput, with a 2-s scan cycle time covering the mass range 400–1000 m/z (one full precursor scan and 24 uDIA MS/MS scans). The authors were able to detect an average of 1301 proteins per sample. Muntel et al. [30] were able to achieve ~8% coefficients of variation (CV) when they used an in-house generated urinary spectral

library. They attribute their high reproducibility to the use of a Q-Exactive HF platform with an ultrahigh-field Orbitrap mass analyzer which allowed them acquire MS/MS fragment ions at 30,000 resolution. This study identified some potential biomarker candidates for ovarian cyst and urinary tract infection (UTI). The high throughput of their method and accurate quantification of DIA allowed the authors to efficiently analyze 87 samples from patients admitted to the ER for abdominal pain in less than 4 days [30]. Indeed, two biomarker candidates appeared quite robust demonstrating high areas under the ROC curves (AUROC), such as AUROC = 0.91 for CYTB (cystatin-B), and which increased by 5.8-fold in the ovarian cyst cohort ($p = 1.3\text{e-}5$) versus all other samples. A potential UTI biomarker was identified as PERM (myeloperoxidase) where the protein amount was 97-fold increased versus control ($p = 1.1\text{e-}6$), and a solid AUROC of 0.968 [30].

To overcome difficulties that may emerge particularly with protein dynamic range in very complex biological matrices such as urine or serum, researchers have developed different techniques to isolate vesicles, such as exosomes, that contain secreted proteins. Exosome isolation was combined with SWATH by Kulkarni et al. to identify urinary and serum exosome biomarkers for radiation exposure [67]. Radiation responsive protein signatures were monitored and quantified using exosomes from various mouse biofluids, such as urine and serum. This quantitative exosome study revealed radiation- and time-dependent protein signatures after whole body irradiation of mice. Overall, the authors identified 23 biomarkers from urine and 24 biomarkers from serum exosomes that were differentially expressed. The high throughput and high multiplexing advantages of SWATH workflows are also exemplified in another recent study of drug-metabolizing enzymes and transporters in human intestine, liver, and kidney microsomes [68]. The authors compared uDIA (in this case SWATH) with targeted DIA methods SRM and PRM [68]. Interestingly, the study showed that absolute protein quantification using SWATH gave similar quantitative values as SRM or PRM, with SWATH data showing the best CV values. In fact, the authors also stated that most of the values determined by SWATH differed by less than 50% from those obtained by SRM or PRM.

Chang et al. reported in 2015 a SWATH analysis of the synaptic proteome in Alzheimer's disease [69]. They found using brain samples from the hippocampus and motor cortex that 30 proteins exhibited significant expression differences between Alzheimer's cases and controls. The proteins that significantly changed were involved in structural maintenance, signal transduction, autophagy, oxidative stress, or proteasome activity.

Many diseases can be monitored by the resulting changes of the immune response. Silva et al. described the generation of reference libraries from peripheral blood mononuclear cells [70] that play important roles in immune homeostasis. Such spectral libraries, this particular one characterizing 1102 proteins involved in diverse human diseases and in immune system-related pathways, are quite useful to the proteomics field because they can be re-used in subsequent uDIA work-flows. Another additional interesting application of uDIA to study the immune system was recently demonstrated by Haverland et al. [71] who quantified >3600 proteins from macrophages. It was revealed that uninfected and HIV-1-infected monocyte-derived macrophages showed altered levels of nucleic acid binding and regulatory proteins.

uDIA has even been applied to monitor longitudinal changes in cervicovaginal fluid (CVF) during menstrual cycle [72]. The authors employed the standard strategy of building an identification library using DDA, followed by quantification of ~200 identified proteins using SWATH. They show some clear examples of protein levels that change depending on the phase of the ovulatory cycle. Insights into proteomic changes of CVF and possible physiological mechanisms may prove useful in the future for biomarkers of fertility status.

The DIA high-throughput and high-multiplexing capabilities have proven particularly useful for projects of human pathogens: For example, uDIA has been used to monitor and quantify the proteome of pathogenic bacterial strains that have sequenced genomes (quantitative proteogenomics) [73]. One such study allowed new insights into dimorphic lifestyles of human gastric pathogen *Helicobacter pylori* bacteria that exist in two morphological forms: the culturable spiral form and the viable but non-culturable coccoid form [74].

In fact, Table 3 provides an overview and some key feature of some of the studies described above presenting some SRM, PRM and uDIA studies that were biologically/clinically relevant and that indicate biological sample matrix, instrument types, scan type, software used, key features, etc. uDIA studies typically perform relative, label-free quantification comparing multiple conditions versus absolute quantification as often done with SRM; however, the very large comprehensiveness of uDIA remains a key advantage.

More recently, uDIA workflows have been extended to quantification of PTMs [32,76–78]. Many diseases are known to cause changes in their dynamic PTM signaling profile, specifically in cancer affecting phosphorylation patterns which provides interesting experimental designs when using uDIA for their analysis [76,78,79]. Other studies have focused on histone modifications, such as in an elegant PRM study published by Jaffe and colleagues describing the targeted quantification histone PTMs [75], but also recent uDIA studies [80,81]. uDIA acquisitions for PTM work also enable strategies to determine PTM site occupancy as recently presented by Meyer et al. [52].

When utilizing DIA together with innovative sample preparation, such as pressure cycling technology (PCT)-assisted sample preparation [82], uDIA has proven a vital component in next-generation low-variability proteome quantification. In addition, the latter study defines and extends the boundary of PCT-SWATH for proteomic analysis of minute clinical samples (from ~ 50,000 cells or 0.2–0.5 mg of wet tissue) and surprisingly found that variability was lowest using the lowest amounts of input tissue. Newly emerging sample preparation technologies are definitely an important step in moving novel mass spectrometric scan types toward clinical use. In summary, DIA has proven to be a very reproducible proteomic quantification method that is appropriate to analyze clinical samples, for example, biopsy-level tissue samples or biofluids.

Robustness and repeatability has recently been shown in the large-scale cross-laboratory study by Hunter [83] and Collins et al. [26] that showed average CV of <15–20% in a large scale study. The average response curves performed as part of the latter study were very similar between sites, all exceeding 4.45 orders of linear dynamic range including all data points, with an average across sites of 4.6 orders of magnitude [83]. QC is implemented

similar to approaches chosen for SRM studies [55]. Typically, these large-scale DIA/SWATH experiments are very robust, and as they can be performed in a label-free environment, sample processing can be held to a minimum, which is important in clinical settings.

As demonstrated by the studies described above, uDIA workflows have been applied to samples from nearly all types of biological matrixes in diverse clinically relevant settings. Key advantages of uDIA methods over more traditional quantitative methods are improved throughput, and high multiplexing in an unbiased approach. The DIA method requires no upfront assay development time, as is necessary for SRM assays, and post-acquisition data processing allows for in-depth data mining after all acquisitions are finished, even years after the original MS study. Finally, since all spectral libraries acquired in any DDA or DIA study can potentially be used as library for quantification in subsequent studies, these studies have greater intrinsic value as shared resources.

4. Expert commentary

One major driving force behind many proteomics techniques is the application to translational research. There are two connected aspects of translational research well suited for proteomic studies: characterization of the protein profiles that represent a diseased state relative to healthy controls, and the ultimate goal, discovery, and verification of clinical biomarkers that define specific diseases or disease progression. uDIA methods are uniquely well suited to reach these goals. The feasibility of using uDIA has been demonstrated in many cases as described in Section 3.

uDIA offers several advantages over other quantitative MS assays are (1) very high multiplexing capabilities and comprehensiveness allowing relative quantification of thousands of analytes, (2) no need for pre-acquisition assay development—all assays can be either developed or changed post-acquisition due to the comprehensiveness of the data structure, (3) high analytical specificity and selectivity, (4) lack of requirement for isotopic labeling allowing application to human samples and reducing cost, and (5) reusability of once acquired DIA data, as the data can always be re-interrogated with different or updated spectral libraries. One other important aspect is that uDIA workflows have encouraged sharing reusable chromatographic libraries [84] and spectral libraries. For example, a spectral library containing assays for half of all predicted human proteins is freely available to the public [58]. These libraries are constantly expanded using newly published data-sets based on endogenous peptides, or even large collections of synthetic peptides [85]. The quantitative performance criteria toward robustness and repeatability were excellent with average CV of <15–20% in a large-scale cross-laboratory study [83] and with a dynamic range of >4.45 orders of magnitude across all sites. QC aspects and system suitability are in general very comparable to SRM system suitability considerations [55] to achieve optimal data quality, which is particularly important for clinical types of samples.

The biggest challenge currently to gain acceptance of this new technology among clinicians is to provide the necessary instrumentation with suitable software interface that will allow even non-expert operation of the MS system. Much more robust system operation including

automatic quality control (Panorama AutoQC) [57] and more stable electrospray, for example, by switching from nano-flow to micro-flow HPLC [86], will provide sufficient reliability for the typical day-today use. Altogether, continued improvements to reliability and usability of uDIA workflows will open numerous possibilities for clinical and pre-clinical studies.

Since the benefits of uDIA stem from the fact that fragmentation data is collected for all analytes in the sample, areas of particular research interest in this field are those that promise to better leverage the comprehensiveness of the data. One such research area of high value is the continued development of software for untargeted data analysis, i.e. analysis without preexisting spectral libraries, such as with DIA-Umpire and Group-DIA. At the same time another somewhat opposing research area that is of utmost importance is the generation of more comprehensive spectral libraries as resources for the community. Workflow approaches, i.e. either building libraries on the fly (directly from the investigated sample set) or using already existing comprehensive spectral libraries from past experiments, can typically be tailored specifically to the type of project, and the most promising approach can then be chosen.

5. Five-year view

Many novel advances both in instrument technology as well as in the high-throughput data processing have made the success of uDIA experimental design possible and the field has widely adopted this new technology. In the next 5 years, we anticipate this growth in adoption of DIA workflows will accelerate. As MS instruments become faster, more accurate, and more sensitive, we expect uDIA will be capable of monitoring very large numbers of proteins simultaneously (>10,000 in single acquisitions) across larger dynamic ranges of protein copy numbers. Such unparalleled comprehensive proteome profiling will make uDIA assays the method of choice for MS-based protein quantification. The implementation of SRM assays into clinical workflows that has already started [4] will ease the adoption of more highly multiplexed uDIA measurements in clinical analysis. Additional global, inter-laboratory studies investigating reproducibility [83], DIA data processing tools [46] will also contribute to expedite the acceptance of uDIA technologies.

One area that is expected to experience heightened exploration using uDIA technology are patient-derived samples, such as fibroblasts, primary tumor cell lines or in cells differentiated from pluripotent stem cells (patient-derived iPSC). Ultimately, however, given regulatory approval, quantitative DIA assays might be well suited for diagnostic tests and patient-specific prognosis. uDIA specifically allows for monitoring biomarkers comprised of many protein changes to reflect very complex disease biology and heterogeneity. For example, Cantonwine et al. performed SRM assays on plasma samples from pregnant women to assess the risk of spontaneous preterm birth and to evaluate proteomic biomarkers [87], while Zhang et al. used measured differential gene expression and changes from peripheral blood of breast cancer patients and control, and used support vector machine, a learning algorithm to develop 'multi-marker panels' [88]. In general, such nontraditional biomarker panels can be easily monitored by uDIA; however, these measurements would not be easily implemented using competing clinical methods, such as ELISA.

Additional sample preparation developments coupled to novel bioinformatic methods will also drive the flexibility and widespread adoption of uDIA. For example, improvement of algorithms that generate pseudo-MS/MS spectra from uDIA will facilitate the application to studies where preexisting spectral libraries are lacking or are insufficient, such as for PTM studies or studies in new model organisms.

Acknowledgments

Funding

This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases (R24DK085610, PI: E. Verdin). Jesse G. Meyer was supported by an NIH T32 (T32G000266, PI: J. Campisi). We acknowledge support from the NIH shared instrumentation grant for the TripleTOF system at the Buck Institute (1S10 OD016281, PI: B.W. Gibson).

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Method of the year 2012. *Nature Methods*. 2013; 10(1):1–1. Editorial. [PubMed: 23547284]
2. Gillette MA, Carr SA. Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. *Nat Methods*. 2013; 10(1):28–34. [PubMed: 23269374]
3. Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat Methods*. 2012; 9(6):555–566. [PubMed: 22669653]
4. Percy AJ, Byrns S, Pennington SR, et al. Clinical translation of MS-based, quantitative plasma proteomics: status, challenges, requirements, and potential. *Expert Rev Proteomics*. 2016; 13(7):673–684. [PubMed: 27341553]
5. Parker CE, Borchers CH. Mass spectrometry based biomarker discovery, verification, and validation—quality assurance and control of protein biomarker assays. *Mol Oncol*. 2014; 8(4):840–858. [PubMed: 24713096]
6. Abbatiello SE, Schilling B, Mani DR, et al. Large-scale interlaboratory study to develop, analytically validate and apply highly multiplexed, quantitative peptide assays to measure cancer-relevant proteins in plasma. *Mol Cell Proteomics*. 2015; 14(9):2357–2374. [PubMed: 25693799]
7. Carr SA, Abbatiello SE, Ackermann BL, et al. Targeted peptide measurements in biology and medicine: best practices for mass spectrometry-based assay development using a fit-for-purpose approach. *Mol Cell Proteomics*. 2014; 13(3):907–917. [PubMed: 24443746]
8. <http://www.srmatlas.org>.
9. Whiteaker JR, Halusa GN, Hoofnagle AN, et al. Using the CPTAC assay portal to identify and implement highly characterized targeted proteomics assays. *Methods Mol Biol*. 2016; 1410:223–236. [PubMed: 26867747]
10. assays.cancer.gov
11. Paulovich AG, Whiteaker JR. Quantifying the human proteome. *Nat Biotechnol*. 2016; 34(10):1033–1034. [PubMed: 27727212]
12. Kusebauch U, Campbell DS, Deutsch EW, et al. Human SRMatlas: A resource of targeted assays to quantify the complete human proteome. *Cell*. 2016; 166(3):766–778. [PubMed: 27453469]
13. Huttenhain R, Soste M, Selevsek N, et al. Reproducible quantification of cancer-associated proteins in body fluids using targeted proteomics. *Sci Transl Med*. 2012; 4(142):142ra194.
14. Percy AJ, Chambers AG, Yang J, et al. Advances in multiplexed MRM-based protein biomarker quantitation toward clinical utility. *Biochim Biophys Acta*. 2014; 1844(5):917–926. [PubMed: 23806606]
15. Gallien S, Bourmaud A, Kim SY, et al. Technical considerations for large-scale parallel reaction monitoring analysis. *J Proteomics*. 2014; 100:147–159. [PubMed: 24200835]

16. Gallien S, Domon B. Detection and quantification of proteins in clinical samples using high resolution mass spectrometry. *Methods*. 2015; 81:15–23. [PubMed: 25843604]
17. Gallien S, Duriez E, Crone C, et al. Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer. *Mol Cell Proteomics*. 2012; 11(12):1709–1723. [PubMed: 22962056]
18. Schilling B, MacLean B, Held JM, et al. Multiplexed, scheduled, high-resolution parallel reaction monitoring on a full scan QqTOF instrument with integrated data-dependent and targeted mass spectrometric workflows. *Anal Chem*. 2015; 87(20):10222–10229. [PubMed: 26398777]
- 19••. Gillet LC, Navarro P, Tate S, et al. 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*. 2012; 11(6):O111 016717. Groundbreaking manuscript for practically implementing untargeted data-independent acquisitions into nowadays widely adopted proteomics workflows for accurate quantification.
20. Aebersold R, Bensimon A, Collins BC, et al. Applications and developments in targeted proteomics: from SRM to DIA/SWATH. *Proteomics*. 2016; 16(15–16):2065–2067. [PubMed: 27542680]
21. Faktor J, Michalova E, Bouchal P. p SRM, SWATH and HRM - targeted proteomics approaches on TripleTOF 5600+ mass spectrometer and their applications in oncology research. *Klin Onkol*. 2014; 27(Suppl 1):S110–115. [PubMed: 24945547]
22. <http://www.swathatlas.org>.
23. Sajic T, Liu Y, Aebersold R. Using data-independent, high-resolution mass spectrometry in protein biomarker research: perspectives and clinical applications. *Proteomics Clin Appl*. 2015; 9(3–4): 307–321. [PubMed: 25504613]
24. Zhang Y, Bilbao A, Bruderer T, et al. The use of variable Q1 isolation windows improves selectivity in LC-SWATH-MS acquisition. *J Proteome Res*. 2015; 14(10):4359–4371. [PubMed: 26302369]
25. Hunter, C., Collins, B., Liu, Y., et al. Multi laboratory reproducibility and performance of SWATH™ acquisition for proteomic analyses. Proceedings of the 63rd Annual ASMS Conference on Mass Spectrometry & Allied Topics; St. Louis, MO. May 31 June 4, 2015; 2015.
- 26••. Collins, BC., Hunter, C., Liu, Y., et al. Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Biorxiv Preprint in Revision*. 2016. (<http://biorxiv.org/content/early/2016/09/14/074567>) Demonstration of high reproducibility, qualitative and quantitative performance of SWATH mass spectrometric workflows in a large-scale, multilaboratory study
27. Schilling B, Gibson BW, Hunter CL. Generation of high-quality SWATH(R) acquisition data for label-free quantitative proteomics studies using tripleTOF(R) Mass Spectrometers. *Methods Mol Biol*. 2017; 1550:223–233. [PubMed: 28188533]
28. Egertson JD, MacLean B, Johnson R, et al. Multiplexed peptide analysis using data-independent acquisition and skyline. *Nat Protoc*. 2015; 10(6):887–903. [PubMed: 25996789]
29. Novy, K., Balagtas, C., Muntel, J., et al. High content protein quantification of liver tumor biopsies by hyper reaction monitoring mass spectrometry. The Molecular Medicine Tri-Conference; San Francisco, CA, USA. 2017;
30. Muntel J, Xuan Y, Berger ST, et al. Advancing urinary protein biomarker discovery by data-independent acquisition on a quadrupole-orbitrap mass spectrometer. *J Proteome Res*. 2015; 14(11):4752–4762. [PubMed: 26423119]
31. Kockmann T, Trachsel C, Panse C, et al. Targeted proteomics coming of age - SRM, PRM and DIA performance evaluated from a core facility perspective. *Proteomics*. 2016; 16(15–16):2183–2192. [PubMed: 27130639]
32. Collins BC, Gillet LC, Rosenberger G, et al. Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-3 system. *Nat Methods*. 2013; 10(12):1246–1253. [PubMed: 24162925]
33. Liu Y, Buil A, Collins BC, et al. Quantitative variability of 342 plasma proteins in a human twin population. *Mol Syst Biol*. 2015; 11(1):786. [PubMed: 25652787]

34. Selevsek N, Chang CY, Gillet LC, et al. Reproducible and consistent quantification of the *Saccharomyces cerevisiae* proteome by SWATH-mass spectrometry. *Mol Cell Proteomics*. 2015; 14(3):739–749. [PubMed: 25561506]
35. Panchaud A, Scherl A, Shaffer SA, et al. Precursor acquisition independent from ion count: how to dive deeper into the proteomics ocean. *Anal Chem*. 2009; 81(15):6481–6488. [PubMed: 19572557]
36. Carvalho PC, Han X, Xu T, et al. XDIA: improving on the label-free data-independent analysis. *Bioinformatics*. 2010; 26(6):847–848. [PubMed: 20106817]
37. Silva JC, Gorenstein MV, Li GZ, et al. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics*. 2006; 5(1):144–156. [PubMed: 16219938]
38. Bruderer R, Bernhardt OM, Gandhi T, et al. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophentreated three-dimensional liver microtissues. *Mol Cell Proteomics*. 2015; 14(5):1400–1410. [PubMed: 25724911]
39. Rost HL, Rosenberger G, Navarro P, et al. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat Biotechnol*. 2014; 32(3):219–223. [PubMed: 24727770]
40. MacLean B, Tomazela DM, Shulman N, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*. 2010; 26(7):966–968. [PubMed: 20147306]
41. Lambert JP, Ivosev G, Couzens AL, et al. Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nat Methods*. 2013; 10(12):1239–1245. [PubMed: 24162924]
42. Wang J, Tucholska M, Knight JD, et al. MSPLIT-DIA: sensitive peptide identification for data-independent acquisition. *Nat Methods*. 2015; 12(12):1106–1108. [PubMed: 26550773]
- 43•. Tsou CC, Avtonomov D, Larsen B, et al. DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat Methods*. 2015; 12(3):258–264. 257 p following 264. Presentation of a comprehensive computational workflow and open-source software for DIA data processing that has been adopted by many groups and allows for database searching of DIA data without spectral libraries. [PubMed: 25599550]
44. Tsou CC, Tsai CF, Teo GC, et al. Untargeted, spectral library-free analysis of data-independent acquisition proteomics data generated using Orbitrap mass spectrometers. *Proteomics*. 2016; 16(15–16):2257–2271. [PubMed: 27246681]
45. Li Y, Zhong CQ, Xu X, et al. Group-DIA: analyzing multiple data-independent acquisition mass spectrometry data files. *Nat Methods*. 2015; 12(12):1105–1106. [PubMed: 26436481]
46. Navarro P, Kuharev J, Gillet LC, et al. A multicenter study benchmarks software tools for label-free proteome quantification. *Nat Biotechnol*. 2016; 34(11):1130–1136. [PubMed: 27701404]
47. Ting YS, Egertson JD, Payne SH, et al. Peptide-centric proteome analysis: an alternative strategy for the analysis of tandem mass spectrometry data. *Mol Cell Proteomics*. 2015; 14(9):2301–2307. [PubMed: 26217018]
48. Teo G, Kim S, Tsou CC, et al. mapDIA: preprocessing and statistical analysis of quantitative proteomics data from data independent acquisition mass spectrometry. *J Proteomics*. 2015; 129:108–120. [PubMed: 26381204]
49. Rost HL, Liu Y, D'Agostino G, et al. TRIC: an automated alignment strategy for reproducible protein quantification in targeted proteomics. *Nat Methods*. 2016; 13(9):777–783. [PubMed: 27479329]
50. Choi M, Chang CY, Clough T, et al. MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics*. 2014; 30(17):2524–2526. [PubMed: 24794931]
51. Bilbao A, Varesio E, Luban J, et al. Processing strategies and software solutions for data-independent acquisition in mass spectrometry. *Proteomics*. 2015; 15(5–6):964–980. [PubMed: 25430050]
- 52•. Meyer JG, D'Souza AK, Sorensen DJ, et al. Quantification of lysine acetylation and succinylation stoichiometry in proteins using mass spectrometric data-independent acquisitions (SWATH). *J Am Soc Mass Spectrom*. 2016; 27(11):1758–1771. A new approach of using DIA and fragment

ion peak areas to determine PTM site occupancy in a highly accurate quantitative approach. [PubMed: 27590315]

53. Meyer, JG., Schilling, B., Gibson, BW. Computational workflows to improve DIA data analysis without reference libraries for quantitative proteomics and PTM analysis. Proceedings of the 64th Annual ASMS Conference on Mass Spectrometry & Allied Topics; San Antonio, TX. June 5–9, 2016;
54. Broudy D, Killeen T, Choi M, et al. A framework for installable external tools in Skyline. *Bioinformatics*. 2014; 30(17):2521–2523. [PubMed: 24813211]
55. Abbatiello SE, Mani DR, Schilling B, et al. Design, implementation and multisite evaluation of a system suitability protocol for the quantitative assessment of instrument performance in liquid chromatography-multiple reaction monitoring-MS (LC-MRM-MS). *Mol Cell Proteomics*. 2013; 12(9):2623–2639. [PubMed: 23689285]
56. Bereman MS. Tools for monitoring system suitability in LC MS/MS centric proteomic experiments. *Proteomics*. 2015; 15(5–6):891–902. [PubMed: 25327420]
57. Bereman MS, Beri J, Sharma V, et al. An automated pipeline to monitor system performance in liquid chromatography-tandem mass spectrometry proteomic experiments. *J Proteome Res*. 2016; 15:4763–4769. [PubMed: 27700092]
- 58•. Rosenberger G, Koh CC, Guo T, et al. A repository of assays to quantify 10,000 human proteins by SWATH-MS. *Sci Data*. 2014; 1:140031. The compendium of highly specific assays covering more than 10,000 human proteins enable their targeted analysis in SWATH-MS datasets acquired from research or clinical specimens. [PubMed: 25977788]
59. Parker SJ, Rost H, Rosenberger G, et al. Identification of a set of conserved eukaryotic internal retention time standards for data-independent acquisition mass spectrometry. *Mol Cell Proteomics*. 2015; 14(10):2800–2813. [PubMed: 26199342]
60. Zolg DP, Wilhelm M, Schnatbaum K, et al. Building ProteomeTools based on a complete synthetic human proteome. *Nat Methods*. 2017; 14(3):259–262. [PubMed: 28135259]
61. <http://www.proteometools.org>.
62. Liu Y, Huttenhain R, Surinova S, et al. Quantitative measurements of N-linked glycoproteins in human plasma by SWATH-MS. *Proteomics*. 2013; 13(8):1247–1256. [PubMed: 23322582]
- 63•. Liu Y, Chen J, Sethi A, et al. Glycoproteomic analysis of prostate cancer tissues by SWATH mass spectrometry discovers N-acyl ethanolamine acid amidase and protein tyrosine kinase 7 as signatures for tumor aggressiveness. *Mol Cell Proteomics*. 2014; 13(7):1753–1768. Identification and Quantification of potential tissue biomarkers to avoid overtreatment of non-aggressive prostate cancer types using DIA. [PubMed: 24741114]
64. Hou G, Lou X, Sun Y, et al. Biomarker discovery and verification of esophageal squamous cell carcinoma using integration of SWATH/MRM. *J Proteome Res*. 2015; 14(9):3793–3803. [PubMed: 26224564]
65. Ortea I, Rodriguez-Ariza A, Chicano-Galvez E, et al. Discovery of potential protein biomarkers of lung adenocarcinoma in bronchoalveolar lavage fluid by SWATH MS data-independent acquisition and targeted data extraction. *J Proteomics*. 2016; 138:106–114. [PubMed: 26917472]
66. Nassar AF, Williams BJ, Yaworsky DC, et al. Rapid label-free profiling of oral cancer biomarker proteins using nano-UPLC-Q-TOF ion mobility mass spectrometry. *Proteomics Clin Appl*. 2016; 10(3):280–289. [PubMed: 26684082]
- 67•. Kulkarni S, Koller A, Mani KM, et al. Identifying urinary and serum exosome biomarkers for radiation exposure using a data dependent acquisition and SWATH-MS combined workflow. *Int J Radiat Oncol Biol Phys*. 2016; 96(3):566–577. This study demonstrates the feasibility of defining biomarkers that could elucidate tissue-associated and systemic response caused by high-dose ionizing radiation. [PubMed: 27485285]
68. Nakamura K, Hirayama-Kurogi M, Ito S, et al. Large-scale multiplex absolute protein quantification of drug-metabolizing enzymes and transporters in human intestine, liver, and kidney microsomes by SWATH-MS: comparison with MRM/SRM and HR-MRM/PRM. *Proteomics*. 2016; 16(15–16):2106–2117. [PubMed: 27197958]
69. Chang RY, Etheridge N, Nouwens AS, et al. SWATH analysis of the synaptic proteome in Alzheimer's disease. *Neurochem Int*. 2015; 87:1–12. [PubMed: 25958317]

70. Silva C, Santa C, Anjo SI, et al. A reference library of peripheral blood mononuclear cells for SWATH-MS analysis. *Proteomics Clin Appl*. 2016; 10(7):760–764. [PubMed: 27188786]
71. Haverland NA, Fox HS, Ciborowski P. Quantitative proteomics by SWATH-MS reveals altered expression of nucleic acid binding and regulatory proteins in HIV-1-infected macrophages. *J Proteome Res*. 2014; 13(4):2109–2119. [PubMed: 24564501]
72. Vaswani K, Ashman K, Reed S, et al. Applying SWATH mass spectrometry to investigate human cervicovaginal fluid during the menstrual cycle. *Biol Reprod*. 2015; 93(2):39. [PubMed: 26108790]
73. Malmstrom L, Bakochi A, Svensson G, et al. Quantitative proteogenomics of human pathogens using DIA-MS. *J Proteomics*. 2015; 129:98–107. [PubMed: 26381203]
74. Loke MF, Ng CG, Vilashni Y, et al. Understanding the dimorphic lifestyles of human gastric pathogen *Helicobacter pylori* using the SWATH-based proteomics approach. *Sci Rep*. 2016; 6:26784. [PubMed: 27222005]
75. Creech AL, Taylor JE, Maier VK, et al. Building the connectivity map of epigenetics: chromatin profiling by quantitative targeted mass spectrometry. *Methods*. 2015; 72:57–64. [PubMed: 25448295]
76. Zawadzka AM, Schilling B, Held JM, et al. Variation and quantification among a target set of phosphopeptides in human plasma by multiple reaction monitoring and SWATH-MS2 data-independent acquisition. *Electrophoresis*. 2014; 35(24):3487–3497. [PubMed: 24853916]
77. Griffiths JR, Chicooree N, Connolly Y, et al. Mass spectral enhanced detection of UbIs using SWATH acquisition: MEDUSA—simultaneous quantification of SUMO and ubiquitin-derived isopeptides. *J Am Soc Mass Spectrom*. 2014; 25(5):767–777. [PubMed: 24549895]
78. Keller A, Bader SL, Kusebauch U, et al. Opening a SWATH window on posttranslational modifications: automated pursuit of modified peptides. *Mol Cell Proteomics*. 2016; 15(3):1151–1163. [PubMed: 26704149]
79. Schmidlin T, Garrigues L, Lane CS, et al. Assessment of SRM, MRM(3), and DIA for the targeted analysis of phosphorylation dynamics in non-small cell lung cancer. *Proteomics*. 2016; 16(15–16):2193–2205. [PubMed: 27219855]
80. Krautkramer KA, Reiter L, Denu JM, et al. Quantification of SAHA-dependent changes in histone modifications using data-independent acquisition mass spectrometry. *J Proteome Res*. 2015; 14(8):3252–3262. [PubMed: 26120868]
81. Sidoli S, Lin S, Xiong L, et al. Sequential window acquisition of all theoretical mass spectra (SWATH) analysis for characterization and quantification of histone post-translational modifications. *Mol Cell Proteomics*. 2015; 14(9):2420–2428. [PubMed: 25636311]
82. Shao S, Guo T, Koh CC, et al. Minimal sample requirement for highly multiplexed protein quantification in cell lines and tissues by PCT-SWATH mass spectrometry. *Proteomics*. 2015; 15(21):3711–3721. Presentation of a sample preparation technology that addresses and offers solutions to overcome problems related to sample amount available for clinical and biological proteomic research which is often limited and thus can significantly restricts clinical and translational research. [PubMed: 26287124]
83. Hunter, CL., Collins, B., Gillet, L., et al. Increasing depth of coverage in data independent acquisition with acquisition improvements and higher sample loads. *Proceedings of the 61st Annual ASMS Conference on Mass Spectrometry & Allied Topics*; Baltimore, MD. June 15–19, 2014;
84. Sharma V, Eckels J, Taylor GK, et al. Panorama: a targeted proteomics knowledge base. *J Proteome Res*. 2014; 13(9):4205–4210. [PubMed: 25102069]
85. Doerr A. Proteomics: peptides aplenty. *Nat Methods*. 2013; 10(7):609. [PubMed: 23967485]
86. Vowinkel J, Zelezniak A, Kibler A, et al. Precise label-free quantitative proteomes in highthroughput by microLC and data-independent SWATH acquisition. *Biorxiv Preprint*. 2016; doi: 10.1101/073478
87. Cantonwine DE, Zhang Z, Rosenblatt K, et al. Evaluation of proteomic biomarkers associated with circulating microparticles as an effective means to stratify the risk of spontaneous preterm birth. *Am J Obstet Gynecol*. 2016; 214(5):631 e631–631 e611. [PubMed: 26874302]

88. Zhang F, Deng Y, Drabier R. Multiple biomarker panels for early detection of breast cancer in peripheral blood. *Biomed Res Int*. 2013; 2013:781618. [PubMed: 24371830]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Key issues

- uDIA, such as SWATH, provides global, comprehensive and highly quantitative assays that are beginning to find applications in clinically relevant projects, such as biomarker discovery and validation.
- Practical application of uDIA workflows has been enabled by recent improvements in instrument speed and resolution, diverse data analysis software, and the growing public availability of comprehensive spectral libraries.
- Compared to relative quantitation using SRM or PRM, uDIA currently allows greater breadth at a similar accuracy.
- Generation of comprehensive spectral libraries for targeted analysis of uDIA data has recently been challenged by software for untargeted uDIA analysis that enables identification and quantification without preexisting spectral libraries.
- The unbiased nature of uDIA uniquely lends itself to data reanalysis in the future as new analysis tools continue to be released.
- Additional efforts are needed to implement FDA guidelines for uDIA assays, which will open up new clinical opportunities.

Figure 1.

Schematic showing the most common DIA methods and their hierarchical relationship. Each section depicts one scan within the scan cycle. SRM and PRM are both types of targeted DIA, where the instrument isolates predefined precursor masses for fragmentation. In SRM, the third quadrupole monitors a single fragment with high sensitivity and selectivity. This precursor/fragment pair is often referred to collectively as a transition. Many transitions are often measured in series throughout the experiment. In PRM, the instrument records all fragments with high mass resolution. Many cycles isolating distinct precursor masses are combined PRM methods. In contrast, with uDIA or SWATH, the instrument first measures all intact precursor ions (not pictured) followed by isolation of a much wider precursor window (between 2 and 200 m/z) for fragmentation that usually contains several peptide precursors. The instrument then measures all fragment ions within the observable m/z window. This process is repeated until MS/MS from all windows covering all measured precursor masses have been fragmented and measured.

Table 1

Overview of typical instrument settings and parameters for data-independent acquisitions: SRM, PRM, untargeted DIA (most cited methods).

MS System	Instrument type	Scan type	Range (in m/z)	Precursor selectivity	Time MS2	Segm.	Cycle time	MS2 Resolution	#Analytes (proteins)	Ref.
QTRAP 5500	QqQ	SRM	N/A	0.7 m/z	10 ms	N/A	3–3.5 s	0.7 m/z	100–200	[13]
6490 QQQ	QqQ	SRM	N/A	0.7 m/z	13–256 ms	N/A	0.51 s	0.7 m/z	142	[14]
Q-Exactive,	Q-Orbitrap	PRM	N/A	2 m/z for screening	50–100 ms	N/A	2–3 s	17,500–35,000 screening	606 peptide pairs of endogenous and SIS	[16]
Q-Exactive Plus /HF	Q-Orbitrap	PRM	N/A	1–2 m/z for quantification	250–500 ms	N/A		70,000–1,40,000 quant.		
TripleTOF 5600/6600	QqTOF	PRM	N/A	2 m/z	60 ms	N/A	3.3 s	15–20,000	466 pept.	[18]
TripleTOF 5600	QqTOF	uDIA	400–1200	5–90 m/z	45 ms	64	3.2 s	15,000	~4500	[26,27]
TripleTOF 6600	QqTOF	uDIA	400–1250	5–50 m/z	25 ms	100	2.8 s	20,000	4000–5000	
QExactive	Q-Orbitrap	uDIA	500–900	20 m/z	100 ms	20	~2 s	17,500 at 200 m/z	5000–7500	[28,29]
QExactive HF	Q-Orbitrap	uDIA	400–1000	20 m/z for 20 MS2, then 40–60 m/z for last 4 MS2	~80 ms	24	~2 s	30,000 at 200 m/z	~2500 *	[30]
Orbitrap Fusion Tribrid	Q-IT-Orbi	uDIA	500–900	20 m/z	50–65 ms	20	~1.8–2.2 sec	30,000 @ 200	1000–7000	**
Orbitrap Fusion Lumos Tribrid	Q-IT-Orbi	uDIA	500–900	10 m/z	17–25 ms	40	~1.8–2.2 sec	15,000 @ 200	1000–7000	**

uDIA: untargeted DIA;

* urine samples;

** personal communication Dr. Jarrett Egerton.

Table 2

Bioinformatic tools for DIA analysis.

Name features	Reference	Acquisition type	Extraction	Description
Skyline	[28,40]	SRM, PRM, uDIA	Targeted	Post-processing, assay development, statistical testing, protein quantification from preexisting spectral libraries
OpenSWATH, Spectronaut, SWATH 2.0	[38,39,41]	uDIA	Targeted	Protein quantification using preexisting spectral libraries
DIA-Umpire	[43,44]	uDIA	Untargeted	Pseudo-MS/MS spectra generation for identification by database search, protein quantification using IDs from pseudo-MS/MS and preexisting libraries
Group-DIA	[45]	uDIA	Untargeted	Pseudo-MS/MS spectra generation for identification by database search, benefits from many acquisitions
MSPLIT-DIA	[42]	uDIA	Untargeted	Peptide Identification using preexisting spectral libraries, spectral matching
PECAN	[47]	uDIA	Targeted	Peptide-centric analysis tool to assessing the significance of peptide presence
XDIA	[36]	uDIA	n/a	Data collection strategy and pseudo-MS/MS spectra generation for identification by database searching using data from middle-down electron-transfer dissociation
PacIFIC	[35]	uDIA	n/a	Targeted fragmentation of 10 small m/z precursor ranges covering 15 m/z repeated many times to cover the entire precursor mass range
mapDIA	[48]	SRM, PRM, uDIA	n/a	Post-quantification fragment-level filtering, significance testing
TRIC	[49]	uDIA	n/a	Automated alignment strategy for reproducible protein quantification
MSstats	[50]	SRM, PRM, uDIA	n/a	Post-quantification significance testing
NOFI ranking	[51]	SRM, PRM, uDIA	n/a	Post-quantification multivariate fragment filtering based on ion rank and chromatographic attributes
StoichiolyzeR	[52]	uDIA	Targeted	Post-identification re-quantification and computation of PTM site-level stoichiometry from peptides containing 1–2 modified sites

Table 3

Selected clinical applications using SRM, PRM, and uDIA.

Study	Sample type (matrix)	Instrument type	Scan type	Processing software	Biomarker candidates, other findings (mutations)	Comments	Ref.
Biomarkers for various cancers	Plasma urine	QTRAP 5500 TSQ Vantage	SRM SRM	Vendor software, mProphet	Cancer-associated proteins (CAP) detected in plasma (182) and urine (408)	Reference map of SRM assays for CAP	[13]
Biomarkers for NCDs	Plasma	6490 QQQ	SRM	MassHunter	142 biomarkers from undepleted plasma	Verification of protein panels, SRM assays	[14]
Clinical samples	Plasma urine	Q-Exactive Q-Exactive Plus /HF	PRM PRM	Pinpoint Skyline	606 peptide pairs of endogenous and SIS	Screening of peptides will trigger very high-resolution quantification scan	[16]
Histones, PTMs Epigenetic regulators	Cultured cells	Q-Exactive	PRM	Skyline	Detailed method description of histone modification profiling	cornerstone in the NIH Library of Integrated Network-based Cellular Signatures (LINCS)	[75]
Protein variability in human twins	Plasma	TripleTOF 5600+	uDIA	OpenSWATH	13 cis-SNPs	SNP influence the level of specific plasma proteins	[33]
Human prostate cancer (non-aggressive, aggressive, metastatic)	tissue biopsy	TripleTOF 5600	uDIA	OpenSWATH	N-acyl/ethanolamine Acid Amidase Tyrosine Kinase 7	Indicators for tumor aggressiveness, combined AUROC = 0.801	[63]
Ovarian cyst, UTI in children (abdominal pain)	Urine	QExactive HF	uDIA	Spectronaut 7.0	Cystatin-B (ovarian cyst) myeloperoxidase (UTI)	5.8-fold up ($p = 1.3e-5$) 97-fold up ($p = 1.1e-6$)	[30]
Microsomes Comparison of SRM, PRM, and SWATH	Intestine, liver, kidney	API5000, TripleTOF 5600 TripleTOF 5600	SR M PRM uDIA	Custom in-house	Similar results from all three methods		[68]

NCD = 'Group II Diseases' and encompass various disorders (e.g. endocrine), diseases (e.g. cardiovascular, respiratory) and congenital anomalies (e.g. Down syndrome).