HHS Public Access

Author manuscript

J Proteome Res. Author manuscript; available in PMC 2023 May 17.

Published in final edited form as:

J Proteome Res. 2023 May 05; 22(5): 1510–1519. doi:10.1021/acs.jproteome.3c00088.

Definitive Screening Designs to Optimize Library-Free DIA-MS Identification and Quantification of Neuropeptides

Ashley Phetsanthad,
Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

Austin V. Carr,
Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

Lauren Fields,
Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

Lingjun Li
Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States; School of Pharmacy and Lachman Institute for Pharmaceutical Development, School of Pharmacy, University of Wisconsin-Madison, Madison, Wisconsin 53705, United States

Abstract

Method optimization is crucial for successful mass spectrometry (MS) analysis. However, extensive method assessments, altering various parameters individually, are rarely performed due to practical limitations regarding time and sample quantity. To maximize sample space for optimization while maintaining reasonable instrumentation requirements, a definitive screening design (DSD) is leveraged for systematic optimization of data-independent acquisition (DIA) parameters to maximize crustacean neuropeptide identifications. While DSDs require several injections, a library-free methodology enables surrogate sample usage for comprehensive

Corresponding Author: Lingjun Li – Phone: +1-608-265-8491; lingjun.li@wisc.edu; Fax: +1-608-262-5345.

The authors declare no competing financial interest.

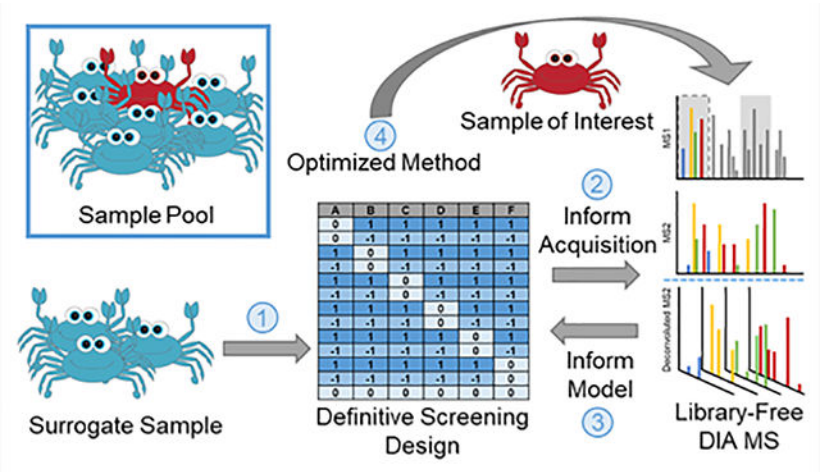
Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.3c00088

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.3c00088.
Table S1. Combined model parameter estimates identified as main or even-order effects to create a model. Figure S1. Distribution of *m/z* precursor values for neuropeptides identified through each MS acquisition; Figure S2. Actual ion injection times from A) MS1 and Bi) MS2 spectra. Bii) Percentage of MS2 spectra where the max IT was reached is shown; Figure S3. Representative A) fragmentation spectra and B) precursor profiles for the peptide HDSPHVFLRFamide are shown; Figure S4. Representative A) fragmentation spectra and B) precursor profiles for the peptide MFAPLSGLPGNLRT are shown; Figure S5. Representative A) fragmentation spectra and B) precursor profiles for the peptide acetyl-VYGPRDIANLY are shown; Figure S6. Histograms describing the peptide spectral count distributions are shown containing data acquired from each of the four methods. PSMs from peptides observed over the triplicate injections were combined; Figure S7. Evaluation of quantitative performance of DIA methods. A) Distribution of quantifiable peptides identified in different numbers of replicate injections (i) with MS2 evidence for identification or (ii) additionally supplemented through match between runs. B) Log2 transformed ratios of 1:1, 1:2, 1:5, 1:10 diluted neuropeptide samples from fDSD and KD methods. C) Distribution of integrated area CVs between replicates for quantified peptides from (i) fDSD or (ii) KD methods; Figure S8. Quantitative data from original data prior to rescuing through removal of an outlier triplicate value. A) Log2 transformed ratios of 1:1, 1:2, 1:5, 1:10 diluted neuropeptide samples from fDSD and KD methods. B) Distribution of integrated area CVs between replicates for quantified peptides from fDSD or KD methods (PDF)

optimization of MS parameters to assess biomolecules from limited samples. We identified several parameters contributing significant first- or second-order effects to method performance, and the DSD model predicted ideal values to implement. These increased reproducibility and detection capabilities enabled the identification of 461 peptides, compared to 375 and 262 peptides identified through data-dependent acquisition (DDA) and a published DIA method for crustacean neuropeptides, respectively. Herein, we demonstrate a DSD optimization workflow, using standard material, not reliant on spectral libraries for the analysis of any low abundance molecules from previous samples of limited availability. This extends the DIA method to low abundance isoforms dysregulated or only detectable in disease samples, thus improving characterization of previously inaccessible biomolecules, such as neuropeptides. Data are available via ProteomeXchange with identifier PXD038520.

Graphical Abstract.



Keywords

neuropeptides; design of experiments; label-free quantitation; DIA; data-independent acquisition; peptidomics

INTRODUCTION

Significant evidence has demonstrated that data-independent acquisition (DIA) mass spectrometry (MS) methods provide greater reproducibility than data-dependent acquisition (DDA) methods. As ion selection in DDA methods is stochastic, replicate injections can have low overlap in peptide and protein identifications, obscuring real biological changes. In contrast, DIA methods define *m/z* windows within a selected precursor range and fragment all precursors within each window simultaneously to ensure that all parent ions are fragmented.[1–4] The incorporation of DIA methods into traditional MS workflows has enabled the discovery and identification of more biomolecules in a single MS acquisition run than with traditional DDA experiments. DIA has been shown to be beneficial for proteomics and peptidomics.[5,6]

One challenge in DIA experiments is that the complex fragmentation spectra of DIA require intensive downstream processing using either library or library-free approaches. Spectral libraries make identifications using a peptide's fragmentation pattern observed in a high resolution DDA experiment to identify the same fragments from a DIA experiment.[1,7] Spectral library identifications are limited to only previously observed peptides, making them incompatible with discovery of new peptides.[8,9] The extra instrument time and large sample amount required to generate a high-quality spectral library can prohibit the analysis of precious samples that are not abundantly available. Library-free software, in contrast, transforms DIA spectra into DDA pseudospectra, compatible with existing DDA search engines.[4]

To perform DIA successfully, acquisition parameters must be carefully selected based on sample type and complexity. These decisions can directly impact the library-free software deconvolution ability. For example, if a selected isolation window width is too large, the resulting MS/MS (MS2) spectra can display overlapping fragmentation patterns from multiple peptides. When the selected isolation window width is too small, the collected points corresponding to a particular peptide decrease. The design of an effective, comprehensive DIA experiment requires careful planning of acquisition parameters for more effective use of deconvolution algorithms, including maximum ion injection time (IT), target automatic gain control (AGC), collision energy (CE), m/z range, and others.[8] Optimizing n parameters comprehensively requires $n!$ experiments when considering synergistic response between any two parameters, requiring extremely large amounts of sample and instrument operation time.[10–13] An alternative approach is the use of design of experiments (DoEs), which leverages statistical power to optimize multiple parameters in tandem using minimal experimental trials.[14,15] DoEs are widely used in engineering[16] but are less commonly applied to liquid chromatography-mass spectrometry (LC-MS) experiments, although usage of different DoEs has increased over the years.[11,13,17–27]

Combining DoE and library-free DIA methods can significantly decrease the sample and instrument time requirements to fully optimize a DIA method. To demonstrate this, we applied DoE and library-free DIA to the analysis of neuropeptides. Neuropeptides are signaling molecules crucial for neuronal communication. The subfemtomolar concentrations of neuropeptides observed *in vivo*, along with limited sample quantities gained from animal models, makes them poor candidates for a spectral-library-based DIA approach. Additionally, neuropeptide identification is not well-suited for DDA's top n selection criteria, as neuropeptides tend to coelute with several higher abundance competing matrix components.[5,28] To minimize sample requirements and maximize identifications, we applied a specific class of DoE, a definitive screening design (DSD), to decrease required experimental runs while maintaining a high level of statistical power to interpret the effects of each parameter. This model enables detection of both independently impactful parameters, classified as main effects, and parameters that complement one another to achieve a particular response, classified as two-factor interactions, in addition to predicting optimal parameter values to be used, all from a single round of data acquisition.[10,29,30] DSDs are versatile and compatible with both continuous and categorical, or discrete, parameters. Following the split-plot design of specific combinations of parameter values for experimentation, DSDs prescribe a set of strategically varied parameter combinations

to ensure enough statistical power to impute the optimal combinations while limiting the number of required experiments.[29,31]

Using a DSD, we evaluated seven parameters, *m/z* range, isolation window width, MS1 maximum IT, CE, MS2 maximum IT, MS2 target AGC, and the number of MS1 scans collected per cycle, for the increased identification of neuropeptides through library-free DIA methodology. We demonstrate a workflow for optimizing DIA MS acquisition of precious samples using a DSD and a surrogate sample of similar complexity, amenable to a wild type or control sample. The optimized method can then be applied to limited mutant or experimental samples precious in nature. We also demonstrate the ability for label-free quantitation (LFQ) of the identified neuropeptides and compare the resultant method with a previously established DIA method for neuropeptides.

EXPERIMENTAL SECTION

Sample Preparation

Several sinus gland pairs were obtained from *Callinectes sapidus*. Neuropeptide samples were prepared according to previous protocols.[32] Briefly, sinus glands were homogenized via ultrasonication probe in ice cold acidified methanol (90% methanol/9% water/1% acetic acid). The neuropeptide containing the supernatant was dried using a vacuum concentrator prior to desalting with a C18 solid phase extraction material. The subsequent samples were pooled into one vial which was used for MS acquisitions for the DSD and quantitative data. Throughout this article, peptides or neuropeptides will be considered to refer to all peptides identified through PEAKS contained in the neuropeptide database, including peptides that are truncated and/or modified by post-translational modifications (PTMs).

HPLC-MS/MS Analysis

All experiments were carried out using a Thermo Scientific Q Exactive orbitrap mass spectrometer coupled to a Waters nanoAcquity Ultra Performance LC system. Separation was performed on a 15 cm homemade column, packed with 1.7 μm particle size C18 ethylene bridged hybrid material, at a flow rate of 200 nL/min. HPLC methods were kept constant for all data acquisition. Using solutions of mobile phase A (0.1% formic acid (FA) in water) and B (0.1% FA in acetonitrile (ACN)), the gradient was ramped as follows: starting at 97% A, ramped to 90% A over 0.5 min; 85% A over 19.5 min; 83% A over 20 min; 80% over 15 min; 67% over 55 min; 25% over 10 min. Unaltered MS settings between all methods include MS1 resolution of 70,000, MS1 AGC target of 3e6, and MS2 resolution of 17,500. The default precursor charge state for all methods was set to +2. DDA acquisition was performed with a MS2 AGC target of 2e5, MS2 max IT of 200 ms, loop count of 10, isolation window width of 2, normalized CE of 30, and dynamic exclusion window of 40 s. DIA acquisitions were executed with parameter values (Table 1) as prescribed by the DSD (Figure 1A). For DIA method comparison, a previously published method[28] for crustacean neuropeptide analysis was performed with slight modifications to decrease variability between parameters. Parameters that differ from the other DIA methods used herein are as follows: MS1 AGC target of 1e6, MS1 max IT of 250 ms, MS2 AGC target of 2e5, auto MS2 max IT, 20 *m/z* isolation window width, CE of 30 V, and loop count of

10. The parameters used in the final DIA method post-DSD evaluation were 16 *m/z* isolation window width from 400 to 1034 *m/z*, CE of 25 V, MS2 max IT of 100 ms, MS2 AGC target of 1e6, MS1 max IT of 30 ms, and 4 MS1 spectra per cycle.

Neuropeptides and related peptides were identified through PEAKSxPro software.[33] PEAKS parameters were set to parent mass error tolerance of 20.0 ppm, fragment mass error tolerance of 0.02 Da, unspecific enzyme digestion, variable modifications: amidation (−0.98 Da), oxidation (M) (+15.99 Da), pyro-glu from E (−18.01 Da), pyro-glu from Q (−17.03 Da), acetylation (N-term), and max variable PTM per peptide of 3. Peptides were filtered using a −logP cutoff of 37.6, corresponding to a 5% false-discovery rate (FDR) for the DDA data. While a 1% FDR cutoff is considered standard within the proteomics community, this is not always the case in the field of peptidomics due to the unique considerations of the field. Routinely only a few hundred IDs (100–400) are expected. In this case, a mere two to five false positives from poor quality spectra will cause the FDR to quickly surpass a 1% threshold. For example, Han et al. credited their inability to identify 19 known neuropeptides to the use of a “strict” FDR threshold of 1%.[34] As a result, it is not an uncommon practice by the peptidomics community to use a larger FDR cutoff, such as 5%, and/or substitute FDR with a quality score combined with manual inspection for identification fidelity.[35–38]

Deconvolution of DIA spectra was performed using DIA-Umpire,[4] according to default parameters, and the resulting pseudo-DDA MS2 spectra were consolidated with their corresponding MS1 scans using a home-built C# console line application (https://github.com/avcarr2/DIAConverter-New) to achieve compatibility with our downstream processing workflow. While powerful, the library-free method DIA-Umpire creates pseudo-DDA spectra unsuitable for LFQ by third-party software due to the dissociation and omission of MS1 spectra information. This leads to unintended consequences which include incompatibility with identification software that uses MS1 spectra data to compute various metrics used to describe peptide spectrum match (PSM)/identification quality, such as with newer PEAKS algorithms.

To demonstrate quantitative ability of the optimized method, peptides were diluted 2-, 5-, and 10-fold in 3% ACN/0.1% FA prior to LC-MS/MS injection. The resulting peptide identifications were quantified by analyzing PEAKS database search outputs through another console application (https://github.com/avcarr2/InjectionTimeGetterApp) and FlashLFQ[39] operating with match between runs selected, a parent mass tolerance of 5 ppm, and a 5 min retention time window. A −logP cutoff of 33.1, corresponding to a 5% FDR, was used to filter the DDA data for improved quantitative accuracy. Quality and quantitative assessments of the optimized method were compared against samples of different concentrations using the previously published DIA method described above.[28] The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[40] partner repository with the data set identifier PXD038520.

Design of Definitive Screening Design

A DSD was created, and the results were interpreted using JMP Pro 15.0.0.[41] Selection of the proper factors to include when designing a DSD is crucial to minimize the introduction of aliasing and confounding factors. Therefore, the parameters and values shown in Table 1

were chosen carefully for this model after much discussion with colleagues and experienced experts in the field.[8,42,43] Per specification of DSDs, the majority of included factors are evaluated as continuous factors for effects on improving DIA MS identification: isolation window width, *m/z* range spanning from 400 *m/z*, CE, MS1 max IT, and MS2 max IT. The number of MS1 per DIA duty cycle (through loop count) was a categorical factor, as noninteger values are not possible. MS2 automatic gain control (AGC) was also evaluated as a categorical factor due to selection restraints within the Q Exactive software preventing three evenly distributed values from being selected, a key requirement of DSDs. The design involved 4 extra runs to estimate quadratic interactions and run order was randomized in two blocks to minimize bias. Peptide identification numbers were used as a response variable to assess the methods. The "maximize desirability function" in JMP was used to predict optimal parameter values to increase IDs.

RESULTS AND DISCUSSION

Identification and Reproducibility

Some proteomics optimization schemes prioritize maintaining a constant duty cycle between MS1 scans;[43,44] however, our experiments were designed with varying duty cycles, while still limiting the maximum loop time (time between MS1 scans) to collect at least 8 points across the peak for quantitation to assess the effects of specific parameters, as shown in Figure 1A. This is especially important for label-free quantitative (LFQ) analysis where MS1 peak area is used, such as in this study. Although three experiments had a loop cycle time of 3–5 s, all others were well below this with a mean of 1.54 s, 1.13 s standard deviation, and median of 1.26 s. Instead, the impact of duty cycle was evaluated indirectly based on the effects of window size, range to scan over, and how many MS1 spectra were acquired per full DIA duty cycle.

When performing library-free DIA analysis, consideration of spectral complexity is particularly important for successful deconvolution, thus the parameters that directly affect spectral complexity were prioritized over maintaining a constant duty cycle time. The number of neuropeptide identifications (IDs) using the DSD-prescribed DIA methods ranged from 182 to 434 neuropeptides. The over 2-fold range demonstrates the dependence of identifications on DIA method parameters. Through qualitative analysis, we see that higher number of IDs were largely associated with overall lower duty cycles; however, this was not always the case.

While each DDA replicate produced comparable quantities of IDs to the midperforming DIA methods, poor reproducibility of neuropeptide populations was observed, a hallmark of DDA (Figure 1B). Only 155 peptides were observed across all triplicate injections, while 331 peptides were observed in at least 2 injections. DIA enabled the observation of 153 common peptides across spectra collected from 21 different DIA methods and 331 peptides across 9 different DIA methods, indicated by the asterisks in Figure 1C, demonstrating decreased variability between DIA experiments. This is additionally beneficial in the case of LFQ, where irreproducibility greatly decreases the number of quantifiable peptides across conditions. While DIA already generally outperforms DDA in terms of number of

peptide identifications, performance can be further improved by evaluating the responses and optimizing acquisition methods through the DSD.

DSD Model Creation and Validation

Combined model parameter estimates created during fit definitive screening, shown in SI Table S1, determined the factors to consider when designing models through main and even-order effect estimates. The response residuals are shown in Figure 2A. These effects were considered when building a generalized regression model, where model selection was performed through minimizing the corrected Akaike information criterion (AICc), an estimator of model quality, to 243.4 through forward selection and the assumption of normal data distribution (Figure 2B). The resulting model chosen includes several main and even-order effects shown in Figure 2C. Main effects were thus limited to those including CE, window width, MS2 max IT, and MS2 AGC. Significant even-order effects included window width and MS2 AGC, CE and MS2 max IT, and last, *m/z* range as a quadratic effect. No other quadratic effects were observed as statistically significant through the model evaluation, although in principle they are likely to be present. It is important to note, while other effects on peptide IDs may be present, DSDs can only observe those with large effects. This limitation means that variables may be considered to have isolated effects, or a linear fit associated in exchange for decreasing the number of experiments required to be performed.

We also want to bring attention to blocking being removed as a main effect, indicating any drift or bias associated with running the mass spectrometer for days, did not contribute to the variation in peptide IDs between methods.

The selected model was then statistically validated and shown to have a R-squared value of 0.96 and root-mean-square error of 16.6 (Figure 3A) and considered to be statistically significant in describing the actual observed data (Figure 3B). Externally studentized residuals, shown in Figure 3C, have an even, normal distribution, indicating no left or right skewing of the model. Residuals are within the 95% Bonferroni simultaneous confidence intervals and individual limits demonstrating good model fit. Table 2 describes the parameters used to build the final model. The probability predictor feature was then used to maximize desirability of the model response, in the form of neuropeptide IDs (Figure 3D). It is thus predicted that acquiring MS2 spectra from collecting ions with an AGC target of 1e6 and max MS2 of 100 ms, where precursor ions are fragmented with a static CE of 25 V across a window width of 16 *m/z* over a range of 400 *m/z* to 1034 *m/z*, should yield a high number of neuropeptide IDs.

DSD Results for Optimal Method Creation

The resultant model identifies the presence of main, secondary, and quadratic effects. The *m/z* range was identified as both a main effect and a quadratic effect, meaning there is an optimal value at a local maximum, in this case, spanning 634 *m/z* values from 400 *m/z*. This is consistent with the principle that a larger range will encapsulate more peptides. However, due to a nonconstant distribution where fewer precursor ions are identified at larger *m/z* values, the range needs to be curtailed at the upper end.[8,44] Precursor *m/z* value distributions are shown in Figure S1. Overall, DIA methods are less biased toward higher *m/z* values

that tend to be lower in intensity with decreased ionization efficiency. Probing samples across a wide *m/z* range with DIA methods allows for the characterization of lesser observed neuropeptides, enabling the expansion of gathered information able to be used in the future.

Window width can directly impact identification capabilities in many ways, but most significantly through increasing fragmentation spectra complexity observed with larger window widths, therefore complicating data deconvolution. Here, we find that fragmenting smaller windows (16 *m/z*) of ions lead to increased neuropeptide identification. While carrying out a similarly designed DSD with smaller isolation window widths, we found that 4 *m/z* windows lead to drastically fewer IDs (data not shown). While it has been observed that narrow windows are optimal for the analysis of several analytes,[8,45,46] the associated increase in duty cycle negatively impacted identifications in crustacean neuropeptide samples. Though wider windows, such as 26 and 36 *m/z*, lead to even shorter duty cycles, the increased quantity of peptide precursors available for fragmentation results in more complex spectra and deconvolution is limited by current library-free DIA software.

MS2 max IT was identified as a main effect, which was logical as it directly impacts the duty cycle. When MS2 max IT is set too long, windows with sparse ions occupy the full allotment of time instead of moving to the next isolation window. Actual IT can be seen in Figure S2, demonstrating the necessity of limiting max IT to 100 ms, minimizing extraneous time spent collecting uninformative spectra where the low signal present would either never reach the selected AGC target, or no large benefit is gained. While a larger proportion of acquired MS2 scans reached its AGC target, there was no significant increase in associated identifications ($p = 0.83$).

Additionally, MS2 AGC target was identified as a main effect with a desired value of 1e6, likely to produce higher quality fragmentation spectra where fragment ion intensity is higher. AGC target and max IT work similarly to limit the duty cycle of MS2 spectra, although this even-order effect on peptide IDs was not large enough to be discerned through statistical analysis. While this 1e6 can be considered a large target for collecting MS2 spectra, the time required is limited by a lower max IT. Optimal MS2 AGC is also dependent on *m/z* isolation window width, identified together to produce a secondary effect. Window size impacts the range of ions measured when estimating AGC, with larger windows enabling more ions to be sampled from the same precursor packet. Increased sampling is not always desirable, however, as determined through the predictor tool that indicated the smallest window of 16 *m/z* and largest AGC target of 1e6 be chosen. The presence of several large ion clouds from different ion populations in an Orbitrap leads to coalescence and error in the resultant calculated *m/z*. Here, peptide identification benefits through the sampling of more ions to increase the intensity of individual ion populations, rather than vast ranges of ion populations.

Normalized collision energy (NCE) is ubiquitously used in the field of proteomics as CE required for adequate fragmentation efficiency is dependent on precursor ion *m/z*. In traditional DDA methods, NCE is calculated in real time to fragment a selected precursor; however, with DIA fragmentation, several precursors of various *m/z* values are cofragmented. As such, a fixed CE must be chosen to fragment all ions in the isolation

window, regardless of *m/z* value. Through the DSD, 25 V was identified as optimal over the wide range of *m/z* ion values in the neuropeptide sample analyzed. While decreasing CE positively impacts identifications, there is likely to be a lower limit with a steep drop off in IDs outside of our sampling range, when there is not enough energy to cause any fragmentation. Considerations need to be made when interpreting these results. Many of these parameters likely reach a limit, demonstrating the dangers in exceeding the statistical power with a given sample space.

The other detected secondary effect was between CE and MS2 max IT. As with many secondary effects, their correlation may not be clear initially and there may be confounding factors involved. A low CE yields inefficient parent ion fragmentation, while a low MS2 max IT means less fragment ions are collected before detection. Combining a low CE and a low MS2 max IT implies that unfragmented parent ions will have a higher signal-to-noise ratio (SNR) than in a traditional DDA experiment. The higher SNR may assist the correlation-based DIA data deconvolution methods that identify correlated parent and fragment ions through unfragmented precursors observed in MS2 spectra.

Two factors evaluated in this design were not determined to significantly affect neuropeptide identification, both related to the acquisition of MS1 spectra. The collection of both 3 and 4 MS1 spectra per cycle provided sufficient precursor information for deconvolution of fragment ion features. It is likely that effect would be seen if a larger range was sampled such as collecting 1 compared to 6 spectra. However, the impact of 3 versus 4 MS1 spectra was not large enough to discern. Similarly, MS1 max IT was also determined to not be significant in affected observed peptide IDs, further supporting the low impact of MS1 scans on overall duty cycle. Again, this observation is limited to this study, where the tested values ranged from 10 to 30 ms. Larger sampling space may lead to different conclusions, further emphasizing the importance of thoughtful experimental design. As neither of these parameters had significant effects indicated by the DSD, and thus the two factors were not considered in the model, we rely on our prior knowledge of mass spectrometry to decide which values to use for each parameter. Moving forward with developing an optimized DIA method for crustacean neuropeptides, 4 MS1 spectra will be collected per duty cycle to increase collected points across the peak for quantitation. Additionally, MS1 IT will have the longer 30 ms limit to increase signal intensity during points where less peptides elute. This will also improve the ability for low abundance precursor ion detection for spectral deconvolution.

DIA Method Assessment

The pooled neuropeptide samples were reinjected into the mass spectrometer using four different data acquisition methods. The optimal method informed by the DSD (fDSD) was assessed against the DSD method providing the highest number of IDs (oDSD), as well as the DDA method, and a DIA method previously optimized without leveraging a DOE for crustacean neuropeptides (KD).[28] As expected, all DIA methods outperformed the DDA method in terms of both peptide identification and reproducibility (Figure 4). The highest percentage of peptides identified in all triplicate injections was achieved using the KD method; however, the absolute number of peptide identification is relatively lower.

This indicates that the n = 3 neuropeptides are likely to be higher abundant or appear in less complex spectra, leading to higher reliability for fragmentation and detection by software. Figure 5A supports this, with KD showing a larger distribution of neuropeptides with a higher area under the curve (AUC) than the other DIA methods. Conversely, fDSD and oDSD methods identified more neuropeptides with lower AUCs, suggesting some improvement in sensitivity by these methods. fDSD and oDSD methods generated many reproducible IDs, as well as many neuropeptides unique to a fraction of the replicates. Overall, these DSD methods lead to a larger variety of peptide identifications.

While DSD usage led to the creation of a DIA method for crustacean neuropeptides capable of outperforming different MS methods, it was comparable to the empirically highest performing DIA method, oDSD. This highlights an important aspect of implementing applied statistics such as DOEs; they are not infallible and not to be implemented without interpretation using expertise in the particular field of application. Empirically determined methods can be used for subsequent analysis; however, one of the prescribed methods cannot be guaranteed to outperform the model predicted method in every DSD implementation. As the model accounts for variability between runs, we will consider the predicted DIA method, fDIA, as the finalized optimal method.

We would like to note that several parameters differed in these methods, limiting the ability for direct comparison; however, we will comment on observed trends. All methods have similar mass accuracy, with KD and DDA methods having mass error distributions closer to 0 ppm (Figure 5B). This can again be attributed to the bias in these methods to identify neuropeptides with higher peak intensities and AUC values. KD also has a lower AGC target, decreasing trapped ions per scan and leading to higher mass accuracy. The DDA method skewed toward fragmenting precursor ions in the low *m/z* range (Figure 5C), unsurprising as higher *m/z* ranges more commonly contain lower charge state and lower signal intensity ions. A wider distribution of precursor *m/z* values are seen for the DIA methods, enabled by the unbiased fragmentation selection scheme. A visible difference in fragmentation patterns of representative neuropeptide precursors can be seen in Figures S3–S5, owing to the different collision energies applied.

MS1 injection times for the four methods fall around 5 to 30 ms, although different upper quartile trends are observed in Figure 5D. Approximately 50% or more of all MS1 scans from fDSD and oDSD reach the max IT, unlike DDA and KD methods, likely because of the 3× larger AGC target in these DIA methods compared to KD. Additionally, the DDA method included trapping lower *m/z* ions, which commonly comprise +1 ions of high signal intensity, reaching the target AGC more quickly. The frequency of MS1 scans directly affects the precursor profile for each neuropeptide, leading to differences in precursor based MS1 quantitative accuracy. Precursor profiles of representative neuropeptides can be found in Figures S3–5. MS2 IT distributions are similar for all methods, with the exception of method KD, which used an automatic max IT setting (Figure 5E). The use of automatic max IT was not considered during our design; however, future experiments may benefit from its inclusion as a level for a categorical parameter. Most scans reached their max IT, showing the importance of balancing trap time with AGC target. Spectra from all four methods were acquired with similar overall MS1 and MS2 injection time distributions, however, different

loop times between MS1 spectra acquisitions are shown in Figure 5F. Most notably, the DDA method varied widely from all DIA methods with the time between MS1 scans. This is attributed to the dynamic exclusion setting; when no candidate for fragmentation is found, MS1 spectra are acquired. This may also contribute to the neuropeptide spectral count distribution observed (Figure S6), as the majority of peptides identified through DDA were identified through a single PSM.

Quantitation Comparison

The presence of a neuropeptide can provide valuable information, although quantitative information is required for correlating biomarkers to diseases and measuring changes due to certain conditions. The quantitative capabilities of the optimized DIA method were assessed using solvent dilution and compared to the previously published DIA method for a metric of comparison. We note that assessing quantitative accuracy through solvent dilutions and not matrix-matched dilutions can produce aberrant abundance ratios. However, herein, we are comparing relative quantitative performance between two methods using the same samples that are affected similarly using solvent dilution. Future absolute quantitative analysis of any biological samples should follow more customary dilution approaches.

LFQ MS analysis does not require a large abundance of samples, nor does it involve multistep labeling procedures where loss of valuable samples can occur, although the method suffers from high variation and missing values between replicate injections. The higher number of unique peptides per injection (Figure 4) leads to a higher number of total identified peptides which is useful for quantitation where software can perform match between runs to integrate peak areas in several chromatograms, given that a PSM was observed in at least one run.[39] We can also see a larger percentage of peptides where only one or two PSMs were detected per peptide using fDSD over KD (Figure S6), indicating that a greater number of scans contributed to increased peptide coverage. An increase in quantified peptides and the number of replicates where integrated area values were obtained can be seen in Figure S7A when match between runs was applied, decreasing missing values across triplicate injections or conditions.

To minimize quantitative inaccuracy observed through LFQ and improve quantitative accuracy and precision of LFQ of limited samples, peptides with a CV above 50%, and observed in all triplicate measurements, were rescued by omitting the peak area contributed by the highest deviation replicate from the median (Figure S7B). While this did not significantly impact the original Log2 ratio distribution (Figure S8A), a shift can be seen in the peptide CVs mean to lower values of 25% and 27% (Figure S7C) from the original 33 and 35% (Figure S8B) for the diluted samples using fDSD and KD DIA methods, respectively. Standard deviations of CV similarly improved, narrowing from 30% to 22% and 37% to 24%, respectively. This indicates increased accuracy in quantitation of individual peptides, compared to the overall sample distribution due to the removal of outlying replicate information. When adjusting for poor quantitative performance by removing 1:10 fold concentration information, an average error of 10% from the theoretical ratio is seen compared to 21% for KD. Overall, KD underperforms, only able to quantify roughly half

the number of neuropeptides as fDSD, having higher mean density distributions of CVs and more outlier CVs.

CONCLUSIONS

We demonstrate the utility of DSDs to create optimized DIA methods for the improved detection and quantitation of neuropeptides. Combining higher identification rates and match-between-runs assisted quantitation improves the number of quantified peptides across sample conditions. The increased total peptide identifications can enable more stringent quality filtering while still leaving many quantifiable PSMs for accurate quantitation. For the samples analyzed in this study, we found that the performance of the optimized method decreases as the fold-change difference increases. Therefore, neuropeptides with large observed changes should be regarded as a qualitative change rather than a quantitative change. Further gains in LFQ accuracy could be achieved by using intrasample normalization or an internal standard. Additional DOE could be performed on the parameters of DIA data processing workflows; such an approach could yield significant improvements in the number of identifications.

While significant sample amounts are required to perform a DSD, surrogate samples such as healthy or control samples can be used. The fDSD DIA-MS method can then be applied to precious or limited samples without requiring a spectral library that may not capture disease/experimental specific peptide isoforms and PTMs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was funded in part by the NIH (R01DK071801) and the NSF (CHE- 2108223). The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531) and Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. A.P. and L.F. were both supported in part by the NIH Chemistry–Biology Interface Training Grant (T32 GM008505). A.P. acknowledges a predoctoral fellowship supported by the National Institutes of Health, under Ruth L. Kirschstein National Research Service Award from the National Institutes of Health-General Medical Sciences (1F31GM143916-01A1). L.L. acknowledges funding support from NIH grants RF1AG052324, R01AG078794, and a Vilas Distinguished Achievement Professorship and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy. We acknowledge the University of Wisconsin-Madison Department of Statistics for statistical consultation during early stage experimental design.

Data Availability Statement

Data are available via ProteomeXchange with identifier PXD038520.

REFERENCES

(1). Hu A; Noble WS; Wolf-Yadlin A Technical advances in proteomics: new developments in data-independent acquisition. F1000Res. 2016, 5.

(2). Demichev V; Messner CB; Vernardis SI; Lilley KS; Ralser M DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. Nat. Methods 2020, 17 (1), 41–44. [PubMed: 31768060]

(3). Saidi M; Kamali S; Beaudry F Neuropeptidomics: Comparison of parallel reaction monitoring and data-independent acquisition for the analysis of neuropeptides using high-resolution mass spectrometry. Biomed Chromatogr 2019, 33 (7), e4523. [PubMed: 30821847]

(4). Tsou CC; Avtonomov D; Larsen B; Tucholska M; Choi H; Gingras AC; Nesvizhskii AI DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. Nat. Methods 2015, 12 (3), 258–264. 7 p following 264. [PubMed: 25599550]

(5). Bruderer R; Bernhardt OM; Gandhi T; Xuan Y; Sondermann J; Schmidt M; Gomez-Varela D; Reiter L Optimization of Experimental Parameters in Data-Independent Mass Spectrometry Significantly Increases Depth and Reproducibility of Results. Mol. Cell Proteomics 2017, 16 (12), 2296–2309. [PubMed: 29070702]

(6). Li KW; Gonzalez-Lozano MA; Koopmans F; Smit AB Recent Developments in Data Independent Acquisition (DIA) Mass Spectrometry: Application of Quantitative Analysis of the Brain Proteome. Front Mol. Neurosci 2020, 13, 564446. [PubMed: 33424549]

(7). Barkovits K; Pacharra S; Pfeiffer K; Steinbach S; Eisenacher M; Marcus K; Uszkoreit J Reproducibility, Specificity and Accuracy of Relative Quantification Using Spectral Library-based Data-independent Acquisition. Mol. Cell Proteomics 2020, 19 (1), 181–197. [PubMed: 31699904]

(8). Pino LK; Just SC; MacCoss MJ; Searle BC Acquiring and Analyzing Data Independent Acquisition Proteomics Experiments without Spectrum Libraries. Mol. Cell Proteomics 2020, 19 (7), 1088–1103. [PubMed: 32312845]

(9). Tsou CC; Tsai CF; Teo GC; Chen YJ; Nesvizhskii AI Untargeted, spectral library-free analysis of data-independent acquisition proteomics data generated using Orbitrap mass spectrometers. Proteomics 2016, 16 (15–16), 2257–71. [PubMed: 27246681]

(10). Aburaya S; Aoki W; Minakuchi H; Ueda M Definitive screening design enables optimization of LC-ESI-MS/MS parameters in proteomics. Biosci Biotechnol Biochem 2017, 81 (12), 2237–2243. [PubMed: 29068257]

(11). Elmiger MP; Poetzsch M; Steuer AE; Kraemer T Parameter Optimization for Feature and Hit Generation in a General Unknown Screening Method-Proof of Concept Study Using a Design of Experiment Approach for a High Resolution Mass Spectrometry Procedure after Data Independent Acquisition. Anal. Chem. 2018, 90 (5), 3531–3536. [PubMed: 29397691]

(12). Hecht ES; Oberg AL; Muddiman DC Optimizing Mass Spectrometry Analyses: A Tailored Review on the Utility of Design of Experiments. J. Am. Soc. Mass Spectrom. 2016, 27 (5), 767–85. [PubMed: 26951559]

(13). Randall SM; Cardasis HL; Muddiman DC Factorial experimental designs elucidate significant variables affecting data acquisition on a quadrupole Orbitrap mass spectrometer. J. Am. Soc. Mass Spectrom. 2013, 24 (10), 1501–12. [PubMed: 23913023]

(14). Dawes ML; Bergum JS; Schuster AE; Aubry AF Application of a design of experiment approach in the development of a sensitive bioanalytical assay in human plasma. J. Pharm. Biomed Anal 2012, 70, 401–7. [PubMed: 22776736]

(15). Goldman JM; More HT; Yee O; Borgeson E; Remy B; Rowe J; Sadineni V Optimization of Primary Drying in Lyophilization During Early-Phase Drug Development Using a Definitive Screening Design With Formulation and Process Factors. J. Pharm. Sci. 2018, 107 (10), 2592–2600. [PubMed: 29890172]

(16). Hundie KB; Akuma DA Optimization of biodiesel production parameters from Prosopis julifera seed using definitive screening design. Heliyon 2022, 8 (2), e08965. [PubMed: 35243085]

(17). Andrews GL; Dean RA; Hawkridge AM; Muddiman DC Improving proteome coverage on a LTQ-Orbitrap using design of experiments. J. Am. Soc. Mass Spectrom. 2011, 22 (4), 773–83. [PubMed: 21472614]

(18). Fauvelle V; Mazzella N; Morin S; Moreira S; Delest B; Budzinski H Hydrophilic interaction liquid chromatography coupled with tandem mass spectrometry for acidic herbicides and metabolites analysis in fresh water. Environ. Sci. Pollut Res. Int. 2015, 22 (6), 3988–96. [PubMed: 24859693]

(19). Gionfriddo E; Naccarato A; Sindona G; Tagarelli A A reliable solid phase microextraction-gas chromatography-triple quadrupole mass spectrometry method for the assay of selenomethionine

and selenomethylselenocysteine in aqueous extracts: difference between selenized and not-enriched selenium potatoes. Anal. Chim. Acta 2012, 747, 58–66. [PubMed: 22986136]

(20). Jadhav SB; Kumar CK; Bandichhor R; Bhosale PN Development of RP UPLC-TOF/MS, stability indicating method for omeprazole and its related substances by applying two level factorial design; and identification and synthesis of non-pharmacopoeial impurities. J. Pharm. Biomed Anal 2016, 118, 370–379. [PubMed: 26600119]

(21). Pan L; Qiu Y; Chen T; Lin J; Chi Y; Su M; Zhao A; Jia W An optimized procedure for metabonomic analysis of rat liver tissue using gas chromatography/time-of-flight mass spectrometry. J. Pharm. Biomed Anal 2010, 52 (4), 589–96. [PubMed: 20185264]

(22). Sauer CS; Li L Mass Spectrometric Profiling of Neuropeptides in Response to Copper Toxicity via Isobaric Tagging. Chem. Res. Toxicol. 2021, 34 (5), 1329–1336. [PubMed: 33706502]

(23). Zheng H; Clausen MR; Dalsgaard TK; Mortensen G; Bertram HC Time-saving design of experiment protocol for optimization of LC-MS data processing in metabolomic approaches. Anal. Chem. 2013, 85 (15), 7109–16. [PubMed: 23841659]

(24). Eliasson M; Rannar S; Madsen R; Donten MA; MarsdenEdwards E; Moritz T; Shockcor JP; Johansson E; Trygg J Strategy for optimizing LC-MS data processing in metabolomics: a design of experiments approach. Anal. Chem. 2012, 84 (15), 6869–76. [PubMed: 22823568]

(25). Szarka S; Prokai-Tatrai K; Prokai L Application of screening experimental designs to assess chromatographic isotope effect upon isotope-coded derivatization for quantitative liquid chromatography-mass spectrometry. Anal. Chem. 2014, 86 (14), 7033–40. [PubMed: 24922593]

(26). Thorsteinsdottir UA; Thorsteinsdottir M Design of experiments for development and optimization of a liquid chromatography coupled to tandem mass spectrometry bioanalytical assay. J. Mass Spectrom 2021, 56 (9), e4727. [PubMed: 33860573]

(27). Bernardo F; Gonzalez-Hernandez P; Ratola N; Pino V; Alves A; Homem V Using Design of Experiments to Optimize a Screening Analytical Methodology Based on Solid-Phase Microextraction/Gas Chromatography for the Determination of Volatile Methylsiloxanes in Water. Molecules 2021, 26 (11), 3429. [PubMed: 34198808]

(28). DeLaney K; Li L Data Independent Acquisition Mass Spectrometry Method for Improved Neuropeptidomic Coverage in Crustacean Neural Tissue Extracts. Anal. Chem. 2019, 91 (8), 5150–5158. [PubMed: 30888792]

(29). Hecht ES; McCord JP; Muddiman DC Definitive Screening Design Optimization of Mass Spectrometry Parameters for Sensitive Comparison of Filter and Solid Phase Extraction Purified, INLIGHT Plasma N-Glycans. Anal. Chem. 2015, 87 (14), 7305–12. [PubMed: 26086806]

(30). Takagaki K; Ito T; Arai H; Obata Y; Takayama K; Onuki Y The Usefulness of Definitive Screening Design for a Quality by Design Approach as Demonstrated by a Pharmaceutical Study of Orally Disintegrating Tablet. Chem. Pharm. Bull. (Tokyo) 2019, 67 (10), 1144–1151. [PubMed: 31582634]

(31). Jones B; Nachtsheim CJ A Class of Three-Level Designs for Definitive Screening in the Presence of Second-Order Effects. Journal of Quality Technology 2011, 43 (1), 1–15.

(32). Phetsanthad A; Vu NQ; Li L Multi-Faceted Mass Spectrometric Investigation of Neuropeptides in Callinectes sapidus. J. Vis. Exp. 2022, (183). DOI: 10.3791/63322

(33). Zhang J; Xin L; Shan B; Chen W; Xie M; Yuen D; Zhang W; Zhang Z; Lajoie GA; Ma B PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. Mol. Cell Proteomics 2012, 11 (4), M111.010587.

(34). Han B; Fang Y; Feng M; Hu H; Qi Y; Huo X; Meng L; Wu B; Li J Quantitative Neuropeptidome Analysis Reveals Neuropeptides Are Correlated with Social Behavior Regulation of the Honeybee Workers. J. Proteome Res. 2015, 14 (10), 4382–93. [PubMed: 26310634]

(35). Podvin S; Jiang Z; Boyarko B; Rossitto LA; O'Donoghue A; Rissman RA; Hook V Dysregulation of Neuropeptide and Tau Peptide Signatures in Human Alzheimer's Disease Brain. ACS Chem. Neurosci. 2022, 13 (13), 1992–2005. [PubMed: 35758417]

(36). Wu HP; Wang XY; Hu J; Su RR; Lu W; Zheng XL Identification of neuropeptides and neuropeptide receptor genes in Phauda flammans (Walker). Sci. Rep 2022, 12 (1), 9892. [PubMed: 35701459]

(37). Teixeira CMM; Correa CN; Iwai LK; Ferro ES; Castro LM Characterization of Intracellular Peptides from Zebrafish (Danio rerio) Brain. Zebrafish 2019, 16 (3), 240–251. [PubMed: 31017557]

(38). Oliphant A; Alexander JL; Swain MT; Webster SG; Wilcockson DC Transcriptomic analysis of crustacean neuropeptide signaling during the moult cycle in the green shore crab, Carcinus maenas. BMC Genomics 2018, 19 (1), 711. [PubMed: 30257651]

(39). Millikin RJ; Solntsev SK; Shortreed MR; Smith LM Ultrafast Peptide Label-Free Quantification with FlashLFQ. J. Proteome Res. 2018, 17 (1), 386–391. [PubMed: 29083185]

(40). Perez-Riverol Y; Bai J; Bandla C; Garcia-Seisdedos D; Hewapathirana S; Kamatchinathan S; Kundu DJ; Prakash A; Frericks-Zipper A; Eisenacher M; Walzer M; Wang S; Brazma A; Vizcaino JA The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. 2022, 50 (D1), D543–D552. [PubMed: 34723319]

(41). JMP, Version 15.0.0.; SAS Institute Inc.: Cary, NC, 1989–2021.

(42). Kalli A; Smith GT; Sweredoski MJ; Hess S Evaluation and optimization of mass spectrometric settings during data-dependent acquisition mode: focus on LTQ-Orbitrap mass analyzers. J. Proteome Res. 2013, 12 (7), 3071–86. [PubMed: 23642296]

(43). Serrano-Blesa E; Porter A; Lendrem DW; Pitzalis C; Barton A; Treumann A; Isaacs JD Robust optimization of SWATH-MS workflow for human blood serum proteome analysis using a quality by design approach. Clin Proteomics 2021, 18 (1), 20. [PubMed: 34384350]

(44). Kawashima Y; Watanabe E; Umeyama T; Nakajima D; Hattori M; Honda K; Ohara O Optimization of Data-Independent Acquisition Mass Spectrometry for Deep and Highly Sensitive Proteomic Analysis. Int. J. Mol. Sci. 2019, 20 (23), 5932. [PubMed: 31779068]

(45). Heaven MR; Funk AJ; Cobbs AL; Haffey WD; Norris JL; McCullumsmith RE; Greis KD Systematic evaluation of data-independent acquisition for sensitive and reproducible proteomics-a prototype design for a single injection assay. J. Mass Spectrom 2016, 51 (1), 1–11. [PubMed: 26757066]

(46). Doellinger J; Blumenscheit C; Schneider A; Lasch P Isolation Window Optimization of Data-Independent Acquisition Using Predicted Libraries for Deep and Accurate Proteome Profiling. Anal. Chem. 2020, 92 (18), 12185–12192. [PubMed: 32840101]

A)

Block	Range from 400 m/z	Isolation Window Width (m/z)	MS1 max IT (ms)	Collision Energy (V)	MS2 max IT (ms)	MS2 AGC Target	MS1 per Cycle	IDs	Estimated Duty Cycle (s)	Loop Time (s)
1	600	16	10	25	100	5E+05	3	434	3.78	1.26
1	800	16	30	25	200	1E+06	3	426	10.09	3.36
1	800	36	30	25	100	5E+05	4	385	2.34	0.59
2	400	16	30	25	100	5E+05	4	370	2.62	0.66
2	800	36	10	25	100	1E+06	3	366	2.25	0.75
2	800	26	30	25	300	1E+06	4	357	9.35	2.34
1	400	16	10	35	300	1E+06	3	345	7.53	2.51
2	400	16	10	25	300	1E+06	4	338	7.54	1.89
1	600	26	20	30	200	1E+06	4	322	4.70	1.17
2	400	36	30	25	300	5E+05	3	320	3.42	1.14
2	600	26	20	30	200	5E+05	3	318	4.68	1.56
1	400	36	20	25	300	5E+05	3	316	3.39	1.13
2	400	16	30	30	100	1E+06	3	313	2.59	0.86
1	800	16	10	25	300	5E+05	4	305	15.04	3.76
1	800	16	20	35	100	1E+06	4	296	5.08	1.27
2	800	36	10	30	300	5E+05	4	292	6.71	1.68
1	600	36	30	35	300	1E+06	4	289	5.12	1.28
1	400	36	10	25	100	1E+06	4	269	1.15	0.29
2	800	16	10	35	100	1E+06	4	258	5.04	1.26
2	800	36	30	35	100	5E+05	3	239	2.31	0.77
2	800	36	10	35	300	1E+06	3	237	6.70	2.23
1	800	16	30	35	300	5E+05	3	234	15.09	5.03
1	400	36	10	35	200	5E+05	4	233	2.26	0.57
2	400	16	30	35	300	5E+05	4	201	7.62	1.91
2	400	26	10	35	100	5E+05	3	190	1.57	0.52
1	400	36	30	35	100	1E+06	3	182	1.20	0.40

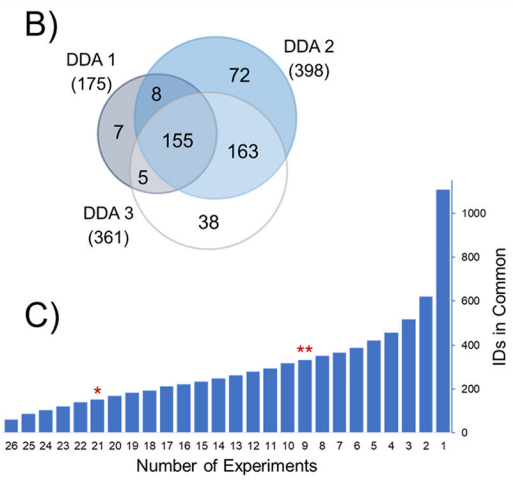
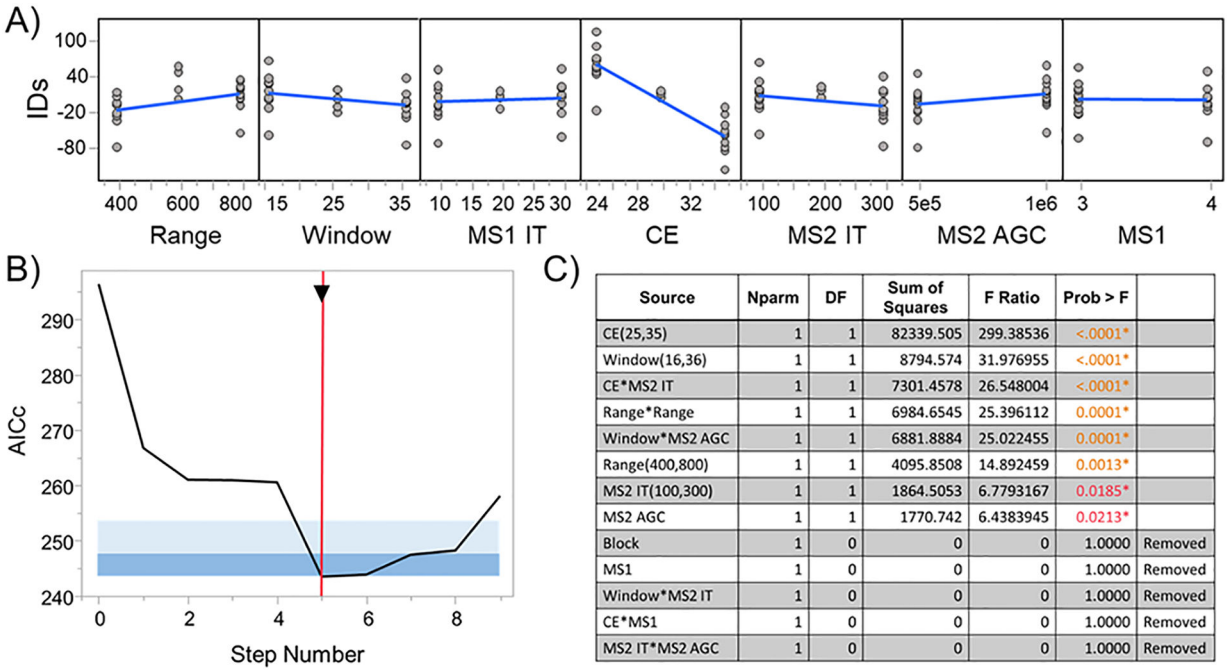


Figure 1.

A) Peptide identifications resulting from 26 DIA acquisitions outlined by the DSD. B) Overlapping peptide identifications from triplicate DDA acquisitions are shown. C) Bar chart describing the number of overlapping peptides in different numbers of DIA acquisitions. Red asterisk indicates where comparable peptide overlaps in DIA data to two (**) or three (*) of the triplicate DDA measurements is observed.



C)

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F	
CE(25,35)	1	1	82339.505	299.38536	<.0001*	
Window(16,36)	1	1	8794.574	31.976955	<.0001*	
CE*MS2 IT	1	1	7301.4578	26.548004	<.0001*	
Range*Range	1	1	6984.6545	25.396112	0.0001*	
Window*MS2 AGC	1	1	6881.8884	25.022455	0.0001*	
Range(400,800)	1	1	4095.8508	14.892459	0.0013*	
MS2 IT(100,300)	1	1	1864.5053	6.7793167	0.0185*	
MS2 AGC	1	1	1770.742	6.4383945	0.0213*	
Block	1	0	0	0	1.0000	Removed
MS1	1	0	0	0	1.0000	Removed
Window*MS2 IT	1	0	0	0	1.0000	Removed
CE*MS1	1	0	0	0	1.0000	Removed
MS2 IT*MS2 AGC	1	0	0	0	1.0000	Removed

Figure 2.

Fit definitive screening to identify parameter effects. A) Main effects residual plots of the tested parameters are shown. B) Solution path for the generalized regression models were evaluated using AICc where C) the minimum AICc value resulted from this table of main and even-order effects.

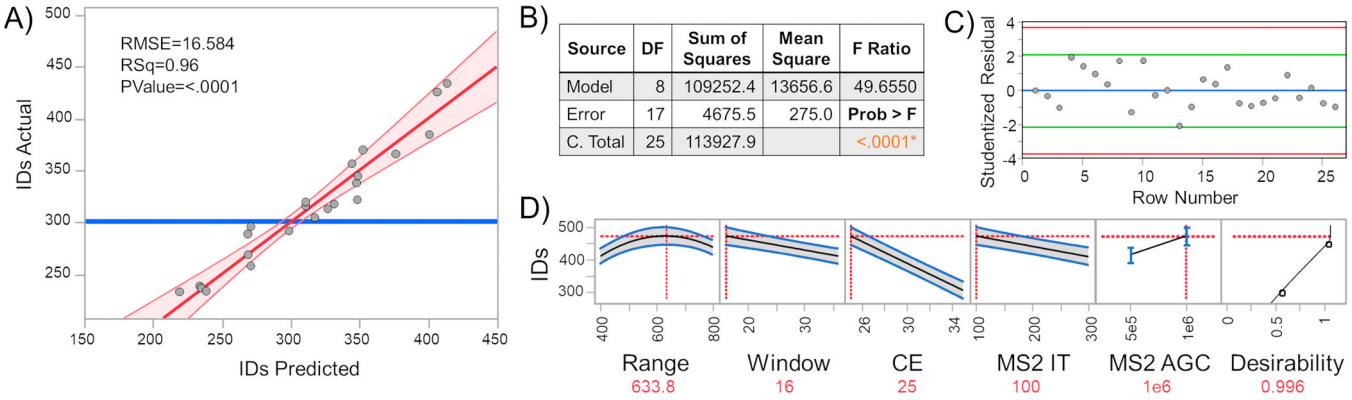


Figure 3.

A) Model validation was performed by comparing the predicted IDs to the observed IDs. B) Analysis of variance was calculated to determine the model as significant. C) Externally studentized residuals are shown with 95% simultaneous limits calculated through the Bonferroni method (red lines) and individual limits (green lines). D) Optimal parameter values were predicted through maximizing the desirability of the model, with confidence limits shown.

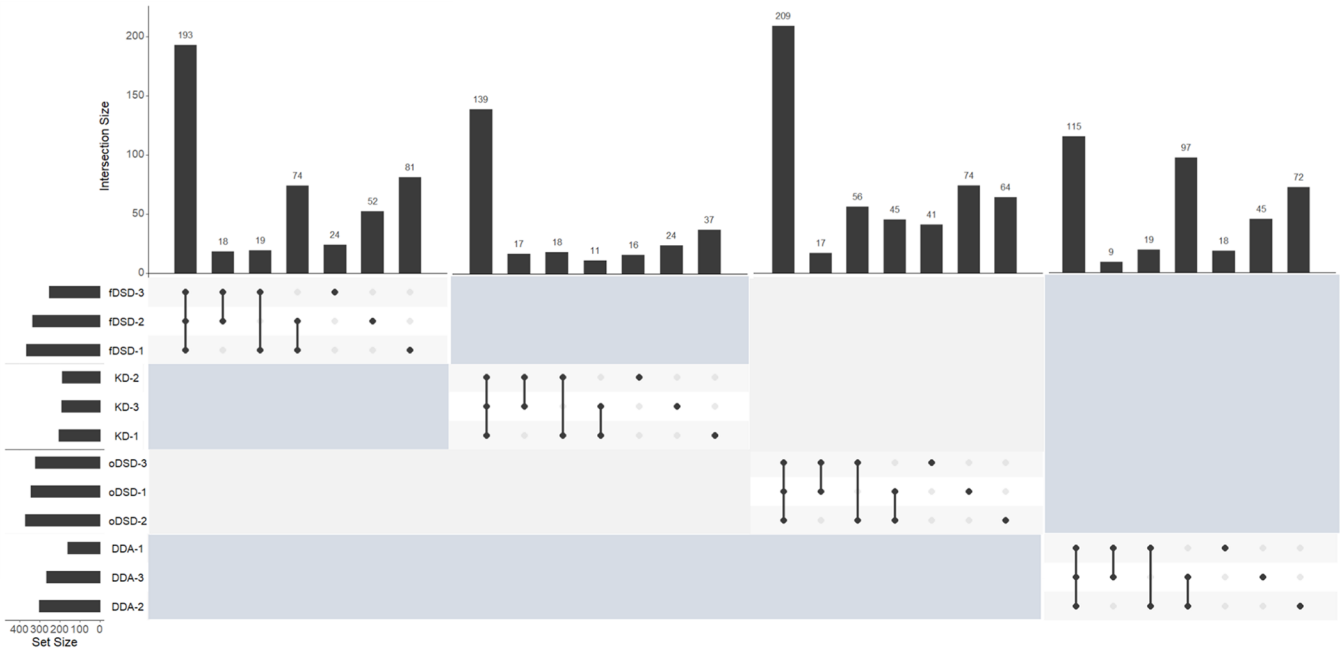


Figure 4.

Reproducibility of four different MS acquisition methods (fDSD, KD, oDSD, DDA) is visualized via upset plots demonstrating the intersection of peptide identifications from triplicate injections.

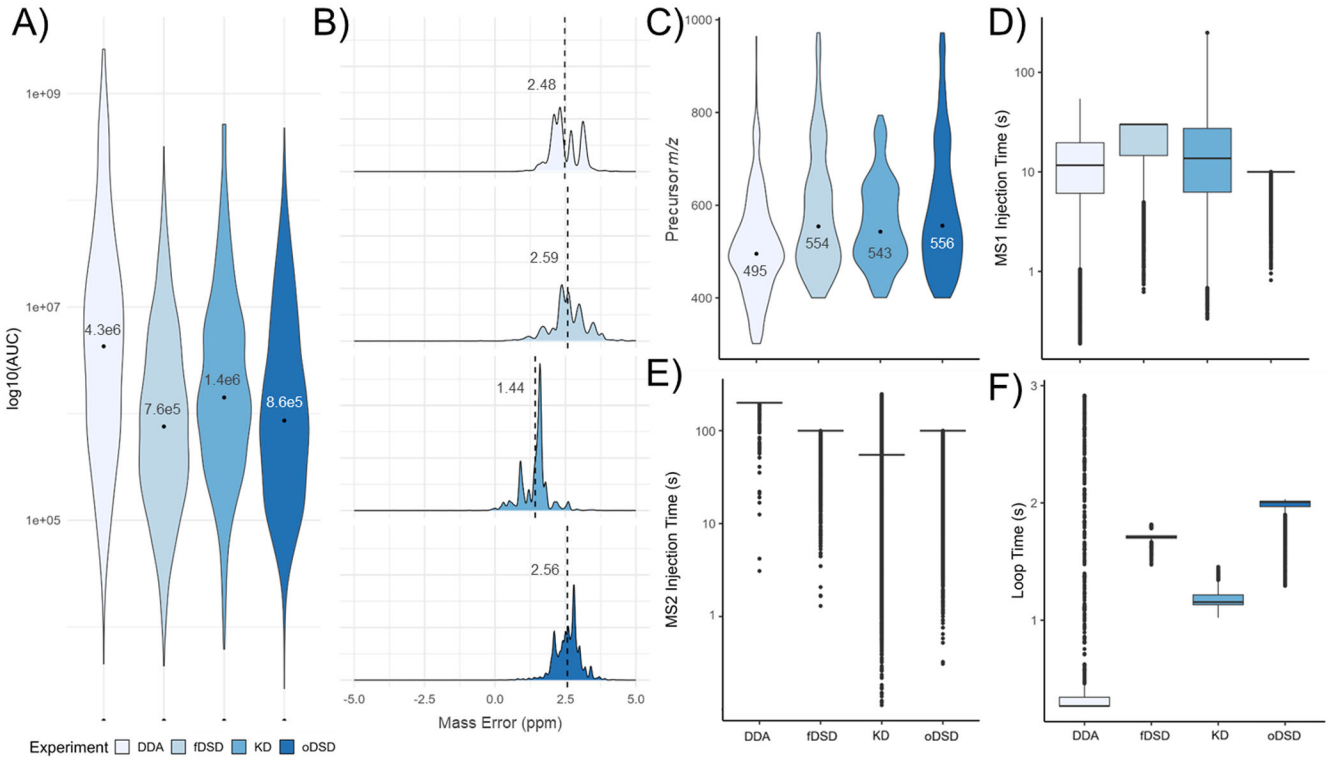


Figure 5.

Comparison of figures of merit between methods and detected neuropeptides. A) Log10 transformed neuropeptide AUC distributions. B) Distributions of identification mass error. C) Precursor *m/z* distributions. D) MS1 and E) MS2 injection times. F) Loop time between MS1 scans.

Table 1.

Parameters Selected as Factors to Evaluate for Optimization[a]

DSD Value	−1	0	1
Continuous Factors			
m/z Range from 400 *m/z*	400	600	800
Isolation Window Width *(m/z)*	16	26	36
MS1 max IT (ms)	10	20	30
MS2 max IT (ms)	100	200	300
Collision Energy (V)	25	30	35
Categorical Factors			
MS2 AGC Target	5e5		1e6
MS1 per Cycle	3		4

[a]Values are used as factor levels to test the effects on response.

Table 2.

Parameter Estimates for Factors Used in Final Model

Term	Estimate	Std Error	t Ratio	Prob>ltl
Intercept	340.04	8.58	39.62	<.0001*
Range (400,800)	13.69	3.55	3.86	0.0013*
Window (16,36)	−10.15	3.55	−2.86	0.0108*
CE (25,35)	−61.40	3.55	−17.30	<.0001*
MS2 IT (100,300)	−9.24	3.55	−2.60	0.0185*
MS2 AGC [5e5]	−8.37	3.30	−2.54	0.0213*
Window*MS2 AGC [5e5]	20.54	4.11	5.00	0.0001*
CE*MS2 IT	21.96	4.26	5.15	<.0001*
Range* Range	−47.60	9.44	−5.04	0.0001*