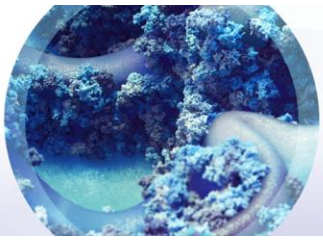


Unlocking deeper DIA performance with empirical spectral libraries generated using ZT Scan DIA 3.0 on the ZenoTOF 8600 system

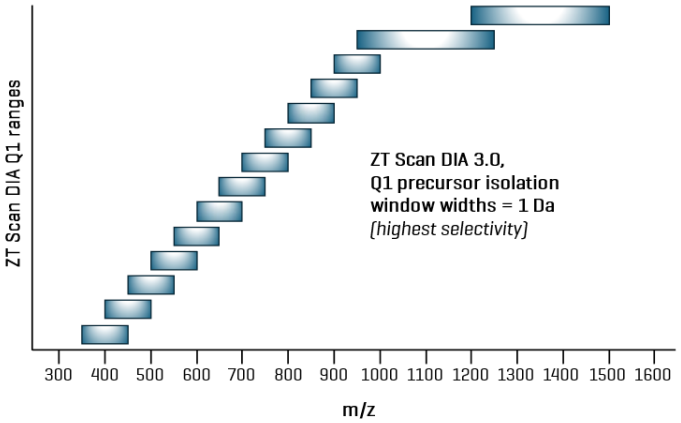
Patrick Pribil[1] and Remco van Soest[2]

[1]SCIEX, Canada; [2]SCIEX, USA

This technical note demonstrates how the combination of gas-phase fractionation and ZT Scan DIA 3.0 on the ZenoTOF 8600 system enables faster generation of deep empirical spectral libraries that enhance the performance of downstream DIA data analysis. By supporting Q1 isolation widths as narrow as 1 Da, ZT Scan DIA 3.0 delivers exceptional precursor selectivity and high-quality MS/MS spectra, maximizing peptide and protein identifications. When coupled with gas-phase fractionation, this approach provides a streamlined alternative to traditional off-line fractionation workflows, enabling the efficient creation of high-coverage spectral libraries with no additional sample preparation. The resulting libraries substantially improve protein group and precursor identifications compared with library-free processing approaches, allowing researchers to achieve deeper proteome coverage, greater quantitative confidence, and maximum analytical performance from the ZenoTOF 8600 system.

Key features of spectral libraries generated using ZT Scan DIA 3.0 on the ZenoTOF 8600 system

- **Ultra-selective DIA for deeper library generation:** ZT Scan DIA 3.0 supports Q1 isolation widths as narrow as 1 Da, delivering unit-resolution precursor selectivity to maximize peptide and protein identifications.
- **Rapid generation of deep empirical spectral libraries:** By combining gas-phase fractionation with narrow-window ZT Scan DIA 3.0, researchers can build high-coverage spectral libraries efficiently, eliminating the need for labor-intensive off-line fractionation workflows.
- **Improved DIA data processing performance:** Empirical spectral libraries generated using gas-phase fractionation and ZT Scan DIA 3.0 increased protein group and precursor identifications by up to 48% compared to library-free DIA processing approaches, enabling deeper and more confident proteome characterization.



ZT Scan DIA Q1 (m/z)	K562		HeLa	
	Protein groups	Precursors	Protein groups	Precursors
1200 - 1500	2,242	6,108	2,129	4,910
950 - 1250	5,518	26,442	5,428	23,203
900 - 1000	5,594	19,103	5,608	17,855
850 - 950	6,110	23,254	6,617	21.414
800 - 900	6,506	26,808	6,625	25,138
750 - 850	7,083	31,977	7,176	30,116
700 - 800	7,579	38,160	7,753	36,612
650 - 750	7,949	44,588	8,114	42,958
600 - 700	8,171	50,683	8,446	49,944
550 - 650	8,368	57,203	8,700	57,109
500 - 600	8,423	62,284	8,783	62,880
450 - 550	8,232	62,199	8,631	64,816
400 - 500	7,725	53,575	8,346	58,232
350 - 450	6,682	34,910	7,505	40,310

Total combined protein groups: **10,596**
Total combined precursors: **396,429**

Figure 1. Spectral library generation for K562/HeLa lysates using gas-phase fractionation (GPF) and ZT Scan DIA 3.0. K562 and HeLa lysate digests were analyzed using ZT Scan DIA (1 Da Q1 window widths) on the ZenoTOF 8600 system. Using GPF, the precursor mass range (350-1500 Da) was divided into 14 partially overlapping fractions for each sample. The combined dataset was processed using PEAKS Studio software version 13.5. Protein groups and precursors identified in each fraction are shown. A total of 10,596 protein groups and 396,429 precursors were identified in the combined fractions, which were exported as a spectral library for further evaluation.

Introduction

Mass spectrometry using data-independent acquisition (DIA) has become a cornerstone technology for quantitative proteomics, offering simplified method development, comprehensive proteome coverage, and robust qualitative and quantitative measurements. While advances in DIA acquisition continue to expand analytical performance, the software tools and data processing strategies used to interpret DIA data remain critical determinants of the final results. DIA datasets are commonly processed using either library-free workflows or spectral library-based approaches. Library-free searching tools and algorithms continue to evolve rapidly and offer the convenience of eliminating the need for pre-existing spectral libraries. However, empirical spectral libraries—built from experimentally observed precursor and fragment ion evidence—are widely recognized for providing greater depth, confidence, and consistency in peptide and protein identification (1).

A key challenge associated with empirical spectral libraries is the effort required to generate them. Traditional workflows often rely on time-consuming off-line fractionation strategies, such as high-pH fractionation, followed by multiple LC-MS analyses to achieve sufficient depth. Gas-phase fractionation (GPF) provides an attractive alternative by dividing the precursor m/z space into multiple narrower windows and analyzing each range separately across sequential runs, thereby reducing spectral complexity and enabling deeper coverage without additional sample preparation steps (2).

In this technical note, GPF is combined with ZT Scan DIA 3.0 on the ZenoTOF 8600 system to generate deep empirical spectral libraries for DIA data processing. ZT Scan DIA uses a continuously scanning quadrupole for precursor isolation, combined with Zeno trapping for enhanced MS/MS sensitivity (3). With ZT Scan DIA 3.0, Q1 isolation widths can be configured as narrow as 1 Da, delivering exceptional precursor selectivity and DDA-like MS/MS spectral quality to maximize peptide and protein identifications (4-6). Using this approach, deep spectral libraries were generated for both human lysate digests (K562 and HeLa) and a hybrid human/yeast/E.coli (HYE) proteome. The resulting libraries were evaluated using multiple DIA datasets acquired across different sample loadings, chromatographic configurations, and instrument platforms, including both the ZenoTOF 8600 and ZenoTOF 7600+ systems. When used for data processing in PEAKS Studio software, these empirical spectral libraries consistently increased the number of identified and quantified protein groups and precursors compared with library-free approaches. The results demonstrate a practical workflow for rapidly generating deep, application-specific spectral libraries using GPF and ZT Scan DIA 3.0, enabling users to maximize proteome coverage and quantitative performance without the need for labor-intensive off-line fractionation.

Methods

Sample preparation: Human K562 and yeast lysate tryptic digests were purchased from Promega. E.coli lysate tryptic digests were purchased from Waters. Human HeLa lysate tryptic digest was purchased from Pierce. Lysate digest dilutions were prepared in buffer containing 0.1% formic acid in water. Human (K562)/Yeast/E.coli (HYE) hybrid proteome lysates were mixed in 3 different weight-percent ratios as previously described (7). The hybrid proteome mixtures were pooled to create a standard reference sample and diluted to a concentration of 200 ng/µL. For spectral library generation using gas-phase fractionation, 200 ng of sample was loaded for all injections.

Chromatography: For spectral library generation, separations were performed using a Waters M-Class UPLC system (Waters). An IonOpticks Ultimate Elite XS C18 nanoflow column (25 cm x 0.075 mm) was used, heated to 50 °C. Samples were analyzed in direct-inject mode with a 38-min active gradient (85-min total run time) as described previously (4).

Mass spectrometry: Spectral libraries were generated using the ZenoTOF 8600 system equipped with the horizontal nanoflow probe. The ion source settings and ZT Scan DIA acquisition parameters used for library generation are summarized in Table 1. To perform gas-phase fractionation (GPF), the precursor mass range was divided into a series of discrete m/z windows. The GPF schemes used for the K562/HeLa digest and HYE hybrid proteome mixture are shown in Figures 1 and 2, respectively. The resulting spectral libraries were subsequently evaluated using previously acquired K562 and HYE datasets. These datasets were generated either on the ZenoTOF 8600 system coupled to the Evosep Eno system (8) or on the ZenoTOF 7600+ system coupled to the Waters M-Class UPLC system (9), as indicated for each experiment. These data were then used to

assess the performance of the newly generated GPF spectral libraries for protein and precursor identification and quantitation.

Table 1: ZenoTOF 8600 system parameters.

Parameter	Value
Method duration	80 min
Curtain gas	40
CAD gas	7
Nano cell temperature	250 °C
Nano gas 1	10 psi
Ion source gas 2	0 psi
Nano spray voltage	2,100 V
Total cycle time	1.1 sec
TOF MS m/z range	400-1500 Da
TOF MS accumulation time	50 msec
MS/MS m/z range	140-1750 Da
MS/MS accumulation time	9.9 msec *
Zeno mode	On
Dynamic collision energy	On (+2)
ZT Scan DIA Q1 isolation range	As indicated [1]
ZT Scan DIA Q1 width	1.0 Da
ZT Scan DIA Q1 scan speed	101 Da/sec *
Total MS/MS scan time	0.991 sec

[1] See Figures 1 and 2.

* For the 950-1250 Da and 1200-1500 Da Q1 windows (Figure 1), the MS/MS accumulation time was 3.3 msec, and the Q1 scan speed was 303 Da/sec.

Data processing: PEAKS Studio software version 13.5 was used for all data processing.

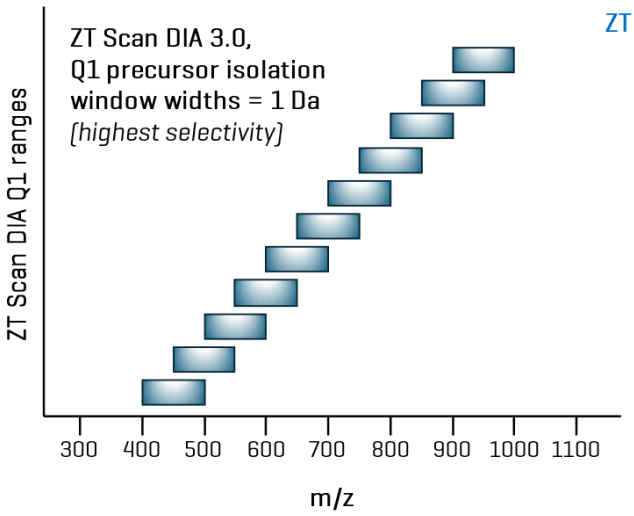
Construction of Human and HYE spectral libraries from acquired ZT Scan DIA 3.0 data (1 Da Q1 window widths): Human K562 and HeLa ZT Scan DIA 3.0 GPF data were searched together against a FASTA containing canonical Human protein sequences, while HYE data were searched together against a FASTA comprising canonical Human, Yeast and E.coli protein sequences, all downloaded from Uniprot (https://www.uniprot.org/). For all searches, carbamidomethylation C was set as a fixed modification, while deamidation N/Q and oxidation M were set as variable modifications. Match Between Runs (MBR) was enabled for all analyses, and LFQ normalization was turned off.

Evaluation of the novel Human and HYE spectral libraries: Previously acquired K562 or HYE hybrid proteome mixture data was used to evaluate the novel spectral libraries (8,9). For a given instrument/on-column loading/sample concentration, triplicate data files were searched together, using either (i) the FASTA for the given sample (i.e., DB searches), (ii) the novel GPF spectral libraries appropriate for that sample (i.e., Lib searches), or (iii) combined library and FASTA (i.e., Lib + DB searches). Fixed and variable modifications were set as above. MBR was used, and LFQ normalization mode was set to TIC (RT dependent). The numbers of identified and quantified protein groups and precursors were determined from the exported Protein and Feature Vector LFQ exports, respectively.

Generation of deep spectral libraries using gas-phase fractionation and ZT Scan DIA 3.0

K562 and HeLa lysate digests were analyzed using ZT Scan DIA 3.0 with 1 Da Q1 window widths to generate a deep empirical spectral library for downstream DIA data processing. To increase precursor coverage while reducing sample complexity per acquisition, a GPF strategy was used to divide the 350-1,500 Da precursor range into the segments shown in Figure 1. The combined fractions (28 total) were processed together in PEAKS Studio software version 13.5 using a FASTA comprising canonical human protein sequences downloaded from Uniprot. Figure 1 summarizes the protein groups and precursors identified in each fraction. Across all runs, this workflow identified 10,596 protein groups and 396,429 precursors, which were exported as a spectral library for further testing.

A second empirical spectral library was generated using a pooled sample of HYE digest analyzed with ZT Scan DIA 3.0, again using methods with 1 Da Q1 window widths. For this hybrid proteome sample, the 400-1,000 Da precursor mass range was divided into 100 Da-wide fractions (11 total), as shown in Figure 2. The 11 fractions were processed together in PEAKS Studio software version 13.5 using a FASTA comprising canonical Human, Yeast, and E.coli protein sequences. Figure 2 summarizes the protein groups and precursors identified in each fraction. In total, 15,015 protein groups and 316,225 precursors were identified and exported as a second spectral library.



ZT Scan DIA Q1 (m/z)	K562/Yeast/E.coli Protein groups	K562/Yeast/E.coli Precursors
900 - 1000	7,921	22,531
850 - 950	8,704	27,608
800 - 900	9,413	31,914
750 - 850	10,443	38,751
700 - 800	11,201	45,705
650 - 750	11,917	53,518
600 - 700	12,444	62,044
550 - 650	12,763	69,558
500 - 600	12,780	73,128
450 - 550	12,365	71,214
400 - 500	11,082	55,974

Total combined protein groups: **15,015**
Total combined precursors: **316,225**

Figure 2. Spectral library generation for a Human (K562)/Yeast/E.coli (HYE) hybrid proteome mixtures using gas-phase fractionation (GPF) and ZT Scan DIA 3.0. A pooled HYE hybrid proteome mixture was analyzed using ZT Scan DIA (1 Da Q1 window widths) on the ZenoTOF 8600 system. Using GPF, the precursor mass range (400-1000 Da) was divided into 11 partially overlapping fractions. The combined dataset was processed using PEAKS Studio software version 13.5. Protein groups and precursors identified in each fraction are shown. 15,015 total protein groups and 316,225 total precursors were identified in the combined fractions, which were exported as a spectral library for further evaluation.

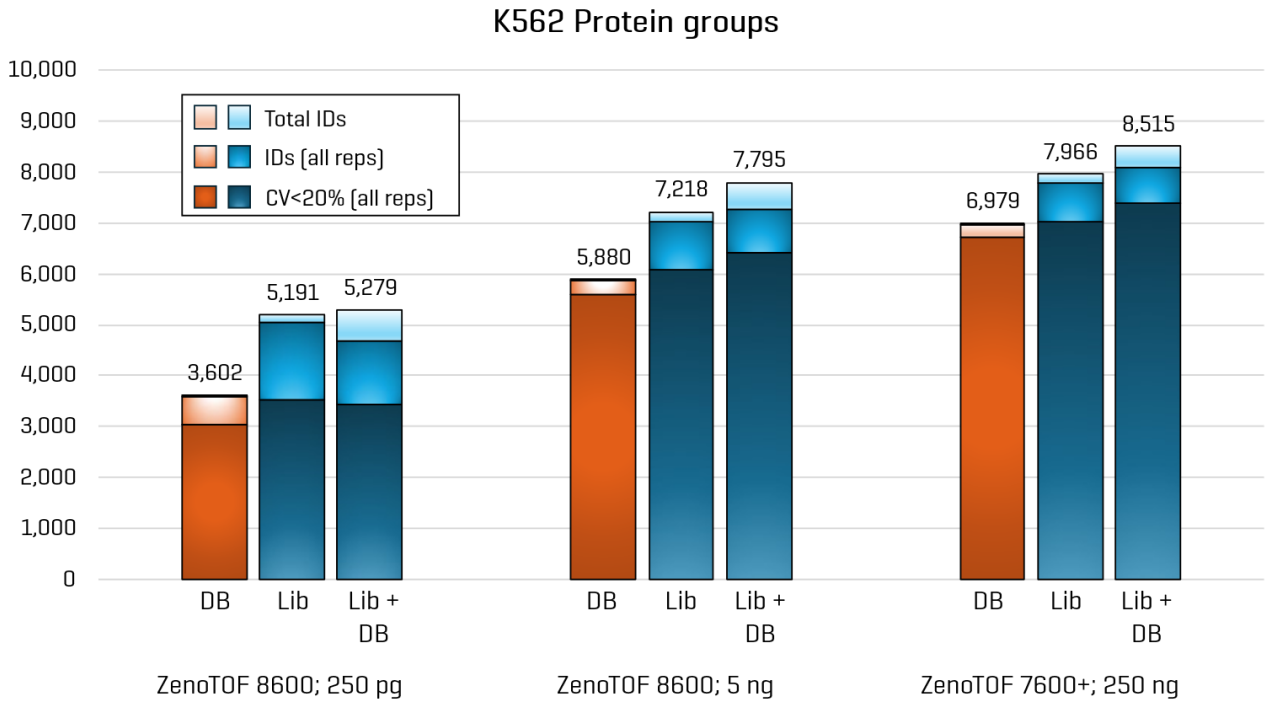


Figure 3. Analysis of various K562 digest ZT Scan DIA datasets using different processing methods – Protein Groups. K562 datasets acquired previously (instrument and sample loading indicated) were processed using PEAKS Studio software version 13.5, using either a library-free approach (DB), the novel K562/HeLa spectral library generated using GPF and ZT Scan DIA 3.0 (Lib), or a combined library + FASTA approach (Lib + DB). The total number of protein groups identified, those identified in all 3 replicates, and those quantified in all 3 replicates with CV<20%, are indicated for each analysis. The novel spectral library significantly increased the number of identified and quantified protein groups compared to library-free searches. Note that different chromatographic separation methods were used for the ZenoTOF 8600 system and ZenoTOF 7600+ system datasets (see References 7 and 8).

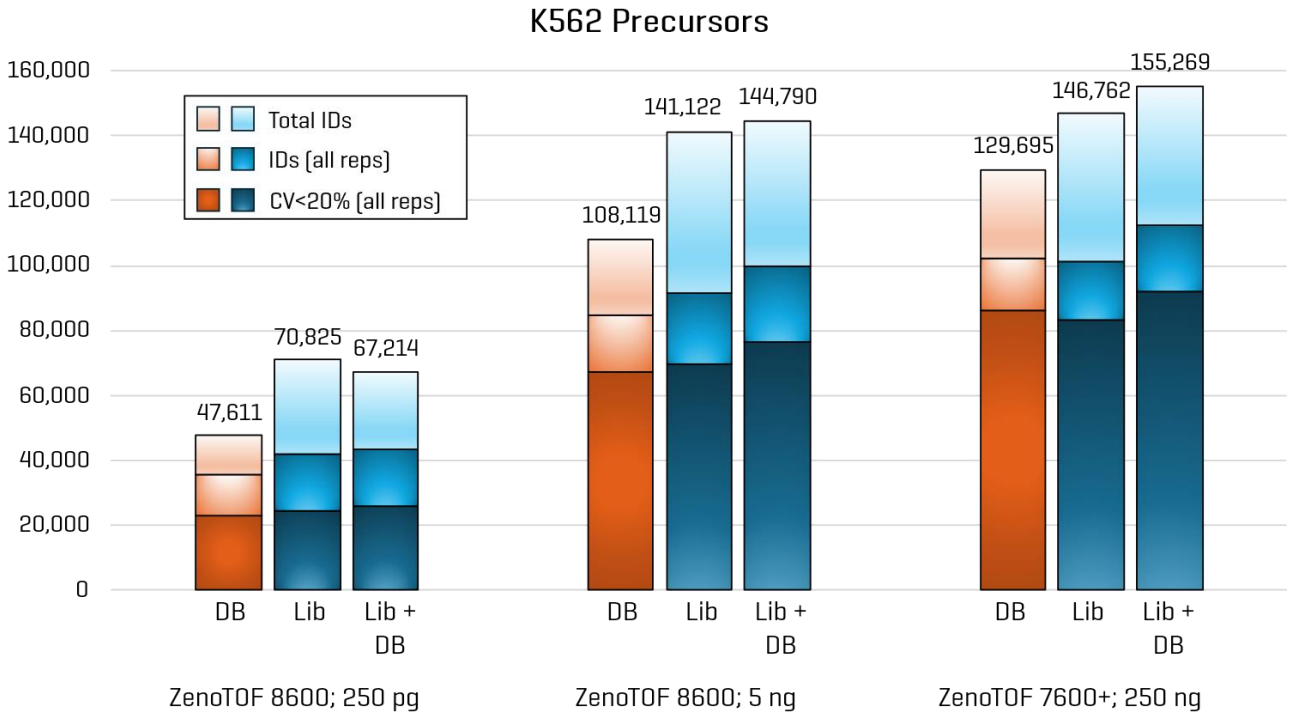


Figure 4. Analysis of various K562 digest ZT Scan DIA datasets using different processing methods – Precursors. Previously acquired K562 datasets were processed using PEAKS Studio software version 13.5, using either a library-free approach (DB), the novel K562/HeLa spectral library generated using GPF and ZT Scan DIA 3.0 (Lib), or a combined library + FASTA approach (Lib + DB). The total number of precursors identified, those identified in all 3 replicates, and those quantified in all 3 replicates with CV<20%, are indicated for each analysis. Note that different chromatographic separation methods were used for the ZenoTOF 8600 system and ZenoTOF 7600+ system datasets (see References 7 and 8).

DIA data processing using empirical spectral libraries improves overall identifications in protein groups and precursors

The novel K562/HeLa gas-phase fractionation spectral library was evaluated across multiple previously acquired K562 DIA datasets to assess its impact on protein group and precursor identifications. These datasets included (i) K562 data, either 250 pg or 5 ng loadings, previously acquired with the Whisper Zoom 40 SPD method on the Evosep Eno system and ZT Scan DIA (9.8 Da Q1 window widths) on the ZenoTOF 8600 system (8), and (ii) K562 data, 250 ng loadings, previously acquired using a nanoflow gradient on the Waters M-Class UPLC system and ZT Scan DIA (6.4 Da Q1 window widths) on the ZenoTOF 7600+ system (9). All datasets were processed in PEAKS Studio software version 13.5 using 3 different search modes: DB searches, Lib searches, and combined Lib + DB searches. The results are summarized in Figures 3 and 4.

Across all datasets, the use of the empirical spectral library increased the number of protein groups and precursors identified compared to DB-only searches. This improvement was observed across total IDs, IDs in all 3 replicates, and IDs quantified in all 3 replicates with CVs<20%. The combined Lib + DB searches generally provided the highest identification depth, increasing total protein groups by 22-47% and precursors by 20-41% compared to DB-only searches. Quantifiable protein groups and precursors were also generally highest with Lib + DB searches, with increases relative to DB searches ranging from 7-15%.

The HYE gas-phase fractionation spectral library was evaluated separately using a different previously-acquired dataset (8): a 250 pg loading of human/yeast/E.coli (65%/15%/20% proportional w/w/w mixture), acquired with the Whisper Zoom 40 SPD method on the Evosep Eno system and ZT Scan DIA (9.8 Da Q1 window widths) on the ZenoTOF 8600 system. As above,

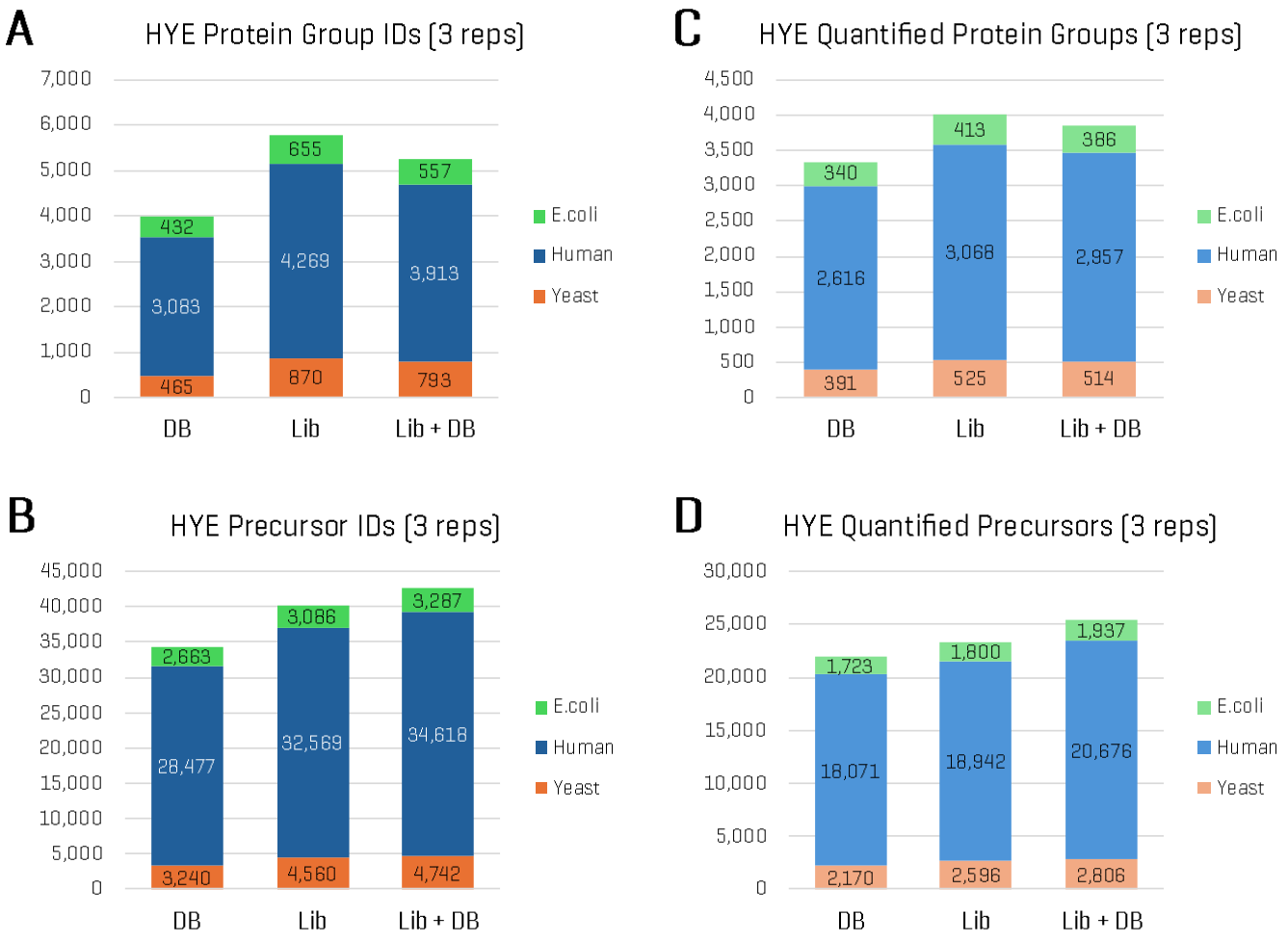


Figure 5. The novel HYE spectral library improves protein group/precursor identification and quantitation in 250 pg HYE digest. A previously acquired dataset from 250 pg of HYE hybrid proteome mixture was processed using PEAKS Studio software version 13.5, using either a library-free approach (DB), the novel HYE spectral library generated using GPF and ZT Scan DIA 3.0 (Lib), or a combined library + FASTA approach (Lib + DB). The number of protein groups (A) and precursors (B) identified in all 3 replicates, and protein groups (C) and precursors (D) quantified in all 3 replicates with CV<20%, are indicated, broken down by species. The novel spectral library significantly increased the number of identified and quantified features compared to library-free searches.

this dataset was searched using DB, Lib, and Lib + DB approaches. The results are shown in Figure 5. Protein groups and precursors identified in all 3 replicates, broken down by species, are shown in Figures 5A and 5B, respectively, while protein groups and precursors quantified with CVs<20% (broken down by species) are shown in Figures 5C and 5D.

Consistent with the K562/HeLa library evaluation, use of the HYE spectral library resulted in substantial gains in both identified and quantified protein groups/precursors compared to DB searches, with gains observed across all species in the hybrid proteome. Identified/quantified protein groups were highest with Lib searches, while identified/quantified precursors were highest with Lib + DB searches, highlighting the value of empirical spectral libraries for improving DIA data processing performance across complex proteomic samples.

Conclusions

The results presented here provide a practical framework for generating deep empirical spectral libraries, enabling researchers to maximize protein and precursor identifications and extract greater biological insight from DIA datasets through enhanced data processing depth and confidence.

- **Deep spectral libraries generated without off-line fractionation:** Gas-phase fractionation combined with 1 Da ZT Scan DIA 3.0 produced libraries containing 10,596 protein groups and 396,429 precursors (for K562/HeLa) and 15,015 protein groups and 316,225 precursors (for HYE hybrid proteome)
- **Improved identification performance:** Use of empirical spectral libraries consistently increased protein group and precursor identifications compared with library-free searches across all tested datasets
- **Maximum gains with combined Lib + DB searches:** The Lib + DB approach generally delivered the highest performance, providing 22–47% increases in total protein groups, 20–41% increases in precursors, and 7–15% improvements in quantifiable protein groups and precursors relative to DB-only searches

References

1. Koudelka T., et al. (2025). Benchmarking of Quantitative Proteomics Workflows for Limited Proteolysis Mass Spectrometry. Molecular & Cellular Proteomics, 24:100945.
2. Pino, L., et al. (2020). Acquiring and Analyzing Data Independent Acquisition Proteomics Experiments without Spectrum Libraries. Molecular & Cellular Proteomics, 19:1088.
3. Continuing the data independent acquisition (r)evolution: Introducing ZT Scan DIA for quantitative proteomics. SCIEX technical note, MKT-31819-A.
4. ZT Scan DIA 3.0: A data-independent acquisition (DIA) method with high-resolution precursor selection improves protein identification and quantitation for proteomics research. SCIEX technical note, MKT- 38093-A.
5. Narrow-window scanning DIA for unbiased label-free quantitation: Improved protein group ratio accuracy with ZT Scan DIA 3.0 on the ZenoTOF 8600 system. SCIEX technical note, MKT-38229-A.
6. Accelerating discovery proteomics: ZT Scan DIA 3.0 powers improved high-throughput proteomics identification and quantitation. SCIEX technical note, MKT- 38277-A.
7. Accurate label-free quantitative (LFQ) proteomics benchmarking using data-independent acquisition (DIA) on the ZenoTOF 8600 system. SCIEX technical note, MKT-38149-A.
8. Pushing performance for low-input proteomics applications using Enhanced Sensitivity Mode on the ZenoTOF 8600 system. SCIEX technical note, MKT-39244-A.
9. ZT Scan DIA 3.0 optimizes protein identification and quantitation on the ZenoTOF 7600+ system. SCIEX technical note, MKT- 39320-A.

Headquarters
250 Forest Street| Marlborough, MA 01752 USA
Phone 508-383-7700
sciex.com

International Sales
For our office locations please call the division headquarters or refer to our website at
sciex.com/offices