Systematic Optimization of Activity-Based Protein Profiling for Identification of Polysorbate-Degradative Enzymes in Biotherapeutic Drug Substance Down to 10 ppb

Taku Tsukidate[1], Anita P. Liu[1], Shannon Rivera[1], Alyssa Q. Stiving[1], Jonathan Welch[2], Xuanwen Li[1]*

1. Analytical Research & Development Mass Spectrometry, Merck & Co., Inc., 126 East Lincoln Avenue, Rahway, New Jersey 07065, United States.

2. Biologics Analytical Research & Development, Merck & Co., Inc., 126 East Lincoln Avenue, Rahway, New Jersey 07065, United States.

ABSTRACT: The identification and control of high-risk host cell proteins (HCPs) in biotherapeutics development is crucial for ensuring product quality and shelf life. Specifically, HCPs with hydrolase activity can cause degradation of excipient polysorbates (PS), leading to a decrease in the shelf life of the drug product. In this study, we systematically optimized every step of an activity-based protein profiling (ABPP) workflow to identify trace amounts of active polysorbate-degradative enzymes (PSDEs) in biotherapeutics process intermediates. Evaluation of various parameters during sample preparation pinpointed optimal pH level and fluorophosphonate (FP)-biotin concentration. Moreover, the combined use of a short liquid chromatography gradient and the fast-scanning Parallel Accumulation Serial Fragmentation (PASEF) methodology increased sample throughput without compromising identification coverage. Tuning Trapped Ion Mobility Spectrometry (TIMS) parameters further enhanced sensitivity. In addition, we evaluated various data acquisition modes including PASEF combined with data-dependent acquisition (DDA PASEF), data-independent acquisition (diaPASEF), or parallel reaction monitoring (prm-PASEF) as well as data processing strategies. By employing the newly optimized ABPP workflow, we successfully identified PSDEs at a concentration as low as 10 parts per billion (ppb) in a drug substance sample. Finally, the new workflow enabled us to detect a PSDE that could not be detected with the original workflow during a PS degradation root-cause investigation.

Biotherapeutics, such as therapeutic monoclonal antibodies (mAbs) products, have become a predominant therapeutic modality for the treatment of a wide range of diseases. The global sales revenue for all mAb drugs alone totaled nearly $115.2 billion in 2018 and is expected to reach $300 billion by 2025.[1,2] Biotherapeutics produced in host cells, such as Chinese hamster ovary (CHO), go through downstream purifications to remove product and process related impurities, such as host cell proteins (HCPs). The identification and control of those high-risk HCPs is critical to biotherapeutics development.[3] For example, certain HCPs with hydrolase activity can cause formulation excipient polysorbate degradation if they are not effectively removed during downstream purification, which can significantly shorten the shelf life of the drug product.[4] If proven in a manufacturing setting even a modest 5–10 % increase in product shelf life resulting from the reduction and control of HCPs would translate to millions of dollars of cost saving, not to mention significant improvement in downstream purification process specifically designed to remove those high-risk HCPs. The increased product shelf life may hopefully reduce the drug cost and improve the equitable access to those life-saving medicines.

Liquid chromatography tandem mass spectrometry (LC–MS/MS)–based proteomics has proven to be a powerful approach to HCP analysis which enables tracking of individual HCPs throughout a purification process.[5,6] HCP analysis poses a unique challenge to conventional abundance-based proteomics because of the large dynamic range of samples that can span more than six orders of abundance between the therapeutic antibody and low-abundance HCPs. Accordingly, the past several years saw the development of various sample preparation methods such as native digestion[7,8], molecular weight cutoff[9], and ProteoMiner enrichment[10] that improved sensitivity for low-abundance HCPs in the range of 0.05–1 ppm. However, abundance-based proteomics does not offer information about the functional state of HCPs that would help determine the cause of an issue at hand such as excipient degradation. In contrast, activity-based protein profiling (ABPP) employs active site–directed chemical reporters to selectively enrich active enzymes in complex samples.[11,12] Thus, we and others have previously reported ABPP protocols for HCP analysis which enables sensitive detection and functional assessment of specific HCPs.[13,14] More recently, Liu et al. developed a polysorbate-inspired chemical reporter and demonstrated the detection of polysorbate-degradative enzymes (PSDEs) down to 80 ppb.[15] Herein, we describe further optimization of the ABPP approach and demonstrate confident identification of active PSDEs at the 10-ppb level. The established ABBP workflow will fill the gap between PSDE abundance and activity, allowing more meaningful polysorbate degradation investigations for biotherapeutic development.

Materials and methods

Protein samples. Recombinant PLA2G7 and SIAE proteins were custom ordered from GenScript. mAb1 harvested cell

culture fluid (HCCF), mAb1 Protein A pool (PAP), mAb1 drug substance (DS), and mAb2 ultrafiltration pool (UFP) were internally sourced.

ABPP sample preparation from HCCF. HCCF (100 μL, ~ 0.5 mg) was buffer-exchanged with either 100 mM sodium acetate pH 5, 100 mM histidine pH 6, DPBS pH 7, or 50 mM tris-HCl pH 8 through a Zeba Spin Desalting Column, 7K MWCO, 0.5 mL (Thermo Fisher Scientific) according to the manufacturer's instructions. The buffer-exchanged material was incubated with fluorophosphonate (FP)-biotin (WuXi AppTec; 0, 0.1, 1, or 10 μM) at 25 °C for 0.5 or 2 h and then buffer-exchanged with DPBS pH 7 through a Zeba Spin Desalting Column to remove excess FP-biotin and adjust pH. To this were added 0.2 % SDS / DPBS (900 μL) and Pierce Streptavidin Magnetic Beads (20 or 40 μL; Thermo Fisher Scientific). After 1-h incubation at room temperature, the beads were washed three times with 0.2 % SDS / DPBS pH 7 (1 mL) and three times with 1 M urea / water (1 mL). Then the beads were incubated with 0 or 10 mM TCEP-HCl or 0 or 20 mM iodoacetamide at 25 °C for 0.5 h and washed twice with 50 mM tris-HCl pH 8 (1 mL). The beads were incubated with 50 mM tris-HCl pH 8 (50 μL) containing 0.1 μg trypsin (Promega) at 25 °C for 20 h. The supernatant (~ 50 μL) was cleaned up with Pierce C18 Spin Column (Thermo Fisher Scientific) according to the manufacturer's instructions, vacuum-dried, and reconstituted with 0.1 % FA / water (5 μL). Two μL was injected. Alternatively, in our optimal protocol, HCCF (100 μL) was directly incubated with FP-biotin (10 μM) at 25 °C for 1 h and filtered through a Zeba Spin Desalting Columns. The filtrate was diluted with 0.2 % SDS / DPBS (900 μL) and incubated with Pierce Streptavidin Magnetic Beads (20 μL) at room temperature for 1 h. The beads were washed three times with 0.2 % SDS / DPBS pH 7 (1 mL) and three times with 1 M urea / water (1 mL), followed by incubation with 50 mM tris-HCl pH 8 (50 μL) containing 0.1 μg trypsin (Promega) at 25 °C for 20 h. The supernatant (20 μL) was loaded on Evotip Pure (Evosep) according to the manufacturer's instructions.

ABPP sample preparation from PAP, UFP or DS. PAP, UFP or DS (5 mg) was diluted to 5 g/L with 100 mM histidine pH 6 and was incubated with 10 μM FP-biotin at 25 °C for 1 h and filtered through a Zeba Spin Desalting Column, 7K MWCO, 2 mL (Thermo Fisher Scientific). To this were added DPBS (4 mL), 10 % SDS (100 μL), and streptavidin magnetic beads (200 μL). After 1-h incubation at room temperature, the beads were washed three times with 0.2 % SDS / DPBS (5 mL) and three times with 1 M urea / water (5 mL). Then the beads were transferred to a microcentrifuge tube using 50 mM tris-HCl (1 mL) and incubated with 50 mM tris-HCl pH 8 (50 μL) containing 1 μg trypsin at 25 °C for 20 h. The supernatant (~ 50 μL) was loaded on Evotip Pure (Evosep). Alternatively, mAb2 UFP (20 mg) was processed as we previously described for comparison.[13]

LC–MS analysis. Peptides were separated on the Evosep One LC system[16] with Pepsep C18 columns (Bruker) using the manufacturer's pre-defined methods denoted by daily sample throughput e.g., 30 samples per day (SPD) and analysed on the Bruker timsTOF Pro 2 system. The DDA PASEF method was optimized based on the standard DDA PASEF method ("DDA PASEF-standard_1.1sec_cycletime.m"). The diaPASEF method was optimized as we previously described and covered 300–1200 m/z with 25 diaPASEF scans and two ion mobility windows per diaPASEF scan (i.e., 50 variable windows).[6] The prm-PASEF method targeted all precursors that were identified for PLA2G7 and SIAE in the DDA PASEF analysis of the mAb1 HCCF ABPP sample. The time and mobility scheduled acquisition boxes were set with 60 s of tolerance on retention time and 0.05 $1/K_0$ on ion mobility. In all methods, the ion mobility range was set as 0.7–1.43 Vs cm^{-2} with 100 ms of accumulation time and 100 ms of ramp time unless otherwise noted. Alternatively, peptides were separated on the EASY-nLC 1200 system and analyzed on Q Exactive HF-X system as we previously described.[13]

Data analysis. DDA PASEF data were processed with MaxQuant v2.4.10 or v2.6.1[17]. diaPASEF data were processed with DIA-NN v1.8.1 or v1.9.1[18]. Protein quantification was performed in R version 4.0.3 using the fast_MaxLFQ function from the iq package[19] based on the Precursor.Normalised column of DIA-NN main reports, which represents the aggregated and normalized MS2 quantities. Prm-PASEF data were processed with Skyline v24.1.0.199[20]. Precursors were filtered by requiring isotope dot product > 0.9 and their intensities were calculated by summing all associated fragment peak areas and subtracting background signals. Protein intensities were then calculated as sums of the intensities of the three most abundant precursors. The precursor and protein FDR thresholds were set at 1 % for HCCF- or PAP-based samples and at 5 % for UFP- or DS-based samples.[21] The FrF2[22] and rsm[23] packages were used to generate and analyze the Plackett-Burman screening design and central composite design, respectively.

High-performance liquid chromatography–charged aerosol detection (HPLC-CAD) for PS degradation analysis. HPLC-CAD was performed on Waters 1200 Series and Thermo Corona Charged Aerosol Detector with Waters Oasis Max column and a 20-min gradient. Mobile phases were (A) 0.5 % acetic acid in water and (B) 0.5 % acetic acid in isopropanol.[24]

Results and discussion

The active site–directed chemical reporter FP-biotin targets over 100 murine and human metabolic serine hydrolases (mSHs) in cell and tissue samples.[25,26] While the standard protocol labels mSHs with FP-biotin at pH 7–8[27], the optimal pH level for enzymatic activity varies for different mSHs and can affect their FP-biotin labeling levels[28,29]. Other sample preparation parameters such as FP-biotin labeling time, bead volume, reductant and alkylation reagent concentrations could also impact assay sensitivity as well as operational cost and/or hands-on time. To identify sample preparation parameters with significant effect on sensitivity towards various mSHs, we performed a twenty-sample Plackett–Burman screening experiment including six aforementioned parameters at two levels (**Table 1**). Indeed, varying pH level resulted in differential enrichment of several mSHs (**Figure 1A**). For example, lysosomal mSHs, namely PRCP and PLA2G15, were significantly enriched in samples labelled at pH 5 whereas several other mSHs including LPL, IAH1 and SIAE were significantly enriched in samples labelled at pH 8, which is consistent with their optimal pH levels for activity[29,30]. In addition, the higher FP-biotin concentration (10 μM) increased protein intensities selectively for mSHs compared to using the lower FP-biotin concentration (1 μM), indicating limited off-target labelling of other inactive proteins. On the other hand, the other four parameters had negligible effect. These results suggest that the pH level during the FP-biotin labelling reaction and the FP-

biotin concentration are the key parameters for sample preparation and merit further optimization.

Table 1. Plackett-Burman design with six factors at two levels and twenty runs for the optimization of ABPP sample preparation.

Run	pH	[FP-biotin] / μM	Labeling time / h	Bead volume / μL	[TCEP] / mM	[IA] / mM
1	5	10	2	20	0	20
2	8	10	2	20	10	0
3	8	1	2	40	0	0
4	5	1	2	40	10	20
5	5	1	2	40	0	20
6	8	10	0.5	20	10	20
7	8	10	0.5	40	0	20
8	5	1	0.5	20	0	0
9	8	1	0.5	20	0	20
10	8	10	2	40	0	20
11	5	10	2	20	10	20
12	5	1	0.5	40	10	0
13	5	10	0.5	20	0	0
14	5	10	2	40	10	0
15	8	1	2	20	0	0
16	5	1	0.5	20	10	20
17	8	1	2	20	10	0
18	8	10	0.5	40	10	0
19	5	10	0.5	40	0	0
20	8	1	0.5	40	10	20

Accordingly, we optimized FP-biotin concentration and pH level. First, we varied FP-biotin concentration from 0.1 μM to 100 μM in mAb1 HCCF at pH 8. As expected, increasing concentrations resulted in more mSH identifications and higher protein intensity up to 10 μM (**Figure 1B**). However, further increase in the FP-biotin concentration reduced protein identification and intensity. Perhaps, some of the excess FP-biotin or its degradant passed through desalting spin column and competed for streptavidin beads. Second, we varied the pH level from 5 to 8 while fixing the FP-biotin concentration at 10 μM. FP-biotin labelling at pH 7 yielded the highest number of mSH identifications, which was closely followed by pH 6 by one identification (**Figure 1C**). Protein intensity for the majority of known PSDEs, namely LPL, LIPA, PLA2G15 and SIAE, peaked between pH 6 and pH 7, which is slightly higher than the pH levels of most biopharmaceutical formulations (**Figure 1D**). On the other hand, somewhat lower pH was preferable for PLA2G7 while higher pH levels were preferable for several other PSDEs such as CES1 homologs and IAH1. Indeed, a recent study reported optimal pH levels for hydrolytic activity of several purified recombinant PSDEs and confirmed our observations.[30] These results suggest that performing the FP-biotin labeling step at pH 6 would increase the chance of identifying active and relevant PSDEs in biopharmaceutical formulations.

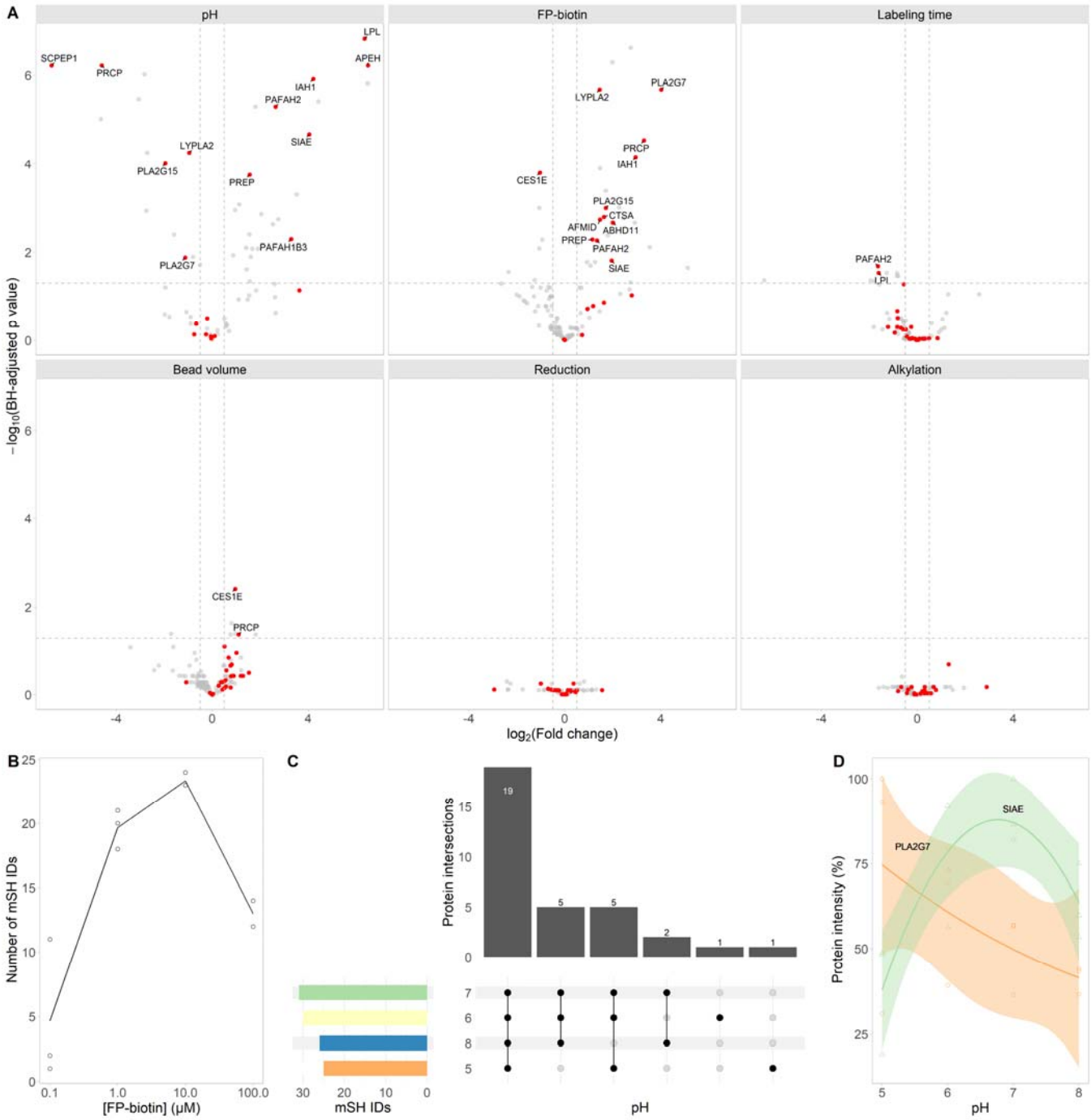

Figure 1. Optimization of ABPP sample preparation. (A) Six parameters were screened according to **Table 1**. Volcano plots indicate how individual mSHs were impacted by the parameter change. Horizontal dashed lines represent adjusted p value < 0.05. Vertical dashed lines represent $|\log_2(\text{fold change})| > 0.5$. (B) Number of mSH identifications at different FP-biotin concentrations. (C) Number and overlap of mSH IDs at different pH levels. (D) Relative quantification of representative PSDEs at different pH levels.. mAb1 HCCF. (B–D) n = 3.

The combined use of short gradients on Evosep One LC and the fast-scanning PASEF methodology on timsTOF Pro 2 allows for increased sample throughput without compromising ID coverage or data quality.[6,31] For example, Jones et al. reported a considerable increase in sample throughput in the ABPP of deubiquitylating enzymes in cell and tissue samples using Evosep One LC coupled with timsTOF Pro compared to a conventional nano-flow LC setup with an Orbitrap MS.[32] Thus, we evaluated several Evosep methods of varying throughput from 200 SPD to 30 SPD in combination with one of the default DDA PASEF methods ("DDA PASEF-short_gradient_0.5sec_cycletime.m") on timsTOF Pro 2 and benchmarked against the conventional nano-flow LC-MS setup consisting of EASY-nLC 1200 and Orbitrap Q Exactive HF-X (**Figure 2**). The number of mSH identifications increased as the gradient became longer and surpassed the level of identification coverage achievable by the conventional setup despite using a gradient that is less than half in length (44 min vs 110 min), representing nearly four-fold improvement in throughput (30 SPD vs ~ 8 SPD). In addition, Evotip can

load, concentrate, and clean up a relatively large volume (e.g., 20–200 μL) of sample and no longer requires offline clean up, concentration, or reconstitution of samples, which simplifies the protocol and minimizes sample loss. These results demonstrate the advantage of combining a short-gradient LC and the PASEF methodology in ABPP of biologics process intermediates.

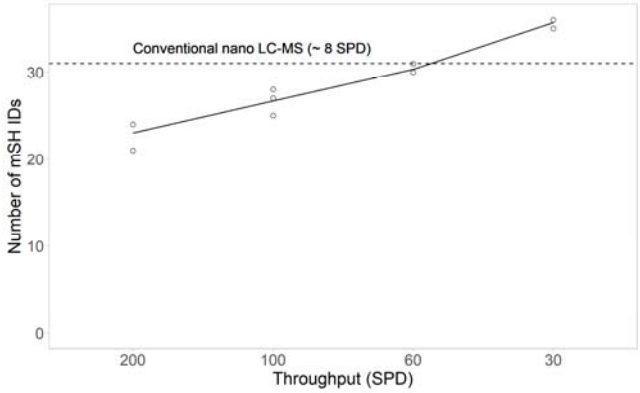

Figure 2. Optimization of Evosep One LC separation method. Conventional nano LC-MS refers to the EASY-nLC 1200–Orbitrap Q Exactive HF-X system. mAb1 HCCF. n = 3.

Previous studies have demonstrated that increasing the TIMS accumulation and ramp times allows fragmenting more precursor ions per PASEF scan and improve sensitivity.[31] For example, Mun et al. reported the benefit of longer TIMS accumulation and ramp times (180 ms) in single-cell proteomics applications.[33] On the other hand, the number of PASEF scans should be varied to control for cycle time and sufficient number of data points per peak for quantification.[34] Thus, we prepared ABPP samples from mAb1 PAP and analyzed these samples under various settings for the TIMS accumulation and ramp times and the number of PASEF scans. Specifically, we applied an eleven-run central composite design centered on the standard DDA PASEF settings and systematically varied the TIMS accumulation and ramp times from the standard values 100 ms to values between 30 ms and 170 ms as well as the number of PASEF scans from the standard value 10 to values between 3 to 17, while operating the instrument near 100% duty cycle (**Table 2**). Indeed, increasing TIMS accumulation and ramp times significantly improved the number of peptide identifications (**Figure 3A**) and representative protein intensity (**Figure 3B**), whereas the number of PASEF scans did not significantly affect either of these measures. In an experiment with somewhat small amount (50 ng) of HeLa protein digest, we observed a similar trend but more pronounced effect of the number of PASEF scans than in the experiment with the ABPP sample, presumably because of increased sample complexity (**Supporting Information Figure S1**). On the other hand, the median number of data points per precursor peaks strongly depended on the TIMS accumulation and ramp times as well as the number of PASEF scans, reflecting the cycle time (**Figure 3C**). These results suggest that DDA PASEF methods for the analysis of biologics ABPP samples should have longer accumulation and ramp times (e.g., 200 ms) as well as smaller number of PASEF scans (e.g., 5 scans) than the standard DDA PASEF method to increase sensitivity while maintaining cycle time (e.g., 1.24 s), hence quantitative accuracy and precision. Analogously, increasing injection time and automated gain control target values on Orbitrap-based instruments would improve sensitivity.[35]

Table 2. Central composite design for the optimization of TIMS parameter settings for the DDA PASEF analysis of ABPP samples.

Run	Accumulation time (ms)	Ramp time (ms)	Number of PASEF scans	Overlap	Cycle time (s)
1	50	50	5	1	0.34
2	150	150	5	1	0.94
3	50	50	15	4	0.90
4	150	150	15	4	2.50
5	100	100	10	4	1.17
6	100	100	10	4	1.17
7	30	30	10	4	0.40
8	170	170	10	4	1.94
9	100	100	3	1	0.42
10	100	100	17	4	1.91
11	100	100	10	4	1.17

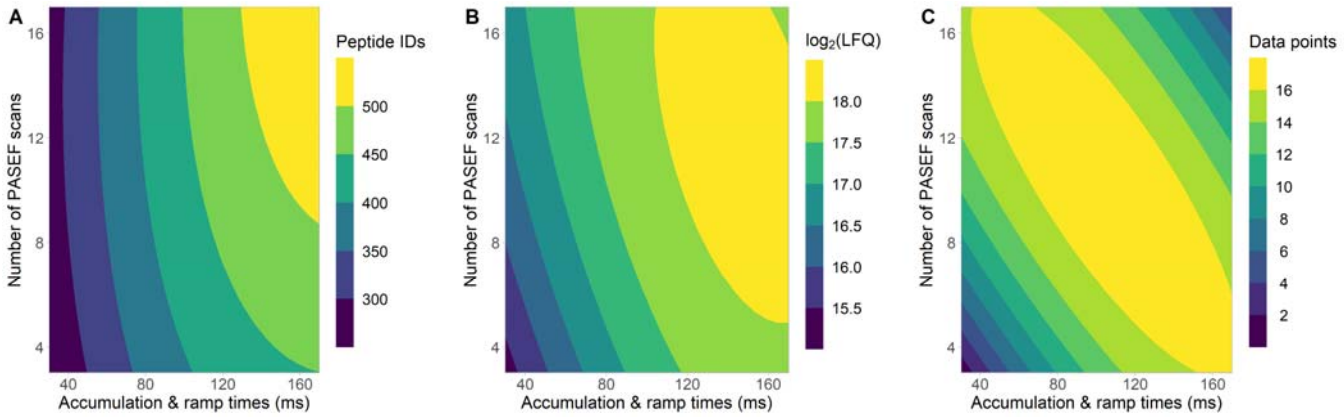

Figure 3. Optimization of TIMS parameter settings for DDA PASEF analysis of ABPP samples. TIMS accumulation and ramp times and number of PASEF scans were varied according to **Table 2**. Response surface models of second order were fitted to (A) number of peptide identifications, (B) label-free quantification of PLA2G7, or (C) median number of data points per peaks ("Number of scans" column of "evidence.txt" file generated by MaxQuant).

Having optimized sample preparation, LC method, and DDA PASEF method, we next evaluated the detection limit of this new workflow. To this end, we spiked mAb1 DS, which lacks measurable PS degradation activity (data not shown), with 1, 10, 100, or 1000 ppb of the purified recombinant PSDEs PLA2G7 and SIAE proteins. Gratifyingly, PLA2G7 was identified between 10–1000 ppb and SIAE was identified between 1–1000 ppb (**Figure 4A**), which was strongly supported by confident identifications of multiple (four for PLA2G7 at 10 ppb and five for SIAE at 1 ppb) unique peptides with high-quality spectra in all replicates (**Figure 4B**). This result updates the lowest limit of detection in literature (80 ppb) by eight-fold while using a commercially available chemical reporter.

To assess the potential for further enhancing sensitivity by employing alternative data acquisition modes that may be less limited by intra-scan dynamic range, we evaluated the diaPASEF[36] and prm-PASEF[37] acquisition modes. We performed a diaPASEF analysis with a general-purpose acquisition scheme[6] and processed the data with DIA-NN using a spectral library generated by refining an *in-silico* predicted spectra with experimental spectra from the diaPASEF analysis of the mAb1 HCCF ABPP samples, after evaluating multiple data processing strategies (**Supporting Information Figure S2**). For example, the diaPASEF analysis of mAb1 PAP-derived ABPP samples combined with our best data processing strategy led to the identification of 20 additional mSHs that were missed by the DDA PASEF analysis of the same ABPP sample. The prm-PASEF method targeted all precursors that were identified for PLA2G7 and SIAE in the DDA PASEF analysis of the mAb1 HCCF samples (**Supporting Information Table S1**). Both diaPASEF and prm-PASEF identified PLA2G7 down to 10 ppb and SIAE down to 1 ppb with multiple unique peptides supported by high-quality spectra and trended to be more precise than DDA PASEF by approximately five percentage points in terms of coefficient of variance. While the DIA data processing software confidently (posterior error probability < 0.01) reported the identification of two unique peptides derived from PLA2G7 even at the 1-ppb level, manual inspections of corresponding spectra found the assignments unconvincing (data not shown), which reiterates the importance of evaluating data quality. Both methods could be further refined to boost sensitivity and expand the coverage. For example, focusing the diaPASEF method on a smaller mass range[38] would allow the use of narrower windows and/or longer TIMS accumulation and ramp times. Alternatively, advanced varieties of PASEF-based DIA methods such as Slice-PASEF[39], which reportedly outperforms diaPASEF for low sample amounts, should be explored. As for the prm-PASEF method, collision energy optimization

could potentially significantly improve sensitivity.[40] In addition, expanding the targets beyond PLA2G7 and SIAE to include other PSDEs is desirable.

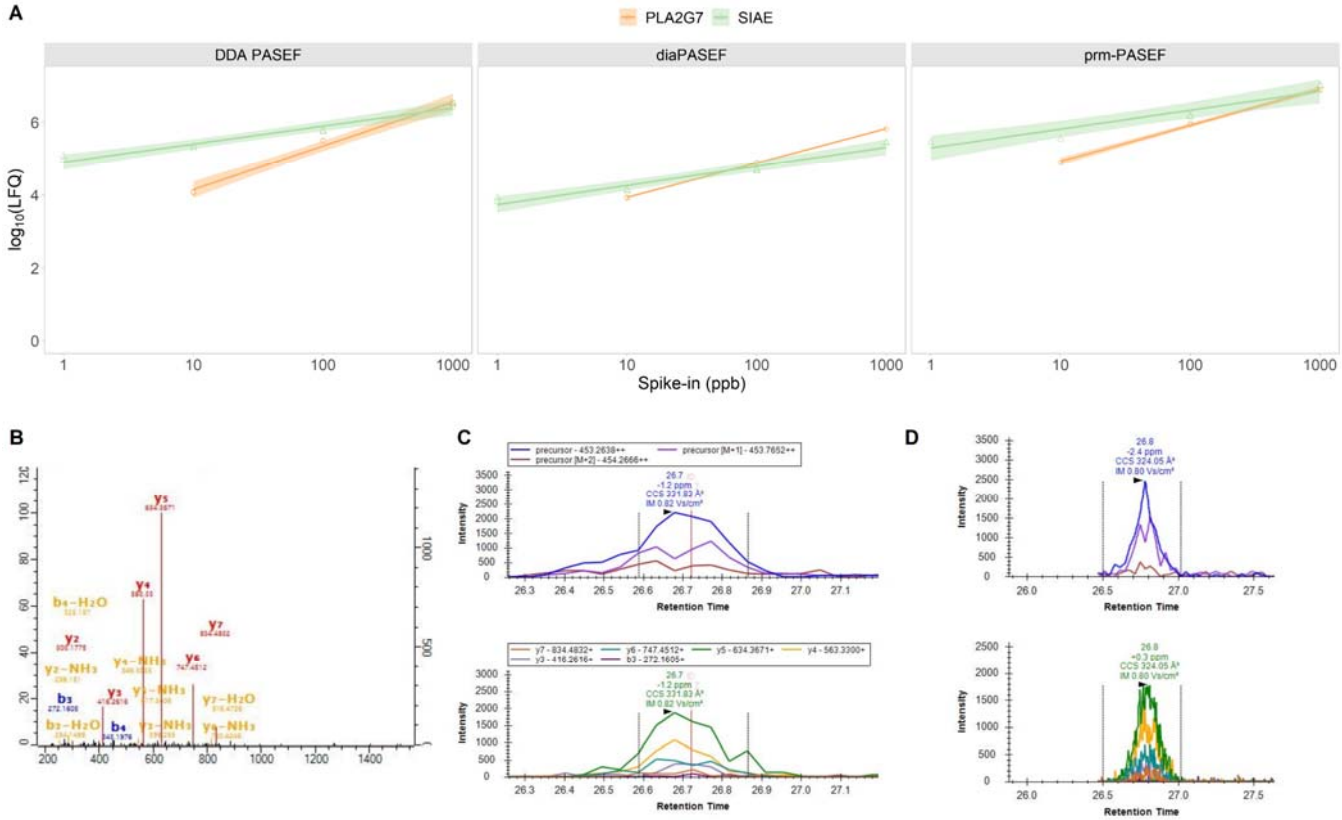

Figure 4. Limit of detection of new ABPP workflow. Purified recombinant PLA2G7 and SIAE were spiked in mAb1 DS at different concentrations of 1, 10, 100, or 1000 ppb. (A) Relative quantification of PLA2G7 and SIAE with three different MS data acquisition modes: DDA PASEF, diaPASEF, or prm-PASEF. Representative MS/MS spectrum or extracted ion chromatograms for the PLA2G7-derived peptide ASLAFLQR detected at the spike-in concentration of 10 ppb with (B) DDA PASEF, (C) diaPASEF, or (D) prm-PASEF. n = 2.

Finally, we leveraged the optimized workflow in a root-cause investigation of PS degradation during the mAb2 process development. mAb2 UFP was incubated with PS-80 for two months and residual PS-80 was monitored with an HPLC-CAD assay according to our standard protocol (**Supporting Information Figure S3**). The material exhibited considerable PS-80 degradation, especially at the elevated temperatures of 25 °C and 40 °C, prompting an investigation. To identify potential enzymatic cause of the PS degradation, we performed an ABPP experiment using the new workflow. For comparison, we also conducted an additional ABPP experiment using our original workflow[13], which requires 20 mg of input per sample, performs FP-biotin labeling at pH 8.0, and acquires data on the conventional nanoflow LC–MS setup with a 110-min gradient (**Table 3**). With the original workflow, we were only able to identify SIAE as an active enzyme with one unique peptide. Unfortunately, SIAE is unlikely to be the cause of observed PS-80 degradation because SIAE hydrolyses PS-20 but not PS-80.[41–43] In contrast, the new workflow enabled us to identify IAH1 (**Figure 5B**) and SIAE (**Figure 5C**) with multiple unique peptides supported by high-quality spectra. From the standard curb in **Figure 4A**, the amount of SIAE in this material was estimated to be 0.03 ppb (95% confidence interval: 0–0.07 ppb). While IAH1 hydrolyses PS-20, whether it also hydrolyses PS-80 remains unknown and warrants further investigation.[30,41] Nevertheless, these results clearly demonstrate that the new workflow outperformed the original method and will advance our understanding the root causes of PS degradation.

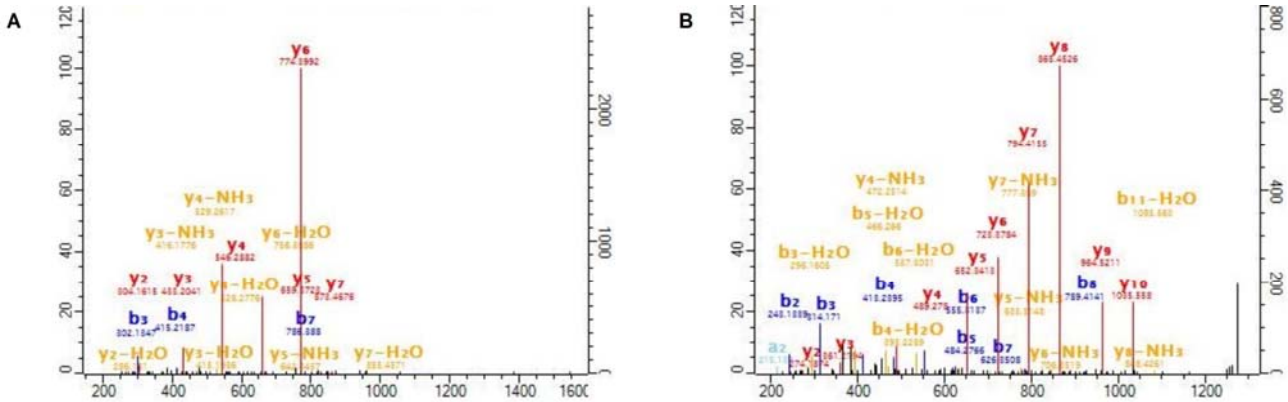

Figure 5. Identification of potential PSDEs in mAb2 UFP. Representative spectra for (A) IAH1-derived peptide SVDILEER and (B) SIAE-derived peptide ELAVAAAYQSVR. n = 3.

Table 3. Summary of identified PSDEs by original and new ABPP workflow.

Material	Workflow	Input (mg)	N	PSDE IDs (number of unique peptides; CV)
mAb2 UFP	New	5	3	IAH1 (3, 4, 3; 35 %); SIAE (2, 2, 1; 14 %)
mAb2 UFP	Original[13]	20	1	SIAE (1; -)

Conclusions

We systematically optimized the FP-biotin–based ABPP workflow for challenging biotherapeutics samples. We comprehensively evaluated every step of the workflow from sample preparation to LC separation and MS data acquisition methods and achieved the identification of as low as 10 ppb of several PSDEs in DS. The new workflow represents the most sensitive analytical tool for PSDE identification. Its enhanced sensitivity as well as reduced sample requirement and instrument time has already proved to be advantageous over the original method[13] in a root-cause investigation. Our future work will optimize and extend the targeted proteomics approach for full coverage of known PSDEs. In summary, the new workflow should advance the understanding of biopharmaceutical process and facilitate process and product development.

ASSOCIATED CONTENT

Supporting Information

(Word Style "TE_Supporting_Informati**on").** A listing of the contents of each file supplied as Supporting Information should be included. For instructions on what should be included in the Supporting Information as well as how to prepare this material for publication, refer to the journal's Instructions for Authors.

The Supporting Information is available free of charge on the ACS Publications website.

brief description (file type, i.e., PDF)

brief description (file type, i.e., PDF)

AUTHOR INFORMATION

Corresponding Author

* XL: xuanwen.li@merck.com

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

TT thanks MRL Postdoctoral Program. The authors thank Kaniz Fatema for executing experiments. The authors thank Douglas D. Richardson and Hillary A. Schuessler for guidance and support. The authors thank our internal Polysorbate Degradation and Lipase Control Strategy (PDLCS) working group for discussions.

REFERENCES

(1) Mullard, A. 2022 FDA Approvals. *Nat. Rev. Drug Discov.* **2023**, *22* (2), 83–88. https://doi.org/10.1038/d41573-023-00001-3.

(2) Walsh, G.; Walsh, E. Biopharmaceutical Benchmarks 2022. *Nat. Biotechnol.* **2022**, *40* (12), 1722–1760. https://doi.org/10.1038/s41587-022-01582-x.

(3) Jones, M.; Palackal, N.; Wang, F.; Gaza-Bulseco, G.; Hurkmans, K.; Zhao, Y.; Chitikila, C.; Clavier, S.; Liu, S.; Menesale, E.; Schonenbach, N. S.; Sharma, S.; Valax, P.; Waerner, T.; Zhang, L.; Connolly, T. "High-Risk" Host Cell Proteins (HCPs): A Multi-Company Collaborative View. *Biotechnol. Bioeng.* **2021**, *118* (8), 2870–2885. https://doi.org/10.1002/bit.27808.

(4) Li, X.; Wang, F.; Li, H.; Richardson, D. D.; Roush, D. J. The Measurement and Control of High-Risk Host Cell Proteins for Polysorbate Degradation in Biologics Formulation. *Antib. Ther.* **2022**, *5* (1), 42–54. https://doi.org/10.1093/abt/tbac002.

(5) Doneanu, C. E.; Anderson, M.; Williams, B. J.; Lauber, M. A.; Chakraborty, A.; Chen, W. Enhanced Detection of Low-Abundance Host Cell Protein Impurities in High-Purity Monoclonal Antibodies Down to 1 Ppm Using Ion Mobility Mass Spectrometry Coupled with Multidimensional Liquid Chromatography. *Anal. Chem.* **2015**, *87* (20), 10283–10291. https://doi.org/10.1021/acs.analchem.5b02103.

(6) Tsukidate, T.; Stiving, A. Q.; Rivera, S.; Sahoo, A.; Madabhushi, S.; Li, X. DiaPASEF Enables High-Throughput Proteomic Analysis of Host Cell Proteins for Biopharmaceutical Process Development. *Anal. Chem.* **2024**, *96* (32), 12999–13006. https://doi.org/10.1021/acs.analchem.4c00977.

(7) Huang, L.; Wang, N.; Mitchell, C. E.; Brownlee, T.; Maple, S. R.; De Felippis, M. R. A Novel Sample Preparation for Shotgun Proteomics Characterization of HCPs in Antibodies. *Anal. Chem.* **2017**, *89* (10), 5436–5444. https://doi.org/10.1021/acs.analchem.7b00304.

(8) Yang, F.; Li, D.; Kufer, R.; Cadang, L.; Zhang, J.; Dai, L.; Guo, J.; Wohlrab, S.; Greenwood-Goodwin, M.; Shen, A.; Duan, D.; Li, H.; Yuk, I. H. Versatile LC–MS-Based Workflow with Robust 0.1 Ppm Sensitivity for Identifying Residual HCPs in Biotherapeutic Products. *Anal. Chem.* **2022**, *94* (2), 723–731. https://doi.org/10.1021/acs.analchem.1c03095.

(9) Chen, I. H.; Xiao, H.; Daly, T.; Li, N. Improved Host Cell Protein Analysis in Monoclonal Antibody Products through Molecular Weight Cutoff Enrichment. *Anal. Chem.* **2020**, *92* (5), 3751–3757. https://doi.org/10.1021/acs.analchem.9b05081.

(10) Zhang, S.; Zhao, B.; Adaniya, S.; Xiao, H.; Li, N. Ultrasensitive Quantification Method for Understanding Biologically Relevant Concentrations of Host Cell Proteins in Therapeutics. *Anal. Chem.* **2023**, *95* (14), 6002–6008. https://doi.org/10.1021/acs.analchem.3c00020.

(11) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **2008**, *77* (1), 383–414. https://doi.org/10.1146/annurev.biochem.75.101304.124125.

(12) Tsukidate, T.; Li, Q.; Hang, H. C. Targeted and Proteome-Wide Analysis of Metabolite–Protein Interactions. *Curr. Opin. Chem. Biol.* **2020**, *54*, 19–27. https://doi.org/10.1016/j.cbpa.2019.10.008.

(13) Li, X.; Chandra, D.; Letarte, S.; Adam, G. C.; Welch, J.; Yang, R.-S.; Rivera, S.; Bodea, S.; Dow, A.; Chi, A.; Strulson, C. A.; Richardson, D. D. Profiling Active Enzymes for Polysorbate Degradation in Biotherapeutics by Activity-Based Protein Profiling. *Anal. Chem.* **2021**, *93* (23), 8161–8169. https://doi.org/10.1021/acs.analchem.1c00042.

(14) Zhang, S.; Xiao, H.; Molden, R.; Qiu, H.; Li, N. Rapid Polysorbate 80 Degradation by Liver Carboxylesterase in a Monoclonal Antibody Formulated Drug Substance at Early Stage Development. *J. Pharm. Sci.* **2020**, *109* (11), 3300–3307. https://doi.org/10.1016/j.xphs.2020.07.018.

(15) Liu, G.-Y.; Nie, S.; Zheng, X.; Li, N. Activity-Based Protein Profiling Probe for the Detection of Enzymes Catalyzing Polysorbate Degradation. *Anal. Chem.* **2022**, *94* (24), 8625–8632. https://doi.org/10.1021/acs.analchem.2c00059.

(16) Bache, N.; Geyer, P. E.; Bekker-Jensen, D. B.; Hoerning, O.; Falkenby, L.; Treit, P. V; Doll, S.; Paron, I.; Müller, J. B.; Meier, F.; Olsen, J. V; Vorm, O.; Mann, M. A Novel LC System Embeds Analytes in Pre-Formed Gradients for Rapid, Ultra-Robust Proteomics. *Mol. Cell. Proteomics* **2018**, *17* (11), 2284–2296. https://doi.org/10.1074/mcp.TIR118.000853.

(17) Sinitcyn, P.; Hamzeiy, H.; Salinas Soto, F.; Itzhak, D.; McCarthy, F.; Wichmann, C.; Steger, M.; Ohmayer, U.; Distler, U.; Kaspar-Schoenefeld, S.; Prianichnikov, N.; Yılmaz, Ş.; Rudolph, J. D.; Tenzer, S.; Perez-Riverol, Y.; Nagaraj, N.; Humphrey, S. J.; Cox, J. MaxDIA Enables Library-Based and Library-Free Data-Independent Acquisition Proteomics. *Nat. Biotechnol.* **2021**, *39* (12), 1563–1573. https://doi.org/10.1038/s41587-021-00968-7.

(18) Demichev, V.; Messner, C. B.; Vernardis, S. I.; Lilley, K. S.; Ralser, M. DIA-NN: Neural Networks and Interference Correction Enable Deep Proteome Coverage in High Throughput. *Nat. Methods* **2020**, *17* (1), 41–44. https://doi.org/10.1038/s41592-019-0638-x.

(19) Pham, T. V; Henneman, A. A.; Jimenez, C. R. Iq□: An R Package to Estimate Relative Protein Abundances from Ion Quantification in DIA-MS-Based Proteomics. *Bioinformatics* **2020**, *36* (8), 2611–2613. https://doi.org/10.1093/bioinformatics/btz961.

(20) Pino, L. K.; Searle, B. C.; Bollinger, J. G.; Nunn, B.; MacLean, B.; MacCoss, M. J. The Skyline Ecosystem: Informatics for Quantitative Mass Spectrometry Proteomics. *Mass Spectrometry Reviews*. 2020, pp 229–244. https://doi.org/10.1002/mas.21540.

(21) Kreimer, S.; Gao, Y.; Ray, S.; Jin, M.; Tan, Z.; Mussa, N. A.; Tao, L.; Li, Z.; Ivanov, A. R.; Karger, B. L. Host Cell Protein Profiling by Targeted and Untargeted Analysis of Data Independent Acquisition Mass Spectrometry Data with Parallel Reaction Monitoring Verification. *Anal. Chem.* **2017**, *89* (10), 5294–5302. https://doi.org/10.1021/acs.analchem.6b04892.

(22) Grömping, U. R Package FrF2 for Creating and Analyzing Fractional Factorial 2-Level Designs. *J. Stat. Softw.* **2014**, *56* (1), 1–56. https://doi.org/10.18637/jss.v056.i01.

(23) Lenth, R. V. Response-Surface Methods in R, Using RSM. *J. Stat. Softw.* **2009**, *32* (7), 1–17. https://doi.org/10.18637/jss.v032.i07.

(24) Tsukidate, T.; Stiving, A. Q.; Mengisen, S.; McKechnie, W. S.; Carrillo, R.; Li, X. Heat Inactivation of Host Cell–Derived Enzymes as a Control Strategy for Polysorbate Degradation. *J. Pharm. Sci.* **2024**, *113* (3), 637–641. https://doi.org/10.1016/j.xphs.2023.10.038.

(25) Bachovchin, D. A.; Ji, T.; Li, W.; Simon, G. M.; Blankman, J. L.; Adibekian, A.; Hoover, H.; Niessen, S.; Cravatt, B. F. Superfamily-Wide Portrait of Serine Hydrolase Inhibition Achieved by Library-versus-Library Screening. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (49), 20941–20946. https://doi.org/10.1073/pnas.1011663107.

(26) Baggelaar, M. P.; Den Dulk, H.; Florea, B. I.; Fazio, D.; Bernabò, N.; Raspa, M.; Janssen, A. P. A.; Scavizzi, F.; Barboni, B.; Overkleeft, H. S.; Maccarrone, M.; Van Der Stelt, M. ABHD2 Inhibitor Identified by Activity-Based Protein Profiling Reduces Acrosome Reaction. *ACS Chem. Biol.* **2019**, *14* (10), 2295–2304. https://doi.org/10.1021/acschembio.9b00640.

(27) Liu, Y.; Patricelli, M. P.; Cravatt, B. F. Activity-Based Protein Profiling: The Serine Hydrolases. *Proc. Natl. Acad. Sci.* **1999**, *96* (26), 14694–14699. https://doi.org/10.1073/pnas.96.26.14694.

(28) Dehghani, A.; Binder, F.; Zorn, M.; Feigler, A.; Fischer, K. I.; Felix, M. N.; Happersberger, P.; Reisinger, B. Investigating PH Effects on Enzymes Catalyzing Polysorbate Degradation by Activity-Based Protein Profiling. *J. Pharm. Sci.* **2023**, *000*, 1–10. https://doi.org/10.1016/j.xphs.2023.09.013.

(29) Long, J. Z.; Cravatt, B. F. The Metabolic Serine Hydrolases and Their Functions in Mammalian Physiology and Disease. *Chem. Rev.* **2011**, *111* (10), 6022–6063. https://doi.org/10.1021/cr200075y.

(30) Maier, M.; Weiß, L.; Zeh, N.; Schmieder-Todtenhaupt, V.; Dehghani, A.; Felix, M. N.; Heinzelmann, D.; Lindner, B.; Schmidt, M.; Studts, J.; Schulz, P.; Reisinger, B.; Otte, K.; Franzreb, M.; Lakatos, D.; Fischer, S. Illuminating a Biologics Development Challenge: Systematic Characterization of CHO Cell-Derived Hydrolases Identified in Monoclonal Antibody Formulations. *MAbs* **2024**, *16* (1). https://doi.org/10.1080/19420862.2024.2375798.

(31) Meier, F.; Brunner, A. D.; Koch, S.; Koch, H.; Lubeck, M.; Krause, M.; Goedecke, N.; Decker, J.; Kosinski, T.; Park, M. A.; Bache, N.; Hoerning, O.; Cox, J.; Räther, O.; Mann, M. Online Parallel Accumulation–Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. *Mol. Cell. Proteomics* **2018**, *17* (12), 2534–2545. https://doi.org/10.1074/mcp.TIR118.000900.

(32) Jones, H. B. L.; Heilig, R.; Fischer, R.; Kessler, B. M.; Pinto-Fernández, A. ABPP-HT - High-Throughput Activity-Based Profiling of Deubiquitylating Enzyme Inhibitors in a Cellular Context. *Front. Chem.* **2021**, *9* (February), 1–13. https://doi.org/10.3389/fchem.2021.640105.

(33) Mun, D. G.; Bhat, F. A.; Ding, H.; Madden, B. J.; Natesampillai, S.; Badley, A. D.; Johnson, K. L.; Kelly, R. T.; Pandey, A. Optimizing Single Cell Proteomics Using Trapped Ion Mobility Spectrometry for Label-Free Experiments. *Analyst* **2023**, *148* (15), 3466–3475. https://doi.org/10.1039/d3an00080j.

(34) Tüshaus, J.; Sakhteman, A.; Lechner, S.; The, M.; Mucha, E.; Krisp, C.; Schlegel, J.; Delbridge, C.; Kuster, B. A Region□resolved Proteomic Map of the Human Brain Enabled by High□throughput Proteomics. *EMBO J.* **2023**, *42* (23), 1–16. https://doi.org/10.15252/embj.2023114665.

(35) Sun, B.; Kovatch, J. R.; Badiong, A.; Merbouh, N. Optimization and Modeling of Quadrupole Orbitrap Parameters for Sensitive Analysis toward Single-Cell Proteomics. *J. Proteome Res.* **2017**, *16* (10), 3711–3721. https://doi.org/10.1021/acs.jproteome.7b00416.

(36) Meier, F.; Brunner, A. D.; Frank, M.; Ha, A.; Bludau, I.; Voytik, E.; Kaspar-Schoenefeld, S.; Lubeck, M.; Raether, O.; Bache, N.; Aebersold, R.; Collins, B. C.; Röst, H. L.; Mann, M. DiaPASEF: Parallel Accumulation–Serial Fragmentation Combined with Data-Independent Acquisition. *Nat. Methods* **2020**, *17* (12), 1229–1236. https://doi.org/10.1038/s41592-020-00998-0.

(37) Lesur, A.; Schmit, P. O.; Bernardin, F.; Letellier, E.; Brehmer, S.; Decker, J.; Dittmar, G. Highly Multiplexed Targeted Proteomics Acquisition on a TIMS-QTOF. *Anal. Chem.* **2021**, *93* (3), 1383–1392. https://doi.org/10.1021/acs.analchem.0c03180.

(38) Fröhlich, K.; Furrer, R.; Schori, C.; Handschin, C.; Schmidt, A. Robust, Precise, and Deep Proteome Profiling Using a Small Mass Range and Narrow Window Data-Independent-Acquisition Scheme. *J. Proteome Res.* **2024**, *23* (3), 1028–1038. https://doi.org/10.1021/acs.jproteome.3c00736.

(39) Szyrwiel, L.; Sinn, L.; Ralser, M.; Demichev, V. Slice-PASEF: Fragmenting All Ions for Maximum Sensitivity in Proteomics. *bioRxiv* **2022**, 2022.10.31.514544. https://doi.org/https://doi.org/10.1101/2022.10.31.514544.

(40) Brzhozovskiy, A.; Kononikhin, A.; Bugrova, A. E.; Kovalev, G. I.; Schmit, P. O.; Kruppa, G.; Nikolaev, E. N.; Borchers, C. H. The Parallel Reaction Monitoring-Parallel Accumulation–Serial Fragmentation (Prm-PASEF) Approach for Multiplexed Absolute Quantitation of Proteins in Human Plasma. *Anal. Chem.* **2022**, *94* (4), 2016–2022. https://doi.org/10.1021/acs.analchem.1c03782.

(41) Kovner, D.; Yuk, I. H.; Shen, A.; Li, H.; Graf, T.; Gupta, S.; Liu, W.; Tomlinson, A. Characterization of Recombinantly-Expressed Hydrolytic Enzymes from Chinese Hamster Ovary Cells: Identification of Host Cell Proteins That Degrade Polysorbate. *J. Pharm. Sci.* **2023**, *112* (5), 1351–1363. https://doi.org/10.1016/j.xphs.2023.01.003.

(42) Zhang, S.; Riccardi, C.; Kamen, D.; Reilly, J.; Mattila, J.; Bak, H.; Xiao, H.; Li, N. Identification of the Specific Causes of Polysorbate 20 Degradation in Monoclonal Antibody Formulations Containing Multiple Lipases. *Pharm. Res.* **2022**, *39* (1), 75–87. https://doi.org/10.1007/s11095-021-03160-3.

(43) Zhang, S.; Xiao, H.; Li, N. Degradation of Polysorbate 20 by Sialate O-Acetylesterase in Monoclonal Antibody Formulations. *J. Pharm. Sci.* **2021**, *110* (12), 3866–3873. https://doi.org/10.1016/j.xphs.2021.09.001.

Authors are required to submit a graphic entry for the Table of Contents (TOC) that, in conjunction with the manuscript title, should give the reader a representative idea of one of the following: A key structure, reaction, equation, concept, or theorem, etc., that is discussed in the manuscript. Consult the journal's Instructions for Authors for TOC graphic specifications.

Insert Table of Contents artwork here