# Advanced mass spectrometry workflows for accurate quantification of trace-level host cell proteins in drug products: Benefits of FAIMS separation and gas-phase fractionation DIA

Corentin Beaumal, Alain Beck, Oscar Hernandez-Alba, Christine Carapito

▶ **To cite this version:**

Corentin Beaumal, Alain Beck, Oscar Hernandez-Alba, Christine Carapito. Advanced mass spectrometry workflows for accurate quantification of trace-level host cell proteins in drug products: Benefits of FAIMS separation and gas-phase fractionation DIA. Proteomics, 2023, 23 (16), 10.1002/pmic.202300172 . hal-04296423

**HAL Id: hal-04296423**
**https://hal.science/hal-04296423**

Submitted on 20 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Advanced mass spectrometry workflows for accurate quantification of trace-level host cell proteins in drug products: benefits of FAIMS separation and gas-phase fractionation DIA

**Authors**

Corentin Beaumal[1,2], Alain Beck[3], Oscar Hernandez-Alba[1,2], Christine Carapito[1,2]

[1] Laboratoire de Spectrométrie de Masse BioOrganique, IPHC UMR 7178, CNRS, Université de Strasbourg, F-67087 Strasbourg, France.

[2] Infrastructure Nationale de Protéomique ProFI - FR2048, F-67087 Strasbourg, France.

[3] IRPF, Centre d'Immunologie Pierre-Fabre (CIPF), F-74160 Saint-Julien-en-Genevois, France.

Corresponding author:

Dr Christine Carapito – Laboratoire de Spectrométrie de Masse BioOrganique, IPHC UMR 7178, CNRS, Université de Strasbourg, 25 rue Becquerel, F-67087 Strasbourg, France

Email : ccarapito@unistra.fr

**Abbreviations**

CQA, critical quality attribute; CV, compensation voltage; DDA, data-dependent acquisition; DIA, data-independent acquisition; DP, drug product; FAIMS, high-field asymmetric waveform ion mobility spectrometry; GPF, gas phase fractionation; HCP, host cell proteins; NIST, National Institute of Standards and Technology

**Keywords**

Data-Independent Acquisition Mass Spectrometry (DIA-MS); High-Field Asymmetric Ion Mobility Spectrometry (FAIMS); Drug Products (DP); Host Cell Protein (HCP) Impurities.

**Total number of words:** 8,536 words

**Abstract**

Therapeutic monoclonal antibodies (mAb) production relies on multiple purification steps before release as a drug product (DP). A few host cell proteins (HCPs) may co-purify with the mAb. Their monitoring is crucial due to the considerable risk they represent for mAb stability, integrity and efficacy and their potential immunogenicity. Enzyme-linked immunosorbent assays (ELISA) commonly used for global HCP monitoring present limitations in terms of identification and quantification of individual HCPs. Therefore, liquid chromatography tandem mass spectrometry (LC-MS/MS) has emerged as a promising alternative. Challenging DP samples show an extreme dynamic range requiring high performing methods to detect and reliably quantify trace-level HCPs. Here, we investigated the benefits of adding high-field asymmetric ion mobility spectrometry (FAIMS) separation and gas phase fractionation (GPF) prior to data independent acquisition (DIA). FAIMS LC-MS/MS analysis allowed the identification of 221 HCPs among which 158 were reliably quantified for a global amount of 880 ng/mg of NIST mAb Reference Material. Our methods have also been successfully applied to two FDA/EMA approved DPs and allowed digging deeper into the HCP landscape with the identification and quantification of a few tens of HCPs with sensitivity down to the sub-ng/mg of mAb level.

**Significance statement**

Host cell protein (HCP) impurities monitoring in biotherapeutics is a critical concern because of their potential adverse actions. In this study, we developed innovative mass spectrometry (MS)-based approaches on a tribrid instrument to overcome the limitations of standard methods for the characterisation of trace level HCP impurities in drug products. We demonstrate the benefits of implementing a High-Field Asymmetric Ion Mobility Spectrometry (FAIMS) separation step in the workflow as well as of the use of Data Independent Acquisition (DIA) combined with a Gas-Phase Fractionation approach to address the challenges of accurate and precise quantification of very low-level HCPs in samples presenting an extreme dynamic range. Major improvements towards robust quantification of trace-level HCPs in Drug Products have been demonstrated both on the standard NIST mAb Reference Material and two FDA/EMA approved drug products, down to the sub-ppm level.

## 1 Introduction

Since the approval of the first monoclonal antibody (mAb) in 1986 [1], the Muromonab-CD3, an immunoglobulin reducing acute rejection in patients with organ transplants, interest for this class of biopharmaceutical products has been growing fast [2,3]. Today, the United States Food and Drug Administration (FDA) and/or the European Medicines Agency (EMA) have accepted more than 125 mAbs and about twenty are currently under review by these regulatory agencies [4,5]. Production of recombinant therapeutic proteins using expression cell lines technology results in the presence of bioprocess-related host cell protein (HCP) impurities in addition to the desired mAb. During the downstream process (*i.e.* Protein A affinity chromatography, ion exchange chromatography, hydrophobic interaction chromatography and multiple filtration steps), the total amount of HCPs and the mixture complexity are drastically reduced [6,7]. However, low levels of HCPs may co-purify with the therapeutic protein and thus remain in the drug products (DPs). Some of these residual HCPs may degrade the drug or excipients in the DP, potentially leading to reduced stability or inactivation of the mAb [8-11], while other impurities may endanger patients' safety by causing an unexpected and damaging immune response weakening the therapeutic protein efficiency [10,12-14]. The potential risks associated with HCPs have conducted the regulatory agencies to consider them as a critical quality attribute (CQA) in DPs, and their total amount should be minimized to trace levels and monitored by highly sensitive analytical methods during downstream process [15-17]. Even if general guidelines indicate that there is no strict limit on the HCPs level in DPs, many biopharmaceutical companies consider that their global amount should not exceed 100 ppm (100 ng/mg of mAb) in the DP [18-20]. Besides, it cannot be excluded that the presence of a specific HCP, even at a very low level, could jeopardize the action of the therapeutic protein [21]. Currently, enzyme-linked immunosorbent assays (ELISA) are commonly used for HCP impurities analysis for their easy-of-use and high precision [22-24].

However, ELISA has some limitations: certain HCP species are not targeted by the assay and thus cannot be detected and only global quantification is possible, with neither individual HCP identification nor quantification possible [7,22,25]. In addition, the development of fit for purpose product-specific immunoassays to measure HCPs in DP is a long, expensive and fastidious process, as they must satisfy numerous criteria (accuracy, precision, range, linearity, etc.) to be validated[26].

Therefore, liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) has emerged as a promising orthogonal and complementary method to characterize HCP impurities over the bioprocess and ultimately in DPs [27-29]. Multiple studies point out LC-MS/MS benefits for HCP monitoring, using targeted approaches (selected reaction monitoring (SRM) and parallel reaction monitoring (PRM)) [30-34], data-dependent acquisition (DDA) [35-37] and data-independent acquisition (DIA) [33,38,39]. DIA presents the advantage of allowing the extraction of quantitative information on all detectable species through MS2 data [40]. Until recently, the use of spectral-libraries and so-called "peptide centric" approaches was the most efficient way to interpret DIA data, while time and sample consuming due to the necessity to generate the most comprehensive spectral library possible in DDA prior to DIA analysis [41,42]. Today, "spectrum centric" approaches are growing in interest, thanks to continuous improvements of acquisition strategies [43-46] and the development of innovative tailored tools, allowing identification and precise quantification without spectral libraries [47-50] or using artificial intelligence (AI) *in silico*-predicted libraries [51-53]. While the dynamic range achievable by latest generation LC-MS/MS couplings reaches 3 to 4 orders of magnitude, quantifying ppm-level HCP in DPs by definition requests 6 orders of dynamics. To achieve this, a few studies have focused on optimizing sample preparation via mAb depletion using native digestion [54] or reversely via HCPs enrichment using several approaches: molecular weight cutoff enrichment [55], protein A affinity chromatography [56,57] or hydrophilic interaction chromatography (HILIC) [58]. Other studies relied on improvements of the LC-MS/MS part including, among others, the multiple chromatography separations [31,37,59,60] or implementing an ion mobility separation step [30,61,62].

In this study, we focused on the most challenging samples, namely DPs. We have evaluated several LC-MS/MS settings for their potential to identify the most exhaustive list of HCPs but also to reliably quantify the largest possible fraction of them. First, we analyzed the NIST mAb standard reference material using DDA and DIA acquisition methods, leading to the identification of hundreds of HCP impurities and the accurate quantification of tens of them. Then, the high field asymmetric ion mobility spectrometry (FAIMS) device allowing gas-phase separation of co-eluting peptides through fast internal compensation voltages (CV) switching, was implemented in front of the mass spectrometer to evaluate its benefits in terms of depth and coverage of HCP impurities detection. In parallel, a gas-phase fractionation DIA strategy was evaluated. Both workflows were benchmarked against results described in the literature. Finally, we applied our methods to monitor HCPs in two FDA/EMA approved mAbs with a very high medical benefits in oncology, trastuzumab and nivolumab, enabling the identification and quantification of several tens of HCPs and demonstrating that these approaches can be used to routinely analyze DPs in a rapid and accurate way.

## 2 Materials and methods

### *2.1 Materials*

The NIST mAb standard reference material RM8671, dithiothreitol (DTT) and Tris-HCl were purchased from Merck. Commercially available mAbs, trastuzumab and nivolumab, were obtained as European Union pharmaceutical-grade drug products from the Pierre Fabre Immunologic Center (CIPF, Saint-Julien-en-Genevois, France). HCP-PROFILER standard peptides were acquired from Anaquant (Villeurbanne, France). Mass spectrometry grade trypsin/Lys-C enzymes were obtained from Promega (Madison, WI, USA).

### *2.2 Sample preparation*

The digestion protocol was adapted from Huang *et al.* [54]. Briefly, a 100 µL aliquot of 10 µg/µL mAb was supplemented with 90 µL of ultrapure $H_2O$ and 5 µL of 1 M Tris-HCl, pH 8. Proteins were digested overnight at 37°C using a solution of Trypsin/Lys C enzymes at a 1:400 enzyme/protein ratio. Digests were reduced with 1.5 µL of 303 mM DTT for 10 min at 90°C and centrifuged at 13 000g for 2 min. Supernatant was finally acidified with 0.5 µL of formic acid (FA). Then, after *vacuum* drying, peptides were solubilized in 2% acetonitrile (ACN) and 0.1% FA to obtain a final protein concentration of 1 µg/µL. For gas-phase fractionation (GPF) chromatogram library generation, technical replicates were pooled together. One HCP-PROFILER bead was dissolved in 150 µL of 5% ACN and 0.1% FA, to obtain standard peptides in a concentration range from 0.5 to 250 fmol/µL. Retention time iRT standards (Biognosys, Schlieren, Switzerland) and 2 µL of HCP-PROFILER solution were mixed to 1 or 2 µL of FDA/EMA approved mAb or NIST mAb Reference Material, respectively, before injection.

### *2.3 Nano LC-MS/MS analysis*

Data-dependent acquisition (DDA) and data-independent acquisition (DIA) analyses were performed on a Dionex UltiMate 3000 RSLC nano system (Thermo Scientific) coupled to an Orbitrap Eclipse™ Tribrid™ (Thermo Scientific) mass spectrometer, equipped with a FAIMS Pro interface (Thermo Scientific). Mobile phase A contained 0.1% FA in water and mobile phase B contained 0.1% FA in 80% ACN/20% water. Peptides were loaded onto an Acclaim PepMap 100 C18 20 mm x 0.1 mm, 5 µm diameter particles precolumn (Thermo Scientific) for 3 min at 10 µL/min, and eluted on an Aurora Series C18 UHPLC (250 mm x 75µm, 1.6 µm diameter particles, IonOpticks) at a 300 nL/min flow rate following a linear gradient: 2.5% B at 0 min, 43.8% B at 95 min and 98% B from 96 to 101 min. The column was finally re-equilibrated in 2.5% B for 15 min.

For DDA analyses, full-scan MS spectra were acquired over a 375-1,500 m/z range at a resolution of 60,000 (at 200 m/z), with an automatic gain control (AGC) target of $1.10^6$ and a maximum injection time of 60 ms. The top 10 most intense precursor ions with an intensity exceeding $5.10^3$ ions per second and charge states between 2 and 6 were automatically selected from each MS spectrum for higher-energy collisional dissociation (HCD) fragmentation at 30% normalized collision energy. MS/MS spectra were collected from 120-1200 m/z in the ion trap in Rapid mode, with an AGC target of $1.10^4$ and a maximum injection time set to Automatic. Experiments with the FAIMS Pro interface were conducted as described in Hebert *et al.* [62]. Briefly, FAIMS carrier gas flow was 4.7 L/min $N_2$, the FAIMS electrodes were set to 100°C, entrance plate voltage was 250 V, asymmetric waveform divert voltage was – 5000 V, compensation voltage (CV) settling time was 25 ms and CVs values of -50, -65 and -85 V were applied successively in the same run.

For DIA analyses, full-scan MS was acquired over a 375-1250 m/z range at a resolution of 60,000 (at 200 *m/z*) with an AGC target of $1.10^6$ and a maximum injection time of 60 ms. Fragment analysis (MS/MS) was subdivided into 42 windows of variable widths, described in Supporting Table S1. A resolution of 15,000 (at 200 m/z), an AGC target of $1.10^6$ and a maximum injection time of 22 ms were used for a total cycle time of 1.5 s. HCD fragmentation normalized collision energy was set to

30%. For DIA-GPF, 6 individual injections were performed, covering the 380-980 m/z mass range. A 100 m/z full-scan MS was acquired with the parameters described above for DIA, and the MS/MS analysis was divided into 25 windows of 4 m/z using a 30,000 (at 200 m/z) resolution, an AGC target of $1.10^6$ and maximum injection time of 60 ms.

*2.4 DDA data analysis*

DDA data analysis was performed with Proteome Discoverer v. 2.5 (Thermo Fisher Scientific) using Sequest HT (Thermo Fisher Scientific) against a database containing all *Mus musculus* entries (17,050 entries, TaxID=10090, 2021/05/26) extracted from UniProtKB/SwissProt for NIST mAb samples and a database containing all *Cricetulus Griseus* entries (78,366 entries, Tax ID=10029, 2022/03/15) extracted from UniProtKB/TrEMBL for trastuzumab and nivolumab samples. Each fasta file included common contaminants, iRT retention time standards, HCP-PROFILER standard sequences and mAbs heavy and light chain sequences. Trypsin/P was used as digestion enzyme and one missed cleavage was allowed. Methionine oxidation and protein N-term acetylation were set as variable modifications. Mass tolerance was set to 5 ppm for precursor ion masses and 0.5 Da for fragment ion masses analyzed in the ion trap. The false discovery rate (FDR) was calculated using the Percolator node set to 1% at PSM, peptide and protein levels. XIC-MS1 quantification was performed using unique peptides only and the chromatographic alignment was set to 10 min and 10 ppm between the 3 technical replicates.

*2.5 DIA data analysis*

DIA data analysis was performed with Spectronaut v. 15.7 (Biognosys) using the fasta files described above and the following settings. Briefly, trypsin/P was set as digestion enzyme, methionine oxidation and protein N-term acetylation were set as variable modifications and one missed cleavage was allowed. Data was extracted using dynamic mass tolerances and a 10 min XIC RT extraction window. Identification was performed using 0.01 precursor and protein Qvalue cutoffs. Quantification was performed using interference correction and at least three fragments per

precursor, without imputation. Quantity is based on $MS^2$ XIC peak areas. Non-identified precursors in rows with at least one q-value below 0.01 were selected for iRT profiling, by enabling carrying over the average template peak position.

For GPF data, a chromatogram library was generated in Spectronaut (v. 15.7). The six raw files were processed altogether with the Pulsar search engine with similar parameters than the ones used for DIA analysis: trypsin/P as digestion enzyme, methionine oxidation and protein N-term acetylation were as variable modifications, one missed cleavage authorized. Additional filters were applied for library generation, including a 1% FDR threshold at PSM, peptide and protein levels and a number of fragments between 3 and 8 per precursor. Then, data acquired in DIA were searched against the chromatogram library with the same parameters as for library-free DIA data analysis described above.

*2.6 HCP quantification*

Validation filters described in Pythoud *et al.*[39] were applied prior to HCP quantification. Briefly, oxidized and acetylated precursors and their counterparts were removed and only host organism or standard protein precursors with charge states of two or three were kept. For DDA analysis, shared and not quantified precursors as well as precursors with rank > 1 and with |RT Sequest – RT Top Apex| >4 min were removed. One missing value and/or one *by matching* value at maximum were authorized over the three replicates and a CV on the triplicate precursor intensity values below 20% must be achieved to consider the precursor for quantification. Finally, a sequence homology filter was applied to remove HCP peptides that can potentially arise from unspecific cleavage or mAb degradation.

After filtering, peptide intensity was obtained by summing all precursor ion intensities and protein intensity by summing the three most intense peptides, based on the Top3 strategy described by Silva *et al.* [63]. Protein mole quantities were estimated using the equation of the HCP Profiler calibration

curve. Individual HCP ng/mg mAb (ppm) amounts were estimated using the molecular weight and injected mAb quantities.

## 3 Results and discussion

### *3.1 LC-MS/MS methods to precisely and reliably monitor trace-level HCPs*

Three technical replicates of NIST mAb reference material were prepared and analyzed in DDA mode on a nanoLC-Orbitrap Tribrid coupling, with fragment detection (MS2) performed in the ion trap. These analyses led to the detection and identification of 325 peptides originating from 133 distinct HCP protein groups (Figure 1.A). We compared the list of protein impurities identified in our study to the lists of HCPs identified and/or quantified in recently published papers from literature, applying comparable sample preparation (based on Huang native digestion protocol) and DDA LC-MS/MS acquisition methods [36,54,64]. The results obtained point out a fairly good overlap between the studies, with 43 proteins (32 % of the proteins detected in our study) identified in common in all four studies, and 21 additional HCPs (16% of the proteins detected in our study) shared by at least two papers and our results (Figure 1.B). Overall, this demonstrates the reproducibility of the native digestion protocol, independently of the LC-MS/MS system used and the location of experiments. This is a critical point regarding the perspective of transferring this protocol in a biopharmaceutical environment. Additionally, Figure 1.B also highlights that we have uniquely identified 54 HCPs thus slightly outperforming the other methods. Then, quantification of these trace-level proteins was performed on the precursor ion intensities extracted at MS1 level. To ensure accurate and robust quantification of the impurities detected, stringent filters based on the signal quality and repeatability were applied as described in Pythoud *et al.* [39]. Briefly, only 2+ and 3+ precursors were

kept, shared peptides, peptides without any quantification value and modified peptides (*i.e.* carrying a methionine oxidation or acetylation) were removed, as well as precursors with a difference between their apex RT and their elution RT higher than 4 min. Then, peptides with missing values in more than one of the three replicates and peptides identified using match between run in more than one over three replicates were removed. Finally, a last filter on the signal repeatability was applied, based on the removal of peptides with a coefficient of variation higher than 20% on measured intensities over the three replicates. Ultimately, to prevent any misidentification of the remaining peptides, a sequence homology filter was applied to remove any peptides that could potentially originate from unspecific mAb cleavage or degradation. Altogether, these stringent filters reduced the number of peptides by 63% (from 325 identified to 121 quantified) and protein groups by 49% (from 133 identified to 68 quantified), as illustrated in Figure 1.A. This decrease between identified and quantified features demonstrates that a major part of the identified HCPs is detected at trace-levels, between the detection and quantification limits. The good overlap between our 68 quantified HCPs and previously cited papers with 65% also found in two or three other papers (29 HCPs founds in our study and the three papers, 14 in our study and two of the three papers, Supporting Figure S1) further strengthens the confidence and reliability on the reported results. Accurate and reproducible characterization and quantification of trace-level HCPs in matrices showing a very high dynamic range with a superabundant mAb such as DPs needs and deserves the use of serious and stringent analytical filters for high-confidence results. Thus, only individual peptides that passed the filters from the replicate associated HCP-PROFILER standard internal calibration curve were calculated. Trauchessec *et al.* [65] already described the internal calibration benefits for quantification reproducibility and robustness (example of calibration curve in Supporting Figure S2). Subsequently, protein group quantities were retrieved based on the Top3 approach, [63] affording an average total amount of HCP impurities of 518 ± 20 ng per mg of mAb or ppm. The quantification accuracy of our results was evaluated against published data in the literature. Figure 1.C represents the correlation between the quantities obtained for each HCP in common to our results and selected papers [54,64].

The Pearson coefficients obtained are 0.97 and 0.99 for the Molden *et al.* and Huang *et al.* papers respectively, demonstrating the strong correlation of the absolute quantities calculated. Altogether these results clearly show that advanced LC-MS/MS workflows are perfectly suited to reliably monitor trace-level HCPs in DPs and thus eventually become the method of choice for quality control of released DPs.

*3.2 Digging deeper into the HCP landscape with the implementation of an additional FAIMS separation*

We explored high-field asymmetric waveform ion mobility spectrometry (FAIMS) benefits for detection of trace-level HCPs in DPs. Indeed, while the native digestion protocol enables the removal of a major part (estimated between 70% and 80% depletion, data not shown) of the mAb and thus reduces the dynamic range in the samples, mAb-specific peptides are still being detected and abundant in comparison to HCP impurities. The use of compensation voltages applied to FAIMS electrodes enables a better resolution of co-eluting peptides, based on the precursor ions gas phase mobility and charge, by filtering the ion population entering the mass spectrometer. To investigate the potential of applying multiple compensation voltages in our LC-MS/MS workflows to separate HCP peptides from co-eluting mAb peptides, the three technical replicates previously analyzed using nanoLC-DDA analysis were analyzed using an optimized FAIMS method. Chromatograms observed without FAIMS and with different FAIMS compensation voltages applied during the experiment are provided in Supporting Figure S3 and clearly illustrate the benefits of ion mobility to clean the signal by removing highly abundant peptides depending on the CV applied (Supporting Figure S3.A). Indeed, as an example, co-elution of the mAb heavy chain peptide STSGGTAALGCLVK and the mouse PPAC (low molecular weight phosphotyrosine protein phosphatase, accession number: Q9D358) HCP peptide SPIAEAVFR prevented the identification of peptide SPIAEAVFR without FAIMS due to the wide dynamic range between those two peptides. On the contrary, when a CV of -85V was applied, a clean signal for peptide SPIAEAVFR allowed its identification and quantification while the mAb heavy

chain peptide STSGGTAALGCLVK was filtered out at this CV value (Supporting Figure S3.B and C). Overall, the use of the optimized FAIMS method lead to an increase of 49% (from 325 to 484) and 66% (from 133 to 221) of identified peptides and protein groups, respectively (Figure 1.D). The validation filters earlier described were applied to quantify HCPs. Figure 1.D illustrates that signal quality and repeatability filters resulted in a drop of 46% (from 484 to 262) and 28% (from 221 to 158) between identified and quantified peptides and protein groups, respectively. Of note is that the latter decrease in identified and quantified peptides with FAIMS is less significant compared to the DDA workflow without the ion mobility dimension in the front-end of the MS (63% and 49%, respectively). The signal repeatability filtering step (coefficient of variation < 20% filter) is the one excluding the highest numbers of peptides and protein groups. Again, the drop is much more pronounced in DDA (26% for proteins and 38% for peptides) compared to FAIMS DDA (10% for proteins and 24% for peptides), demonstrating that data acquired with the FAIMS interface seem to be overall more reproducible. Furthermore, FAIMS addition allowed increasing the number of both peptides and protein groups quantified by more than twofold (from 121 to 262, and from 68 to 158, respectively). Among the 158 quantified protein groups, 53 had also been quantified in DDA without FAIMS, representing 78% of the 68 proteins quantified with DDA only. This significant gain in terms of quantified HCPs logically results in 70% increase of total HCP amount quantified in the NIST mAb, from 518 ± 20 ppm to 880 ± 15 ppm without and with the FAIMS, respectively, thus demonstrating the real asset of adding FAIMS in the workflow. To further describe the benefits of implementing FAIMS, we compared in detail the amounts of each HCPs quantified either with or without the FAIMS. The quantities derived by both methods highly correlate with a Pearson coefficient of 0.97. Moreover, the amount of each individual protein was investigated to show the benefits of FAIMS to afford a more detailed characterization of the trace-level HCP landscape. Figure 2.A represents the density of HCP protein groups as a function of the concentration range. It points out the contribution of FAIMS separation in the quantification of low-abundant impurities since more than 75% of HCPs quantified (82 out of 105) exhibited concentrations lower than 2-ppm. Considering these results,

FAIMS appeared to be very useful to increase the number of quantified HCP, thanks to dynamic range and sensitivity enhancements. The removal of singly charged peptides and better separation of some highly abundant mAb peptides improved the signal to noise ratio, and consequently pushing the detection and quantification limits of the instrument for remaining HCPs. Notwithstanding, the CV values applied also influenced the HCP peptides, resulting in a slight decrease in the average number of peptides per protein, from 1.8 without the FAIMS to 1.6 with the FAIMS. As a conclusion, the implementation of FAIMS in the workflow appears to be highly beneficial to allow digging deeper into the HCP landscape, with quantification of up to 158 protein impurities down to sub-ppm level after stringent validation filters.

*3.3 Digging deeper into the HCP landscape with data independent acquisition supplemented with gas phase fractionation*

We [38,39] and others [33,66] have previously demonstrated the benefits of using DIA methods and MS2-based quantification to accurately monitor HCP impurities, and it is clearly confirmed in the current study. While DDA allowed the identification of the highest number of protein groups, namely 135, against 115 in DIA (Figure 1. D), DIA allowed quantifying 100 HCPs (222 peptides) against 68 in DDA (121 peptides). This can be explained by the stochasticity of DDA causing missing values and CV filtering steps to remove more peptides and proteins. Contrary to DDA, the high reproducibility of DIA, with the systematic fragmentation of all peptides included in a defined mass range, significantly reduces the rate of missing values and the abundance variability. DIA increased by 45% and 83%, respectively, the number of protein groups and peptides quantified, compared to DDA. Consequently, an average of 2.2 peptides per protein was determined using DIA against 1.8 using DDA strengthening the confidence of the identifications. Fifty out of the 68 (74%) HCPs quantified in DDA are also quantified in DIA, representing a good overlap between both approaches (Supporting Table S3). The retrieved quantities of these 50 common proteins also attest the strong correlation between DDA and DIA, with a Pearson coefficient of 0.96. Overall, total HCP amounts retrieved by

both acquisition modes are comparable, with a total HCP amount of 518 ± 20 ppm in DDA and 465 ± 28 ppm in DIA.

As an alternative to sample-fractionation to generate comprehensive spectral libraries, gas-phase fractionation DIA using chromatogram libraries can be applied with significantly reduced sample consumption and analysis time [46]. To evaluate this approach, we have performed six runs of 100 m/z mass range, from 480 to 1080 m/z, to generate a chromatogram library containing 844 precursors from 226 protein groups. Figure 1.D displays the results achieved by processing DIA data previously acquired using this chromatogram library: 526 peptides and 172 protein groups, corresponding to a 59% and 50% increase respectively, compared to the DIA "spectrum-centric" data interpretation. After applying our stringent validation filtering, we quantified 119 HCP protein groups and 248 peptides. In comparison to library-free processing, chromatogram library DIA data processing enables to quantify 20% more protein groups. The total amount of quantified HCPs by GPF was 457 ± 20 ppm, a very similar value compared to the library-free DIA strategy. Furthermore, 88 HCPs were simultaneously identified and quantified using both DIA approaches, representing 88% and 73% of the proteins quantified using the library-free and chromatogram library based approaches, respectively. In addition to this high overlap, the Pearson coefficient of 0.99 also highlights the strong correlation between the two strategies. This result demonstrates that chromatogram library based approach increases the number of quantified proteins, without impairing the individual quantification of the proteins, as evidenced by the similar total HCP amounts reported with both methods. On top of that, we investigated the capabilities of the GPF DIA approach to quantify low-level proteins and digging deeper in the HCP landscape. By looking at the HCPs only quantified using the GPF DIA approach, we highlighted that these proteins are mostly under the ppm level, as depicted in Figure 2.B. The shape of the violin plot clearly underlines that the density of HCPs is more important for low concentration levels using GPF DIA compared to DIA, mainly because of the presence of proteins quantified only with this approach. These elements clearly highlight that the use of GPF DIA allows increasing the number of identified and quantified HCPs, without the current

drawbacks of generating exhaustive spectral library from DDA data after fractionation, requiring huge amounts of sample along with time-consuming sample preparation, and data acquisition of fractionated samples. In addition, FAIMS implementation in DIA can be performed, but data interpretation is still the limiting step. In summary, DIA-based strategies appear to improve classical DDA, by increasing the number of identified and quantified proteins and digging deeper into the HCP landscape.

*3.4 Combined coverage achieved on the NIST and application of the methods to FDA/EMA approved DPs*

The merging of the results obtained using the four methods evaluated in previous sections (*i.e.* DDA, FAIMS DDA, DIA and GPF DIA) allows reaching a total of 278 HCP protein groups identified in the NIST mAb Reference Material , among which 80 are common to the four methods. The list of identified and quantified HCPs by each method is available in Supporting Table S2. We matched these identification results with those obtained in two recent papers showing high numbers of identified HCPs thanks to the combination of innovative technologies, including protein A depletion, native digestion, FAIMS and 2h30 gradient in Johnson *et al.* (2020) [61] or ultralow trypsin concentration, long column and gradient (50 cm, 4 h) and BoxCar acquisition in Nie *et al.* (2021) [67]. We discovered 213 common HCPs with the literature (*i.e.* Johnson *et al.* and Nie *et al.*), representing 77% of the 278 proteins we identified in this work, and we uniquely identified 32 proteins (Supporting Figure S4).

The optimizations performed on the NIST mAb Reference Material demonstrate the enhanced performances achieved when implementing innovative methods such as FAIMS DDA or GPF DIA compared to classical DDA and DIA. However, the NIST mAb has not been designed for clinical use and its purification process is not at the level of the ones employed for FDA/EMA approved DPs. Therefore, we analyzed two high-value approved DPs, trastuzumab and nivolumab, with our advanced MS methods (Figure 3.A and B, peptides results are presented in Supporting Figure S5). Overall, our results confirm the trends observed with the NIST mAb, with an increase of identified

features using the FAIMS DDA and GPF DIA compared to DDA and DIA, respectively. Indeed, 67 HCPs in trastuzumab and 60 HCPs in nivolumab using DDA with FAIMS were detected, while only 37 and 18 HCP impurities could be reported without FAIMS. In terms of quantification, 31 and 29 protein groups in trastuzumab and nivolumab, respectively, were successfully quantified with our FAIMS DDA method. Thus, FAIMS hyphenation in DDA increased by threefold the number of quantified trace level proteins compared to DDA without FAIMS in trastuzumab (from 10 to 31). Similarly, the use of GPF quantified 4 more HCPs, from 8 in DIA to 12 in GPF DIA.

These few proteins quantified account for a total HCP amount between 6 and 57 ppm, depending on the methods. Briefly, quantities are around 10 ppm in trastuzumab and nivolumab using DDA, DIA and GPF DIA methods, and up to 57 ppm using DDA FAIMS (Figure 3.C, Supporting Table S4 and S5). In a previous study by Pythoud *et al.* (2021) [39], comparable numbers and quantities of HCPs were measured for nivolumab and trastuzumab in both DDA and DIA mode, with 3 and 10 HCPs quantified representing a total HCP amount between 20 and 80 ppm. Additionally, a study by Molden *et al.* [64] about HCPs profiling in various mAb-based therapeutic proteins, including nivolumab and trastuzumab, reported comparable results. Indeed, among the 28 therapeutic proteins analyzed, 7 HCPs on average were identified for an average total amount of 20 ppm in each DP. These findings are consistent with the current results, confirming the low number and abundance of HCP impurities in approved DPs and being in line with the current target of many biopharmaceuticals companies and authorities recommendations, with admitted quantities of HCPs in DP tolerated between 1 and 100 ppm [18]. Our work highlights the capabilities of LC-MS/MS to quantify routinely and accurately individual HCPs in challenging FDA/EMA approved DPs demonstrating trace-level impurities. Based on their individual identification and quantification, investigations on the potential immunogenic effect of the monitored HCPs can be undertaken.

## 4 Concluding remarks

To summarize, we compared LC-MS/MS acquisition methods to characterize and monitor HCPs in DPs and achieved individual identification and quantification of hundreds of these impurities. First, using DDA analysis on the NIST mAb Reference Material, we proved the extreme necessity to apply stringent validation filtering steps in order to precisely and accurately quantify individual trace-level HCPs, by removing low-confidence and non-reproducible signals. Then, we evaluated optimized LC-MS/MS acquisition methods including DDA with and without FAIMS implementation, DIA, and gas-phase fractionation DIA. The benchmarking of DIA against DDA confirmed its benefits for HCPs monitoring in DP, with an increase of 47% in the number of impurities quantified. The addition of the chromatogram library-based approach for DIA data analysis also revealed very promising results, allowing quantification of 119 HCPs. Finally, the FAIMS implementation in DDA provided enhanced capabilities with the quantification of 158 protein impurities for a total amount of 880 ppm. Finally, we applied the aforementioned methods to trastuzumab and nivolumab, two FDA/EMA approved DPs known and presumably highly-pure products. FAIMS implementation to DDA reveals to be a powerful approach, with the quantification of around 30 HCPs for a total amount around 55 ppm in

both trastuzumab and nivolumab. This work demonstrated that LC-MS/MS based approaches followed by stringent validation filters can be employed for precise individual HCP impurities characterization in very challenging DPs.

**Associated Data**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

The dataset was deposited with the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD039582 (for DIA data) and PXD039585 (for DDA data). [68].

**Acknowledgments**

This work was supported by the “Agence Nationale de la Recherche” via the French Proteomic Infrastructure (ProFI FR2048; ANR-10-INBS-08-03) and by the French Ministry of Higher Education and Research for the PhD fellowship of C.B.

**Conflict of interest statement**

The authors have declared no conflict of interest.

**5 References**

[1] Reichert, J. M. (2012). Marketed therapeutic antibodies compendium. *MAbs, 4*(3), 413-415. doi: 10.4161/mabs.19931

[2] Todd, P. A., & Brogden, R. N. (1989). Muromonab CD3. A review of its pharmacology and therapeutic potential. *Drugs, 37*(6), 871-899. doi: 10.2165/00003495-198937060-00004

[3] Emmons, C., & Hunsicker, L. G. (1987). Muromonab-CD3 (Orthoclone OKT3): the first monoclonal antibody approved for therapeutic use. *Iowa Med, 77*(2), 78-82.

[4] The Antibody Society. Therapeutic monoclonal antibodies approved or in review in the EU or US. Retrieved 23/12/2022, from www.antibodysociety.org/resources/approved-antibodies

[5] Kaplon, H., Crescioli, S., Chenoweth, A., Visweswaraiah, J., & Reichert, J. M. (2023). Antibodies to watch in 2023. *MAbs, 15*(1), 2153410. doi: 10.1080/19420862.2022.2153410

[6] Liu, H. F., Ma, J., Winter, C., & Bayer, R. (2010). Recovery and purification process development for monoclonal antibody production. *MAbs, 2*(5), 480-499. doi: 10.4161/mabs.2.5.12645

[7] Walker, D. E., Yang, F., Carver, J., Joe, K., Michels, D. A., & Yu, X. C. (2017). A modular and adaptive mass spectrometry-based platform for support of bioprocess development toward optimal host cell protein clearance. *MAbs, 9*(4), 654-663. doi: 10.1080/19420862.2017.1303023

[8] Robert, F., Bierau, H., Rossi, M., Agugiaro, D., Soranzo, T., Broly, H., & Mitchell-Logean, C. (2009). Degradation of an Fc-fusion recombinant protein by host cell proteases: Identification of a CHO cathepsin D protease. *Biotechnol Bioeng, 104*(6), 1132-1141. doi: 10.1002/bit.22494

[9] Dixit, N., Salamat-Miller, N., Salinas, P. A., Taylor, K. D., & Basu, S. K. (2016). Residual Host Cell Protein Promotes Polysorbate 20 Degradation in a Sulfatase Drug Product Leading to Free Fatty Acid Particles. *J Pharm Sci, 105*(5), 1657-1666. doi: 10.1016/j.xphs.2016.02.029

[10] Vanderlaan, M., Zhu-Shimoni, J., Lin, S., Gunawan, F., Waerner, T., & Van Cott, K. E. (2018). Experience with host cell protein impurities in biopharmaceuticals. *Biotechnol Prog, 34*(4), 828-837. doi: 10.1002/btpr.2640

[11] Li, X., An, Y., Liao, J., Xiao, L., Swanson, M., Martinez-Fonts, K., . . . Richardson, D. D. (2021). Identification and characterization of a residual host cell protein hexosaminidase B associated with N-glycan degradation during the stability study of a therapeutic recombinant monoclonal antibody product. *Biotechnol Prog, 37*(3), e3128. doi: 10.1002/btpr.3128

[12] Jawa, V., Joubert, M. K., Zhang, Q., Deshpande, M., Hapuarachchi, S., Hall, M. P., & Flynn, G. C. (2016). Evaluating Immunogenicity Risk Due to Host Cell Protein Impurities in Antibody-Based Biotherapeutics. *Aaps j, 18*(6), 1439-1452. doi: 10.1208/s12248-016-9948-4

[13] Fischer, S. K., Cheu, M., Peng, K., Lowe, J., Araujo, J., Murray, E., . . . Song, A. (2017). Specific Immune Response to Phospholipase B-Like 2 Protein, a Host Cell Impurity in Lebrikizumab Clinical Material. *Aaps j, 19*(1), 254-263. doi: 10.1208/s12248-016-9998-7

[14] Jones, M., Palackal, N., Wang, F., Gaza-Bulseco, G., Hurkmans, K., Zhao, Y., . . . Connolly, T. (2021). "High-risk" host cell proteins (HCPs): A multi-company collaborative view. *Biotechnol Bioeng, 118*(8), 2870-2885. doi: 10.1002/bit.27808

[15] Guidance for Industry Q6B Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products (1999).

[16] *<1132> Residual Host Cell Protein Measurement in Biopharmaceuticals*. (2016).

[17] DNA and host cell protein impurities, routine testing versus validation studies. (1997). (CPMP/BWP/382/97).

[18] Gilgunn, S., El-Sabbahy, H., Albrecht, S., Gaikwad, M., Corrigan, K., Deakin, L., . . . Bones, J. (2019). Identification and tracking of problematic host cell proteins removed by a synthetic, highly functionalized nonwoven media in downstream bioprocessing of monoclonal antibodies. *J Chromatogr A, 1595*, 28-38. doi: 10.1016/j.chroma.2019.02.056

[19] Wang, X., Hunter, A. K., & Mozier, N. M. (2009). Host cell proteins in biologics development: Identification, quantitation and risk assessment. *Biotechnol Bioeng, 103*(3), 446-458. doi: 10.1002/bit.22304

[20] Champion, K. M., Madden, H., Dougherty, J., & Shacter, E. (2005). Defining Your Product Profile and Maintaining Control Over It, Part 2

Challenges of Monitoring Host Cell Protein Impurities. *BioProcess International*.

[21] Chon, J. H., & Zarbis-Papastoitsis, G. (2011). Advances in the production and downstream processing of antibodies. *N Biotechnol, 28*(5), 458-463. doi: 10.1016/j.nbt.2011.03.015

[22] Zhu-Shimoni, J., Yu, C., Nishihara, J., Wong, R. M., Gunawan, F., Lin, M., . . . Vanderlaan, M. (2014). Host cell protein testing by ELISAs and the use of orthogonal methods. *Biotechnol Bioeng, 111*(12), 2367-2379. doi: 10.1002/bit.25327

[23] Tscheliessnig, A. L., Konrath, J., Bates, R., & Jungbauer, A. (2013). Host cell protein analysis in therapeutic protein bioprocessing - methods and applications. *Biotechnol J, 8*(6), 655-670. doi: 10.1002/biot.201200018

[24] Bracewell, D. G., Francis, R., & Smales, C. M. (2015). The future of host cell protein (HCP) identification during process development and manufacturing linked to a risk-based management for their control. *Biotechnol. Bioeng., 112*(9), 1727-1737. doi: 10.1002/bit.25628

[25] Pilely, K., Johansen, M. R., Lund, R. R., Kofoed, T., Jørgensen, T. K., Skriver, L., & Mørtz, E. (2021). Monitoring process-related impurities in biologics-host cell protein analysis. *Anal Bioanal Chem*, 1-12. doi: 10.1007/s00216-021-03648-2

[26] European Pharmacopeia Monograph 2.6.34 Host-Cell Protein Assays. 9. (2017). *European Pharmacopeia Monograph 2.6.34 Host-Cell Protein Assays. 9, 2017.*

[27] Valente, K. N., Levy, N. E., Lee, K. H., & Lenhoff, A. M. (2018). Applications of proteomic methods for CHO host cell protein characterization in biopharmaceutical manufacturing. *Curr Opin Biotechnol, 53*, 144-150. doi: 10.1016/j.copbio.2018.01.004

[28] Hogwood, C. E., Bracewell, D. G., & Smales, C. M. (2014). Measurement and control of host cell proteins (HCPs) in CHO cell bioprocesses. *Curr Opin Biotechnol, 30*, 153-160. doi: 10.1016/j.copbio.2014.06.017

[29] Chen, Y., Xu, C. F., Stanley, B., Evangelist, G., Brinkmann, A., Liu, S., . . . Yeung, B. (2021). A Highly Sensitive LC-MS/MS Method for Targeted Quantitation of Lipase Host Cell Proteins in Biotherapeutics. *J Pharm Sci, 110*(12), 3811-3818. doi: 10.1016/j.xphs.2021.08.024

[30] Doneanu, C., Fang, J., Alelyunas, Y., Yu, Y. Q., Wrona, M., & Chen, W. (2018). An HS-MRM Assay for the Quantification of Host-cell Proteins in Protein Biopharmaceuticals by Liquid Chromatography Ion Mobility QTOF Mass Spectrometry. *J Vis Exp*(134). doi: 10.3791/55325

[31] Doneanu, C. E., Xenopoulos, A., Fadgen, K., Murphy, J., Skilton, S. J., Prentice, H., . . . Chen, W. (2012). Analysis of host-cell proteins in biotherapeutic proteins by comprehensive online two-dimensional liquid chromatography/mass spectrometry. *MAbs, 4*(1), 24-44. doi: 10.4161/mabs.4.1.18748

[32] Gao, X., Rawal, B., Wang, Y., Li, X., Wylie, D., Liu, Y. H., . . . Richardson, D. D. (2020). Targeted Host Cell Protein Quantification by LC-MRM Enables Biologics Processing and Product Characterization. *Anal Chem, 92*(1), 1007-1015. doi: 10.1021/acs.analchem.9b03952

[33] Kreimer, S., Gao, Y., Ray, S., Jin, M., Tan, Z., Mussa, N. A., . . . Karger, B. L. (2017). Host Cell Protein Profiling by Targeted and Untargeted Analysis of Data Independent Acquisition Mass

Spectrometry Data with Parallel Reaction Monitoring Verification. *Anal Chem, 89*(10), 5294-5302. doi: 10.1021/acs.analchem.6b04892

[34] E, S. Y., Hu, Y., Molden, R., Qiu, H., & Li, N. (2022). Identification and Quantification of a Problematic Host Cell Protein to Support Therapeutic Protein Development. *J Pharm Sci*. doi: 10.1016/j.xphs.2022.10.008

[35] Krawitz, D. C., Forrest, W., Moreno, G. T., Kittleson, J., & Champion, K. M. (2006). Proteomic studies support the use of multi-product immunoassays to monitor host cell protein impurities. *Proteomics, 6*(1), 94-110. doi: 10.1002/pmic.200500225

[36] Ma, J., & Kilby, G. W. (2020). Sensitive, Rapid, Robust, and Reproducible Workflow for Host Cell Protein Profiling in Biopharmaceutical Process Development. *J Proteome Res, 19*(8), 3396-3404. doi: 10.1021/acs.jproteome.0c00252

[37] Yang, F., Walker, D. E., Schoenfelder, J., Carver, J., Zhang, A., Li, D., . . . Michels, D. A. (2018). A 2D LC-MS/MS Strategy for Reliable Detection of 10-ppm Level Residual Host Cell Proteins in Therapeutic Antibodies. *Anal Chem, 90*(22), 13365-13372. doi: 10.1021/acs.analchem.8b03044

[38] Husson, G., Delangle, A., O'Hara, J., Cianferani, S., Gervais, A., Van Dorsselaer, A., . . . Carapito, C. (2018). Dual Data-Independent Acquisition Approach Combining Global HCP Profiling and Absolute Quantification of Key Impurities during Bioprocess Development. *Anal Chem, 90*(2), 1241-1247. doi: 10.1021/acs.analchem.7b03965

[39] Pythoud, N., Bons, J., Mijola, G., Beck, A., Cianférani, S., & Carapito, C. (2021). Optimized Sample Preparation and Data Processing of Data-Independent Acquisition Methods for the Robust Quantification of Trace-Level Host Cell Protein Impurities in Antibody Drug Products. *J Proteome Res, 20*(1), 923-931. doi: 10.1021/acs.jproteome.0c00664

[40] Gillet, L. C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., . . . Aebersold, R. (2012). Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics, 11*(6), O111.016717. doi: 10.1074/mcp.O111.016717

[41] Zhang, F., Ge, W., Ruan, G., Cai, X., & Guo, T. (2020). Data-Independent Acquisition Mass Spectrometry-Based Proteomics and Software Tools: A Glimpse in 2020. *Proteomics, 20*(17-18), e1900276. doi: 10.1002/pmic.201900276

[42] Barkovits, K., Pacharra, S., Pfeiffer, K., Steinbach, S., Eisenacher, M., Marcus, K., & Uszkoreit, J. (2020). Reproducibility, Specificity and Accuracy of Relative Quantification Using Spectral Library-based Data-independent Acquisition. *Mol Cell Proteomics, 19*(1), 181-197. doi: 10.1074/mcp.RA119.001714

[43] Ludwig, C., Gillet, L., Rosenberger, G., Amon, S., Collins, B. C., & Aebersold, R. (2018). Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol, 14*(8), e8126. doi: 10.15252/msb.20178126

[44] Cai, X., Ge, W., Yi, X., Sun, R., Zhu, J., Lu, C., . . . Guo, T. (2021). PulseDIA: Data-Independent Acquisition Mass Spectrometry Using Multi-Injection Pulsed Gas-Phase Fractionation. *J Proteome Res, 20*(1), 279-288. doi: 10.1021/acs.jproteome.0c00381

[45] Amodei, D., Egertson, J., MacLean, B. X., Johnson, R., Merrihew, G. E., Keller, A., . . . MacCoss, M. J. (2019). Improving Precursor Selectivity in Data-Independent Acquisition Using Overlapping Windows. *J Am Soc Mass Spectrom, 30*(4), 669-684. doi: 10.1007/s13361-018-2122-8

[46] Pino, L. K., Just, S. C., MacCoss, M. J., & Searle, B. C. (2020). Acquiring and Analyzing Data Independent Acquisition Proteomics Experiments without Spectrum Libraries. *Mol Cell Proteomics, 19*(7), 1088-1103. doi: 10.1074/mcp.P119.001913

[47] Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S., & Ralser, M. (2020). DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods, 17*(1), 41-44. doi: 10.1038/s41592-019-0638-x

[48] Gotti, C., Roux-Dalvai, F., Joly-Beauparlant, C., Mangnier, L., Leclercq, M., & Droit, A. (2021). Extensive and Accurate Benchmarking of DIA Acquisition Methods and Software Tools Using a Complex Proteomic Standard. *J Proteome Res, 20*(10), 4801-4814. doi: 10.1021/acs.jproteome.1c00490

[49] Tsou, C. C., Avtonomov, D., Larsen, B., Tucholska, M., Choi, H., Gingras, A. C., & Nesvizhskii, A. I. (2015). DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat Methods, 12*(3), 258-264, 257 p following 264. doi: 10.1038/nmeth.3255

[50] van der Spek, S. J. F., Gonzalez-Lozano, M. A., Koopmans, F., Miedema, S. S. M., Paliukhovich, I., Smit, A. B., & Li, K. W. (2021). Age-Dependent Hippocampal Proteomics in the APP/PS1 Alzheimer Mouse Model: A Comparative Analysis with Classical SWATH/DIA and directDIA Approaches. *Cells, 10*(7). doi: 10.3390/cells10071588

[51] Gessulat, S., Schmidt, T., Zolg, D. P., Samaras, P., Schnatbaum, K., Zerweck, J., . . . Wilhelm, M. (2019). Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning. *Nat Methods, 16*(6), 509-518. doi: 10.1038/s41592-019-0426-7

[52] Searle, B. C., Swearingen, K. E., Barnes, C. A., Schmidt, T., Gessulat, S., Küster, B., & Wilhelm, M. (2020). Generating high quality libraries for DIA MS with empirically corrected peptide predictions. *Nat Commun, 11*(1), 1548. doi: 10.1038/s41467-020-15346-1

[53] Van Puyvelde, B., Willems, S., Gabriels, R., Daled, S., De Clerck, L., Vande Casteele, S., . . . Dhaenens, M. (2020). Removing the Hidden Data Dependency of DIA with Predicted Spectral Libraries. *Proteomics, 20*(3-4), e1900306. doi: 10.1002/pmic.201900306

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

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

[56] Madsen, J. A., Farutin, V., Carbeau, T., Wudyka, S., Yin, Y., Smith, S., . . . Capila, I. (2015). Toward the complete characterization of host cell proteins in biotherapeutics via affinity depletions, LC-MS/MS, and multivariate analysis. *MAbs, 7*(6), 1128-1137. doi: 10.1080/19420862.2015.1082017

[57] Thompson, J. H., Chung, W. K., Zhu, M., Tie, L., Lu, Y., Aboulaich, N., . . . Mo, W. D. (2014). Improved detection of host cell proteins (HCPs) in a mammalian cell-derived antibody drug using liquid chromatography/mass spectrometry in conjunction with an HCP-enrichment strategy. *Rapid Commun Mass Spectrom, 28*(8), 855-860. doi: 10.1002/rcm.6854

[58] Wang, Q., Slaney, T. R., Wu, W., Ludwig, R., Tao, L., & Leone, A. (2020). Enhancing Host-Cell Protein Detection in Protein Therapeutics Using HILIC Enrichment and Proteomic Analysis. *Anal Chem, 92*(15), 10327-10335. doi: 10.1021/acs.analchem.0c00360

[59] Doneanu, C. E., Anderson, M., Williams, B. J., Lauber, M. A., Chakraborty, A., & Chen, W. (2015). 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., 87*(20), 10283-10291. doi: 10.1021/acs.analchem.5b02103

[60] Farrell, A., Mittermayr, S., Morrissey, B., Mc Loughlin, N., Navas Iglesias, N., Marison, I. W., & Bones, J. (2015). Quantitative host cell protein analysis using two dimensional data independent LC-MS(E). *Anal Chem, 87*(18), 9186-9193. doi: 10.1021/acs.analchem.5b01377

[61] Johnson, R. O., Greer, T., Cejkov, M., Zheng, X., & Li, N. (2020). Combination of FAIMS, Protein A Depletion, and Native Digest Conditions Enables Deep Proteomic Profiling of Host Cell Proteins in Monoclonal Antibodies. *Anal Chem, 92*(15), 10478-10484. doi: 10.1021/acs.analchem.0c01175

[62] Hebert, A. S., Prasad, S., Belford, M. W., Bailey, D. J., McAlister, G. C., Abbatiello, S. E., . . . Coon, J. J. (2018). Comprehensive Single-Shot Proteomics with FAIMS on a Hybrid Orbitrap Mass Spectrometer. *Anal Chem, 90*(15), 9529-9537. doi: 10.1021/acs.analchem.8b02233

[63] Silva, J. C., Gorenstein, M. V., Li, G. Z., Vissers, J. P., & Geromanos, S. J. (2006). Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics, 5*(1), 144-156. doi: 10.1074/mcp.M500230-MCP200

[64] Molden, R., Hu, M., Yen, E. S., Saggese, D., Reilly, J., Mattila, J., . . . Li, N. (2021). Host cell protein profiling of commercial therapeutic protein drugs as a benchmark for monoclonal antibody-based therapeutic protein development. *MAbs, 13*(1), 1955811. doi: 10.1080/19420862.2021.1955811

[65] Trauchessec, M., Hesse, A. M., Kraut, A., Berard, Y., Herment, L., Fortin, T., . . . Manin, C. (2021). An innovative standard for LC-MS-based HCP profiling and accurate quantity assessment: Application to batch consistency in viral vaccine samples. *Proteomics, 21*(5), e2000152. doi: 10.1002/pmic.202000152

[66] Strasser, L., Oliviero, G., Jakes, C., Zaborowska, I., Floris, P., Ribeiro da Silva, M., . . . Bones, J. (2021). Detection and quantitation of host cell proteins in monoclonal antibody drug products using automated sample preparation and data-independent acquisition LC-MS/MS. *J Pharm Anal, 11*(6), 726-731. doi: 10.1016/j.jpha.2021.05.002

[67] Nie, S., Greer, T., O'Brien Johnson, R., Zheng, X., Torri, A., & Li, N. (2021). Simple and Sensitive Method for Deep Profiling of Host Cell Proteins in Therapeutic Antibodies by Combining Ultra-Low Trypsin Concentration Digestion, Long Chromatographic Gradients, and BoxCar Mass Spectrometry Acquisition. *Anal Chem, 93*(10), 4383-4390. doi: 10.1021/acs.analchem.0c03931

[68] Deutsch, E. W., Csordas, A., Sun, Z., Jarnuczak, A., Perez-Riverol, Y., Ternent, T., . . . Vizcaíno, J. A. (2017). The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. *Nucleic Acids Res, 45*(D1), D1100-d1106. doi: 10.1093/nar/gkw936

**Figures**

Figure 1: A) Number of HCP peptides and protein groups identified (light blue) and quantified after applying validation filters (navy blue) in DDA mode. B) Overlap of host cell proteins identified in Huang et *al.* (2017) [54] (grey), Molden et *al.* (2021) [64] (yellow), Ma et *al.* (2020) [36] (purple) and in the current study (blue). All studies used a native digestion protocol adapted from Huang et *al.* [54]. Data from Molden et *al*. were acquired in LC-MS/MS, vs. nanoLC-MS/MS for Huang et *al*., Ma et *al*. (44 min gradient) and our study. C) Pearson coefficient and correlation between the individual amounts of host cell proteins in ng of HCP per mg of NIST mAb (ppm) in our study using nanoLC-MS/MS DDA analysis after native digestion vs. previously reported quantities using native digestion followed by nano-LC-MS/MS obtained by Huang et *al*. (grey) and Molden et *al*. (yellow). For representation convenience, results from Molden, originally in micromoles to moles ratio, were converted into ng of HCP to mg of mAb ratio. D) Number of HCP peptides and protein groups identified (light colors) and quantified after applying validation filters (dark colors) using DDA (blue), FAIMS DDA (green), DIA (orange) and GPF DIA (red).

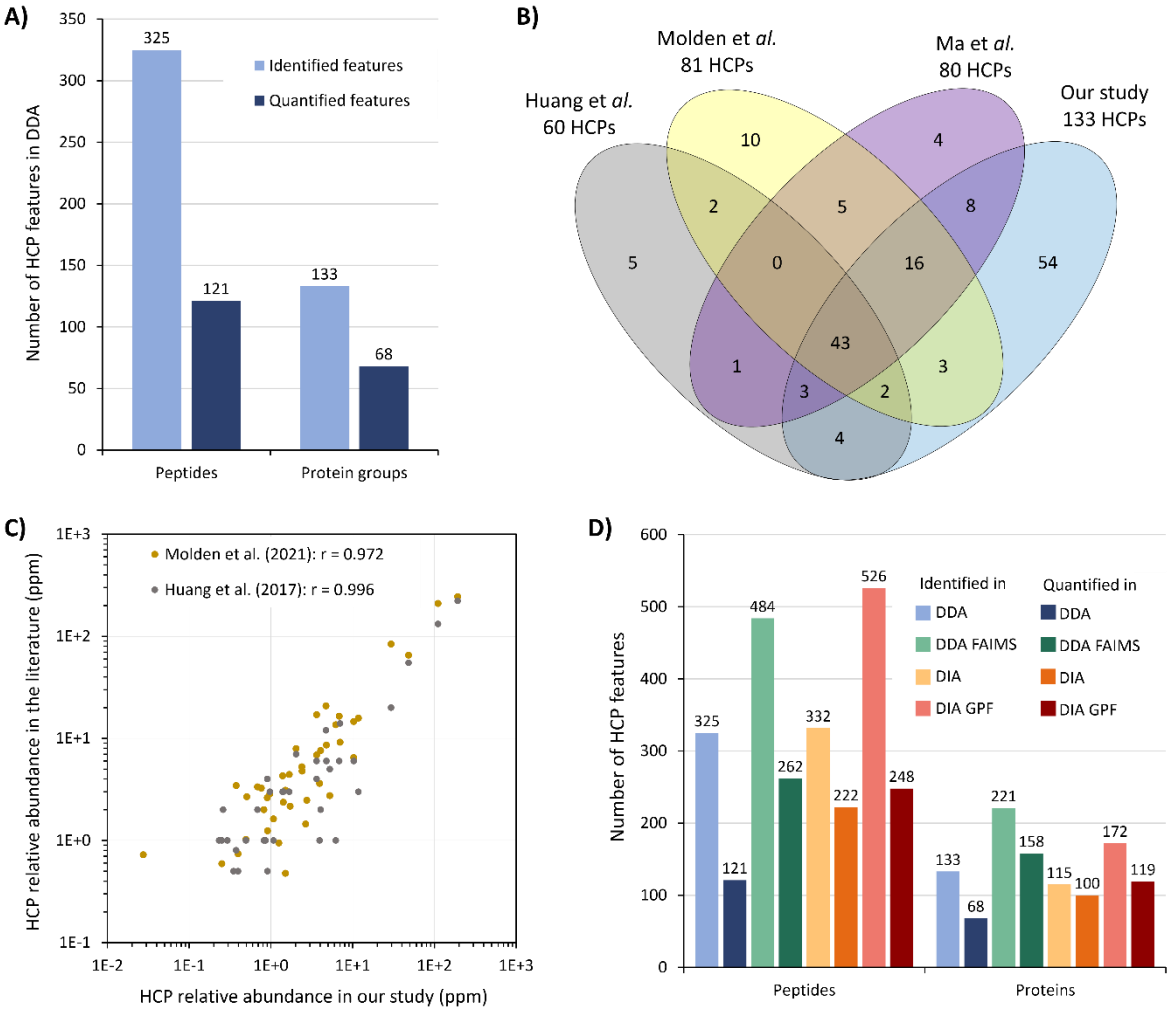

A)
Number of HCP features in DDA
350
300
250
200
150
100
50
0
325
121
133
68
Identified features
Quantified features
Peptides
Protein groups
B)
Molden et al.
81 HCPs
Ma et al.
80 HCPs
Huang et al.
60 HCPs
Our study
133 HCPs
10
4
2
5
8
5
0
16
54
43
1
3
3
2
4
C)
Molden et al. (2021): r = 0.972
Huang et al. (2017): r = 0.996
HCP relative abundance in the literature (ppm)
HCP relative abundance in our study (ppm)
1E+3
1E+2
1E+1
1E+0
1E-1
1E-2
1E-1
1E+0
1E+1
1E+2
1E+3
D)
Number of HCP features
600
500
400
300
200
100
0
325
121
484
262
332
222
526
248
133
68
221
158
115
100
172
119
Identified in
DDA
DDA FAIMS
DIA
DIA GPF
Quantified in
DDA
DDA FAIMS
DIA
DIA GPF
Peptides
Proteins

Figure 2: Violin plots representing the density of host cell proteins as a function of their concentration in ng of HCP per mg of NIST mAb (ppm). Blue dots represent proteins quantified by both approaches and red dots proteins quantified only in one approach, with A) the comparison of DDA and FAIMS DDA and B) the comparison of DIA and GPF DIA.

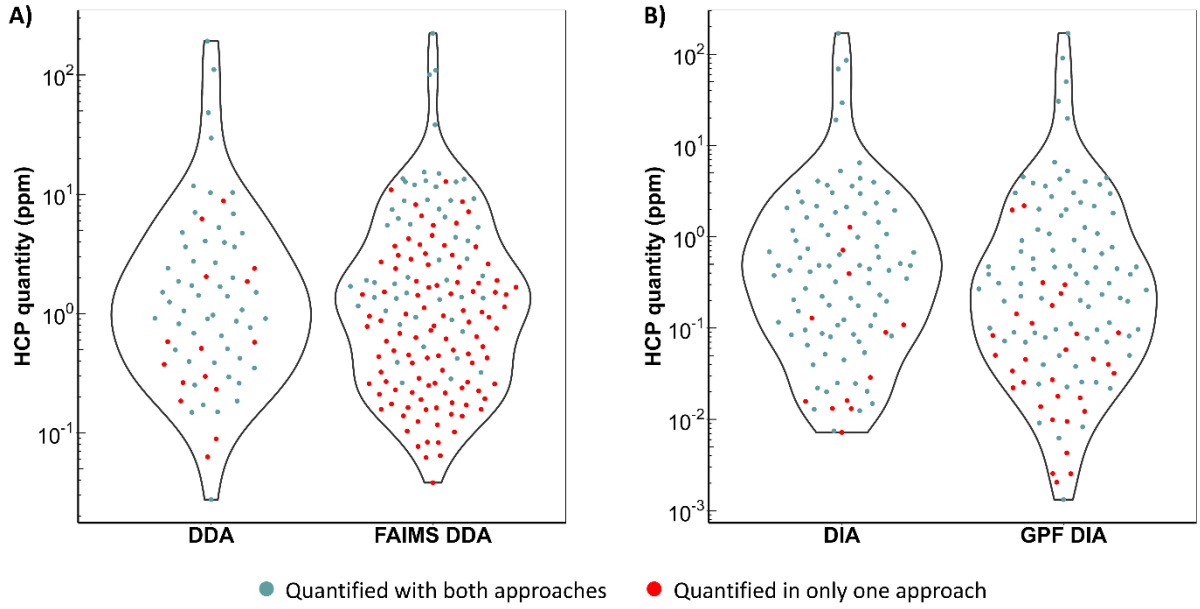
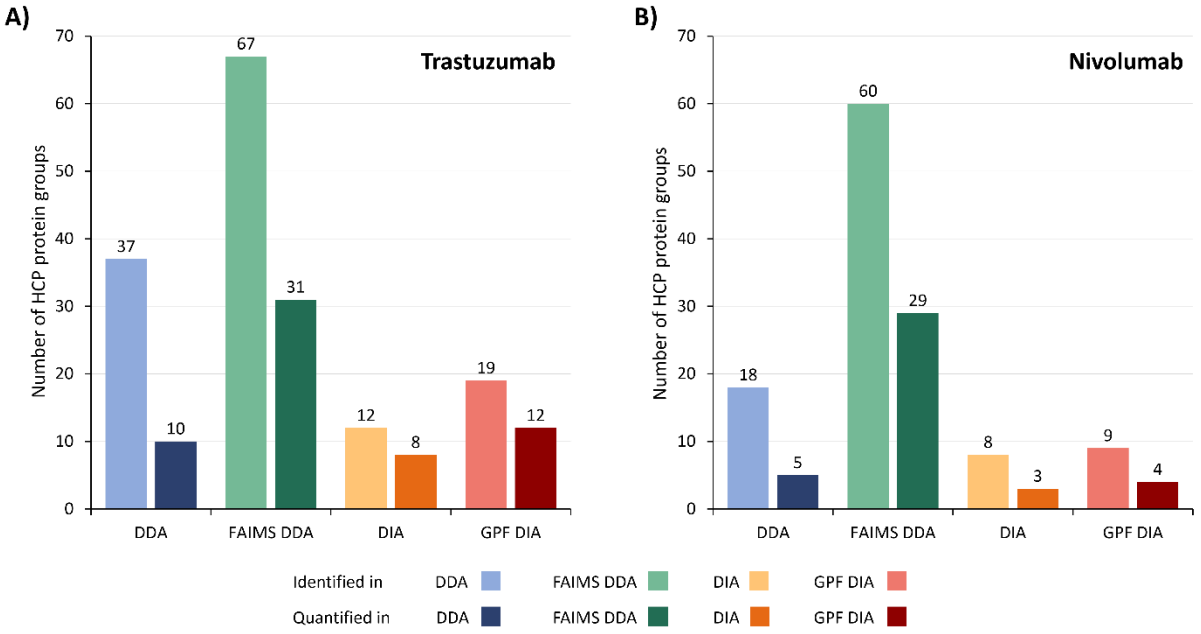


Figure 3: Number of HCP peptides identified (light colors) and quantified after applying validation filters (dark colors) using DDA (blue), FAIMS DDA (green), DIA (orange) and GPF DIA (red) A) in trastuzumab and B) in nivolumab. C) Average quantity of HCPs quantified in nivolumab and trastuzumab drug products using DDA, DDA FAIMS, DIA and DIA GPF.




C)

| Drug product | DDA | FAIMS DDA | DIA | DIA GPF |
|---|---|---|---|---|
| Trastuzumab | 10 ± 1 ppm | 57 ± 9 ppm | 10 ± 2 ppm | 9 ± 2 ppm |
| Nivolumab | 6 ± 0.3 ppm | 54 ± 6 ppm | 10 ± 6 ppm | 11 ± 6 ppm |