Article

Dual Data Independent Acquisition approach combining global HCP profiling and absolute quantification of key impurities during bioprocess development

Gauthier Husson, Aurélie Delangle, John O'Hara, Sarah Cianférani, Annick Gervais, Alain Van Dorsselaer, Daniel G. Bracewell, and Christine Carapito

Anal. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.analchem.7b03965 • Publication Date (Web): 13 Dec 2017

Dual Data Independent Acquisition approach combining global HCP profiling and absolute quantification of key impurities during bioprocess development

Authors list

Gauthier Husson[1], Aurélie Delangle[2], John O'Hara[3], Sarah Cianferani[1], Annick Gervais[2], Alain Van Dorsselaer[1], Dan Bracewell[4], Christine Carapito[1*]

Affiliations

[1]Laboratoire de Spectrométrie de Masse BioOrganique, Université de Strasbourg, CNRS, IPHC, UMR 7178, F-67000 Strasbourg, France

[2]Department of Analytical Sciences Biologicals, UCB Pharma s.a., Chemin du Foriest, B-1420 Braine L'alleud, Belgium

[3]Department of Analytical Sciences Biologicals, UCB Pharma s.a., 216 Bath Road, Slough SL1 4EN, UK

[4]Dept. Biochemical Engineering, University College London, Gower Street, London, WC1E 6BT, UK

*Corresponding author: Christine Carapito, ccarapito@unistra.fr, Laboratoire de Spectrométrie de Masse BioOrganique, Université de Strasbourg, CNRS, IPHC, UMR 7178, 25 Rue Becquerel, F-67087 Strasbourg, France

Keywords

Bioprocess development - Host cell proteins - Mass spectrometry characterisation - Data Independent Acquisition – Absolute quantification

Abstract

Host cell proteins (HCP) are a major class of impurities derived from recombinant protein production process. While HCP are usually monitored by ELISA, mass spectrometry (MS) based approaches are emerging as promising orthogonal methods. Here, we developed an original method relying on data independent acquisition (DIA) coupling global HCP amounts estimation (Top 3) and absolute quantification with isotope dilution (ID). The method named Top 3-ID-DIA was benchmarked against ELISA and a gold standard selected reaction monitoring assay (ID-SRM). Various samples generated at different steps and conditions of the purification process, including different culture durations, harvest procedures and purification protocols were used to compare the methods. Overall, HCP were quantified over 5 orders of magnitude and down to sub-ppm level. The Top 3-ID-DIA strategy proved to be equivalent to the gold standard ID-SRM in terms of sensitivity (1-10 ppm), accuracy and precision. Moreover, 81% of the Top 3 estimations were accurate within a factor of 2 when compared to ID-SRM. Thus, our approach aggregates global HCP profiling for comprehensive process understanding with absolute quantification of key HCP within a single analysis, and provides an improved support for bioprocess development and product purity assessment.

Host cell proteins (HCP) constitute a major class of impurities that must be monitored and efficiently removed during recombinant protein purification process[1]. Remaining HCP in the final drug product can reduce the drug efficacy[2-4] or induce adverse patient reactions[5,6]. Therefore, HCP amounts in the final drug product must be provided to the regulatory authorities[7]. As a rule of thumb, HCP must be quantified below 100 ppm in the final product by enzyme-linked immunosorbent assay (ELISA)[8]. Evidence is growing however that ELISA does not provide comprehensive HCP quantification since it only detects HCP that induced immune response in animals during ELISA development and provides only total HCP amounts without any information on the identity of the detected HCP[1,9-13]. Finally, developing a specific ELISA is costly and time consuming[10].

As an alternative, mass spectrometry (MS) approaches recently revealed to be most promising to characterise HCP contents as they allow unbiased quantification and individual HCP monitoring. Recent advances in the MS field, notably the use of MS2 signals for quantification by targeted methods (selected reaction monitoring SRM[14] or parallel reaction monitoring PRM[15]) or data independent acquisition (DIA) methods, allowed a 2- to 8-fold gain in sensitivity[16], and a significant gain in specificity and dynamic range when compared to the use of MS1 signals. These features are particularly crucial in the HCP field in which very low abundant proteins have to be quantified besides a highly abundant predominant protein.

Targeted SRM assays run on triple-quadrupole type instruments and coupled to isotope dilution have, for long, been the gold standard MS-based quantification technique offering highest sensitivity, accuracy and robustness but with limited multiplexing. More recently, DIA modes based on the collection of MS2 information for all detectable species were introduced on high resolution/accurate mass (HRAM) instruments in order to extract valuable quantitative information from whole complex proteome maps[16]. Two-dimensional liquid chromatography coupled to DIA-type MS^E (2D-LC MS^E) technology has been used in a few studies to quantify HCP in monoclonal antibody (mAb) solutions. A Top 3 quantification strategy, which assumes that the signal of the three best responding peptides

per mole of protein is constant within a coefficient of variation of less than 10%[17], was used in these studies to estimate absolute amounts of HCP down to a lower limit of quantification (LLOQ) of about 10 ppm[18-23]. Time is also a factor, at least 10 hours were necessary for this type of analysis, which is not easily compatible with real time process support. Alternatively, a 1D-LC sequential windowed acquisition of all theoretical fragment ion mass spectra (SWATH) DIA approach was recently shown by Walker and co-workers[24] to achieve equivalent sensitivity in only one hour.

In this work, we developed MS-based quantification approaches to characterise and profile HCP contents in a variety of mAb samples obtained from different steps and conditions of the purification process. We propose an original dual DIA-based HCP quantification method allowing both global HCP profiling and absolute quantification of a subset of key HCP, thereby leveraging the advantages of global and targeted approaches within a single analysis. Our method was benchmarked against ELISA and a reference isotope dilution SRM assay (ID-SRM).

Experimental Section

Cell culture. An IgG4 A33 mAb producing CHO-DG44 cell line (provided by UCB Pharma, Brussels, Belgium) was cultivated in batch mode using a protein free and chemically defined CD CHO medium (Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 6 mM Glutamine (Thermo Fisher Scientific) and 5 nM Methotrexate (Merck, Darmstadt, Germany). Cells were grown in 1 L Erlenmeyer flasks (300 mL working volume) and incubated at 36.5°C with 5% CO2 on an orbital shaker (123 rpm). The cell concentration and viability were monitored every day using a Vi-Cell (Beckman Coulter, Brea, CA, USA). Viable cells were distinguished from dead cells using the trypan blue dye exclusion method.

Samples production. The ultra scale-down (USD) shearing, harvest procedure and clarified cell culture fluid (CCCF) fractions collection protocols are described in Supporting Information.

Protein A chromatography. mAbs were purified using 1 mL HiTrap MabSelect SuRe columns (GE Healthcare Life Sciences, Pittsburgh, PA) on an AKTA Pure system (GE Healthcare Life Sciences). Two purification types were performed at 1 mL/min using either (i) a standard protocol[25] or (ii) a modified protocol[26]. (i) Standard protocol: equilibration step (5 column volumes (CV) of PBS, pH 7.4) followed by loading of an appropriate volume of CCCF for 20 mg mAb. The column was washed with loading buffer, and the mAb was eluted (0.1 M citrate pH 3.6). (ii) Modified protocol: equilibration step (5 CV 25 mM Tris, 100 mM NaCl pH 7.4) followed by loading of an appropriate volume of CCCF for 20 mg mAb. The column was washed with loading buffer. An intermediate wash (5 CV 25 mM Tris, 10% isopropanol, 1 M urea, pH 9) and a pre-elution wash (3 CV 50 mM citrate, pH 4.4) were performed before mAb elution (100 mM acetate, pH 3.6). After elution, post protein A (PPA) fractions were directly neutralised to pH 6 using 2 M Tris HCl pH 8.8. A dedicated new column was used for each purification.

Protein quantification. mAb and global protein quantifications were performed as described in Supporting Information.

HCP-ELISA. The HCP were quantified using the CHO HCP ELISA kit, 3G (Cygnus Technologies, Southport, NC, USA) in technical triplicates according to the manufacturer's protocol.

Sample preparation. Samples were separated using SDS-PAGE for spectral library generation (pooled CCCF and pooled PPA fractions), or stacked in a single band for HCP quantification. Retention time standards (iRT, Biognosys, Schlieren, Switzerland) and four accurately quantified standard proteins (on-column 100 fmol ADH (yeast alcohol dehydrogenase P00330), 20 fmol PYGM (phosphorylase b P00489), 5 fmol BSA (bovin serum albumin P02769) and 2 fmol ENL (yeast enolase P00924) from the MassPREP Digestion Standard Kit, Waters, Milford, MA, USA) were spiked in each samples. For absolute quantification experiments, a concentration-balanced mixture of 20 accurately quantified stable isotope labelled peptides (AQUA peptides, Thermo Fisher Scientific) were spiked.

Mass spectrometry analysis. Data dependent acquisition (DDA) and data independent acquisition-sequential windowed acquisition of all theoretical fragment ion mass spectra (DIA-SWATH) analyses were performed on an Eksigent NanoLC 400 system operated in microLC-mode and coupled to a TripleTOF 6600 quadrupole-time of flight mass spectrometer (both from SCIEX, Framingham, MA, USA). Selected reaction monitoring (SRM) analyses were performed on a Dionex UltiMate 3000 operated in microLC-mode and coupled to a TSQ Vantage triple quadrupole mass spectrometer (both from Thermo Fisher Scientific). On both couplings, 8 µg of peptides were separated on a ZORBAX 300SB-C18 column (150 mm x 300 µm with 3.5 µm diameter particles, Agilent Technologies). Chromatographic gradient and MS settings are given in Supporting Information.

DIA-SWATH targeted data extraction. A spectral library was generated as described in Supporting Information. DIA data were processed using Skyline[27] (version 3.5.9.10061). Validated proteotypic peptides from the spectral library were extracted with following parameters (based on previous work[28] and in-house optimisations on standard samples, data not shown): the 6 most intense 1+ b- and y-type product ions were extracted, from ion 3 to last ion – 1, while the precursors with less than 3 transitions were excluded. Resolving power was set to 50 000, and a retention time tolerance of 5

min (+/- 2.5 min) was used. Retention times were predicted with iRT standards (Biognosys). Peaks were reintegrated using the target decoy approach (reverse sequences) of the mProphet peak-scoring model, and a Q-value was assigned to each peak. Peak integrations were manually checked and curated for HCP of interest. Fragment areas, detection Q value and library dot product were exported for each peptide in .csv files.

Top 3 estimation. Only peptides with Q-value below 0.01 (corresponding to a false discovery rate of 1%) and dot-product above 0.6 were kept. The fragment areas were summed for each peptide and the 3 best responding peptides were summed for each protein. Only proteins quantified in at least two replicates in at least one sample were kept, independently for CCCF and PPA fractions. The universal signal response factor[17] (signal / mol of protein) was calculated using PYGM and was used to estimate mol quantities of all proteins. Using molecular weights and mAb quantifications, individual HCP amounts in ppm were estimated. Only quantifications with a coefficient of variation (CV) below 20% between technical triplicates were used to build a heat map (Table S1) and calculate total HCP amounts in each sample.

Selection of 10 HCP and their proteotypic peptides. Ten HCP were chosen based on their potential immunogenicity, proteolytic activity, purification behaviour, or estimated abundance using preliminary data acquired on the samples and literature knowledge. The selected HCP and their proteotypic peptides are described in Supporting Information and Table S2.

Absolute quantification. Six transitions were analysed for each precursor ion for both SRM and DIA approaches. If comprised in the linear range of the assay as determined by calibration curves (detailed procedures for calibration curves and LLOQ determination are provided in Supporting Information and Figure S1 and S2), the ratios between endogen and stable isotope labelled AQUA peptides were used to calculate the mol amounts of endogenous peptides, which were averaged to calculate the mol amounts of corresponding proteins. Using molecular weights and injected mAb quantities, individual HCP amounts in ppm were calculated.

Results

Overview of MS approaches developed to monitor HCP contents during bioprocess development

MS-based HCP quantification strategies were evaluated according to the global workflow presented in Figure 1. First, we investigated the impact on the HCP content at harvest of the cell culture duration and cell viability, as it was shown to induce significant changes in harvest HCP composition[11,12,29-31]. Cell culture fluid was thus collected at days 7 and 10, corresponding to cell viabilities of 71% and 8%, respectively (Figure S3). Then, different shear stress conditions during harvest were compared using the ultra scale down (USD) shear device[32,33] developed at University College London. Finally, two protein A purification protocols were compared: a standard protocol[25], and a modified protocol[26] including a high pH wash with a combination of 1 M urea and 10% isopropanol in order to disrupt mAb – HCP interactions while preserving mAb – protein A bindings. Overall, 4 clarified cell culture fluid (CCCF) and 8 post protein A (PPA) fractions were collected (Figure 2).

MS-based quantification methods were developed and benchmarked for HCP monitoring on all samples: an original Top 3-ID-DIA approach combining DIA and Top 3 quantification of all detected HCP with isotope dilution for absolute quantification of a subset of 10 selected HCP, and a gold standard isotope dilution SRM assay to absolutely quantify the same 10 selected HCP.

For global profiling, a spectral library was generated as described in Experimental Section and used to extract signals for all detectable HCP. For Top 3 amount estimations, the PYGM protein was used as the reference protein, and ADH, BSA and ENO were used as quantification controls. Stable isotope labelled peptides were spiked into all samples to allow accurate absolute quantification of the 10 selected HCP with both Top 3-ID-DIA and ID-SRM methods.

Global HCP contents estimation over bioprocess steps

We generated a comprehensive HCP spectral library from the mAb-producing CHO cell line containing 25 338 proteotypic peptides corresponding to 3 220 proteins. The use of a SDS-PAGE separation prior to LC-MS/MS analyses allowed limiting interferences from overwhelming mAb peptides and building a large spectral library specifically corresponding to the producing cell line, without having to use a null CHO cell line. Two distinct acquisition methods using 75 variable windows were optimised for CCCF and PPA fractions (Table S3), respectively, to reduce potential interferences and achieve highest specificity in DIA analysis. Transition groups specific to all peptides contained in the spectral library were extracted using their predicted retention times (thanks to retention time standards) as described in the Experimental Section. Noteworthy is that in PPA samples, the low number of target HCP peptides makes the differentiation between targets and decoys challenging, the false discovery rate strategy suboptimal[34] and an additional dot product threshold was thus applied to remove highly interfered peptides in these samples. The universal response factor[17] (signal / mol of protein) allowed the quantification controls ADH, BSA and ENO spiked at 100 fmol, 5 fmol and 2 fmol, to be estimated at 140 ± 12 fmol, 7 ± 2 fmol and 0.5 ± 0.2 fmol, respectively. Only quantifications achieved with a coefficient of variation (CV) of less than 20% between triplicates were summed to calculate total HCP amounts for each sample. On average, 1 454 HCP were quantified in the CCCF fractions representing 288 513 to 389 657 ppm and 119 HCP in the PPA fractions representing 2 646 to 5 386 ppm (Figure 2). These global HCP amounts are in accordance with previous studies focused on HCP quantification by MS[19,23]. We could estimate individual HCP amounts ranging from 0.5 to 16 192 ppm in the CCCF fractions, and from 0.1 to 731 ppm in the PPA fractions, thus covering a dynamic range of 5 orders of magnitude (Table S1). In parallel, HCP were quantified in PPA fractions using ELISA and ranged from 276 to 959 ppm, which is significantly lower when compared to Top 3-ID-DIA estimations (on average 8 times lower).

Absolute quantification of 10 selected HCP

ID-SRM quantification. Working with the same LC gradient as for the Top 3-ID-DIA analyses, a time-scheduled ID-SRM method was developed, first, using 20 crude stable isotope labelled synthetic peptides spiked in CCCF and PPA matrices. Six specific transitions were chosen for each peptide, and collision energy values were individually optimised for each transition. Then, accurately quantified stable isotope labelled AQUA peptides were spiked in known amounts into the samples for absolute quantification. Calibration curves were realised for each peptide to determine the linear quantification range and LLOQ of the assay (Figure S1). FDA-approved criteria[35] were applied for calibration curves interpretation as detailed in Supporting Information. LLOQ values could be determined for 13 out of the 20 peptides. Absolute quantification could be obtained for 8 out of the 10 targeted HCP, ranging from 1.7 to 23 681 ppm thus covering a dynamic range of 4.1 orders of magnitude (Table S4, Figure S4). As expected, Pyruvate kinase was found very abundant in CCCF fractions (from 13 674 to 23 681 ppm), while cytoplasmic Isoleucyl-tRNA synthetase was very low abundant in PPA fractions (below 18 ppm). Difficult to remove HCPs were detected in the PPA fractions from 1.7 to 106 ppm with the exception of Pyruvate kinase found at 157 and 536 ppm in PPA 7 and 8 fractions, respectively. HEAT repeat-containing protein 3 and Eukaryotic translation initiation factor 3 subunit L were not quantified because no valid calibration curve could be built for their corresponding peptides.

Top 3-ID-DIA quantification. The same 10 HCP were absolutely quantified using isotope dilution within the Top 3-ID-DIA experiment. Identical criteria were applied as for the ID-SRM approach to build calibration curves and determine LLOQ in DIA mode. Calibration curves and LLOQ values could be determined for 17 out of the 20 peptides (Figure S2). The 10 HCP were accurately quantified from 0.7 to 26 017 ppm thus covering a dynamic range of 4.6 orders of magnitude (Table S4, Figure S4). Again, as expected, Pyruvate kinase was found highly abundant in CCCF fractions (from 15 494 to 26 017 ppm), while cytoplasmic Isoleucyl-tRNA synthetase was very low abundant in PPA fractions (below 32 ppm). Difficult to remove HCPs were consistently detected in PPA fractions from 0.7 to 120

ppm, excepted for Pyruvate kinase which was quantified at 172 and 456 ppm in PPA 7 and 8 fractions, respectively.

Discussion

Benchmarking of MS methods for HCP quantification

The sensitivity of the Top 3-ID-DIA method was evaluated by comparing LLOQ values (obtained from established calibration curves) to the ones achieved with the gold standard ID-SRM assay (Figure S5a). Most LLOQ values determined for both methods were below 10 ppm, with minima at 0.3 ppm for ID-SRM and 0.1 ppm for ID-DIA. These sensitivities are consistent, even better, when compared to previous works that published LLOQ values of 10 ppm using 1D LC-SWATH[24] and 2D-LC MS[E23]. The accuracy of both Top 3 estimations and ID-DIA absolute quantifications simultaneously achievable with the Top 3-ID-DIA method, was assessed using pairwise comparisons to ID-SRM absolute quantifications (Figure S5b). Absolute quantifications achieved by ID-SRM and ID-DIA were all consistent within a factor of 2. Top 3 estimations presented wider errors attributable to both acquisition mode and quantification strategy changes. It is of note that 81% of the Top 3 estimations were consistent with ID-SRM quantification values within a factor of 2. These results are also in line with previous evaluations of the Top 3 estimation strategy for HCP quantification[22,24]. Finally, the precision of quantification was probed using coefficients of variation (CV) between technical triplicates (Figure S5c). All three approaches, Top 3 estimation, ID-DIA and ID-SRM displayed equivalent and good precision with a vast majority of CV values between technical triplicates below 5%.

Overall, our results demonstrate that Top 3-DIA estimations are reasonably accurate and constitute a good compromise with limited method setup requirements and limited cost, while providing a wider

view and understanding of the HCP content. Though, when accurate quantification of specific HCP of interest needs to be provided (which is certainly the case for problematic HCP such as for instance recognised immunogenic ones), the use of stable isotope dilution combined to a targeted MS assay like ID-SRM is still recommended. In the present work, we proved that combining both strategies within a single Top 3-ID-DIA method is possible without compromising performances. By spiking a reference protein and an optimised mixture of stable isotope labelled AQUA peptides corresponding to key HCP into the samples, we reached equivalent performances compared to the gold standard ID-SRM approach for a subset of HCP (ID-DIA) in addition to provide estimations of all detected HCP amounts within a single analysis. The combined Top 3-ID-DIA strategy thus constitutes a solution of choice that could be generalised in the HCP characterisation field.

Benchmarking of MS quantification against ELISA quantification

In PPA fractions, total HCP contents were quantified by the Top 3-ID-DIA approach and by a generic ELISA. Overall, the MS-HCP quantification raised on average an 8 fold higher total HCP content, which is in line with previous reports[36-38]. This can be explained by the biases intrinsic to ELISA quantification (i) the anti-HCP antibodies only detect a subset of HCP (those who elicit an immune response during ELISA development) while MS allows unbiased quantification of all detectable HCP, (ii) intracellular enzymes including proteases released at harvest (increased probability with increased shear stress[32]) may degrade HCP and thus prevent their recognition by ELISA, while degraded HCP can still be detected by MS[2-4,39], and (iii) the generic ELISA standard HCP sample that is used to generate the standard curve does not contain the same HCP population as the tested samples thus biasing the quantification. Interestingly, the ratio obtained between MS and ELISA quantifications increases when the samples diverge from a “standard” sample. The MS over ELISA amounts ratio is about 4 for “standard” samples (PPA 1, 2 and 3) generated using standard protocols, while it increases for “non-standard” samples, up to 14.4 for PPA 8 fraction. Forty-six HCP were detected uniquely in the PPA 8 fraction representing 904 ppm over a total of 5 386 ppm for 154 HCP.

This observation again argues that ELISA targets only a subset of all possible HCP. Ultimately, this argument could also be raised for our Top 3-ID-DIA approach as it is limited to HCP present in the spectral library used to extract the data. Though, building an HCP spectral library is less demanding of time and resources than developing a new ELISA assay, and improving an HCP spectral library by adding newly identified peptides is possible anytime (provided retention time standards are spiked in newly analysed samples). Thus, an HCP spectral library can be considered as an evolving resource that can be easily shared and collectively enriched to reach the largest proteome coverage of the concerned cell line. Alternatively, non-library-based algorithms are currently being developed by the computational proteomics community to interpret DIA data and, even if the results are still not reaching the quality levels of library-based approaches, the output of these solutions has recently significantly improved in terms of proteome coverage and false discovery proportions control[28,40]. HCP characterisation using MS techniques would also greatly benefit from a better curated CHO protein database[41], as the one that is currently available on public resources contains 99% unreviewed and highly redundant sequences (mostly UniProtKB-TrEMBL entries). Database redundancy is in the end the most limiting factor as only unique peptides are considered for quantification and numerous peptides are therefore unnecessarily discarded based on non-unicity criteria.

MS allows better understanding of process-related behaviours

Beyond global HCP contents estimation and unlike ELISA, the Top 3-ID-DIA approach also allowed precise identification of about 1 450 HCP in CCCF fractions and 120 HCP in PPA fractions. Precisely identifying and individually quantifying HCP is of crucial importance if one aims to understand ongoing mechanisms and eventually improve bioprocess.

For instance, global HCP contents were estimated gradually higher when cell culture fluids were exposed to low or high shear stress or when cells were cultivated for an extended duration. Both observations are in line with previous studies[11,12,29,30]. This tendency was not observed in PPA

fractions which can be tentatively explained by looking at specific HCP behaviours (Figure S6a). A gradual and strong enrichment of ribosomal proteins was observed with increasing shear stress: from 1 804 ppm without shear stress (CCCF 1) to 3 619 and 12 409 ppm with low (CCCF 2) and high (CCCF 3) shear stress, respectively. Shear stress is known to induce cell breakage and therefore intracellular content release among which ribosomal proteins are highly abundant[32]. However, these differences were not observed among PPA fractions (in which ribosomal proteins were quantified around 50 ppm regardless their originating CCCF fraction) demonstrating that the protein A purification step efficiently removed these abundant intracellular proteins.

An extended cell culture duration led to overrepresentation of heat shock proteins family in the CCCF fractions (Figure S6b): Heat shock protein (tr|A0A061ID29|A0A061ID29_CRIGR) was quantified at 24 ppm at 7 days versus 124 ppm at 10 days, Endoplasmin (tr|G3HQM6|G3HQM6_CRIGR) at 8 798 ppm at 7 days versus 16 192 ppm at 10 days, and 78 kDa glucose-regulated protein (tr|G3I8R9|G3I8R9_CRIGR) at 6 059 ppm at 7 days versus 12 009 ppm at 10 days. While Heat shock protein is totally removed (not detected) and 78 kDa glucose-regulated protein is partially removed from all PPA fractions (average 54 ppm), Endoplasmin remains more abundant in PPA fractions obtained at 10 days (on average 110 ppm in all PPA fractions from 7 days compared to 172 and 405 ppm in PPA 7 and 8 fractions from 10 days). On the contrary, several proteins are underrepresented in CCCF fractions after an extended culture duration, like Annexins (tr|A0A061IML2|A0A061IML2_CRIGR, tr|G3I5L3|G3I5L3_CRIGR, tr|G3IG05|G3IG05_CRIGR) which were quantified at 245, 797 and 324 ppm at 7 days versus 0, 51 and 18 ppm at 10 days. Annexins were efficiently removed by the protein A purification and were not detected in any PPA fraction.

Several HCP keep constant over time in CCCF fractions but are significantly more abundant in PPA fractions obtained after 10 days of culture like Pyruvate kinase (tr|A0A098KXC0|A0A098KXC0_CRIGR) which was quantified on average at 60 ppm in PPA fractions from 7 days versus 183 and 445 ppm in PPA 7 and 8 fractions from 10 days, 6-phosphogluconate

dehydrogenase, decarboxylating (tr|G3IHY5|G3IHY5_CRIGR) which was not detected in PPA fractions except in PPA 7 and 8 fractions from 10 days at 32 and 67 ppm, respectively, Heat shock cognate 71 kDa protein (sp|P19378|HSP7C_CRIGR) which was quantified on average at 21 ppm in PPA fractions from 7 days versus 94 and 250 ppm in PPA 7 and 8 fractions from 10 days, respectively. These behaviours are not easy to understand but one could hypothesise that the presence of specific HCP after 10 days in the CCCF fraction could help other HCP to co-purify with the mAb. Thereby, an HCP that is known to be easily removed could become challenging in the presence of certain cofactors.

Finally, the modified protein A purification protocol that we evaluated allowed a more efficient removal of some HCP when compared to the standard purification protocol (Figure S6c) : Heterogeneous nuclear ribonucleoprotein U-like protein 1 (tr|G3IA10|G3IA10_CRIGR) was quantified on average at 298 ppm in PPA fractions obtained using the standard purification protocol, but at 59 and 14 ppm in PPA 4 and 8 fractions obtained using the modified purification protocol; Putative phospholipase B-like 2 (tr|G3I6T1|G3I6T1_CRIGR) was quantified on average at 62 ppm in PPA fractions obtained using the standard purification protocol, but at only 12 ppm using the modified purification protocol. The absolute quantification of Phospholipase B-like 2, or PLBL-2, by ID-SRM and ID-DIA confirmed the Top 3-DIA estimations as shown in Figure 3. This is particularly interesting since PLBL-2 is known for its immunogenicity[5,6] and currently constitutes a major purification challenge[42].

More generally, it becomes obvious that after the protein A purification step, downstream purification process has to face with co-purifying HCP that either specifically bind to the mAb[19,43-46] or bind to protein A and are co-eluted, as it was demonstrated for chromatin-bound histone impurities[47,48]. Though published evidence using Null cell lines indicates, the presence of mAb is the critical factor to seeing elevated HCP levels[26,44]. In this context also, the specific identification of “difficult to remove” HCP by MS and the consecutive development of robust targeted quantification methods constitute tools of choice to help in designing an appropriate HCP clearance strategy.

Altogether, these examples show that MS-based quantification approaches, and especially the proposed combined Top 3-ID-DIA approach, provide additional valuable information to ELISA: HCP contents can be precisely monitored in a more comprehensive manner with high throughput, and can lead process development to release cleaner and safer products.

Supporting Information Available

Supporting Information is available free of charge via http://pubs.acs.org/.

It includes detailed experimental procedures, complete lists of quantified proteins, calibration curves and a series of graphical views of the quantification results.

Notes

The authors declare no competing financial interests.

Acknowledgements

This project was supported by the «Association Nationale de la Recherche et de la Technologie » and UCB Pharma via the CIFRE fellowship of GH. This work was also supported by the “Agence Nationale de la Recherche” (ANR) and the French Proteomic Infrastructure (ProFI; ANR-10-INBS-08-03).

References

(1) Bracewell, D. G.; Francis, R.; Smales, C. M. *Biotechnol Bioeng* **2015**, *112*, 1727-1737.
(2) Gao, S. X.; Zhang, Y.; Stansberry-Perkins, K.; Buko, A.; Bai, S.; Nguyen, V.; Brader, M. L. *Biotechnol Bioeng* **2011**, *108*, 977-982.
(3) Bee, J. S.; Tie, L.; Afdahl, C. D.; Jusino, K. C.; Johnson, D.; Dimitrova, M. N. *Biotechnol Prog* **2015**.
(4) Robert, F.; Bierau, H.; Rossi, M.; Agugiaro, D.; Soranzo, T.; Broly, H.; Mitchell-Logean, C. *Biotechnol Bioeng* **2009**, *104*, 1132-1141.
(5) Hanania, N. A.; Noonan, M.; Corren, J.; Korenblat, P.; Zheng, Y.; Fischer, S. K.; Cheu, M.; Putnam, W. S.; Murray, E.; Scheerens, H.; Holweg, C. T.; Maciuca, R.; Gray, S.; Doyle, R.; McClintock, D.; Olsson, J.; Matthews, J. G.; Yen, K. *Thorax* **2015**, *70*, 748-756.
(6) Fischer, S. K.; Cheu, M.; Peng, K.; Lowe, J.; Araujo, J.; Murray, E.; McClintock, D.; Matthews, J.; Siguenza, P.; Song, A. *The AAPS journal* **2017**, *19*, 254-263.
(7) ICH, Guidance for Industry Q6B Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological products; 1999.
(8) Chon, J. H.; Zarbis-Papastoitsis, G. *N Biotechnol* **2011**, *28*, 458-463.
(9) Tscheliessnig, A. L.; Konrath, J.; Bates, R.; Jungbauer, A. *Biotechnol J* **2013**, *8*, 655-670.
(10) Wang, X.; Hunter, A. K.; Mozier, N. M. *Biotechnol Bioeng* **2009**, *103*, 446-458.
(11) Hogwood, C. E.; Bracewell, D. G.; Smales, C. M. *Curr Opin Biotechnol* **2014**, *30C*, 153-160.
(12) Jin, M.; Szapiel, N.; Zhang, J.; Hickey, J.; Ghose, S. *Biotechnol Bioeng* **2010**, *105*, 306-316.
(13) Zhu-Shimoni, J.; Yu, C.; Nishihara, J.; Wong, R. M.; Gunawan, F.; Lin, M.; Krawitz, D.; Liu, P.; Sandoval, W.; Vanderlaan, M. *Biotechnol Bioeng* **2014**, *111*, 2367-2379.
(14) Picotti, P.; Aebersold, R. *Nat Methods* **2012**, *9*, 555-566.
(15) Gallien, S.; Duriez, E.; Crone, C.; Kellmann, M.; Moehring, T.; Domon, B. *Molecular & cellular proteomics : MCP* **2012**, *11*, 1709-1723.
(16) Gillet, L. C.; Navarro, P.; Tate, S.; Rost, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R. *Molecular & cellular proteomics : MCP* **2012**, *11*, O111 016717.
(17) Silva, J. C.; Gorenstein, M. V.; Li, G. Z.; Vissers, J. P.; Geromanos, S. J. *Molecular & cellular proteomics : MCP* **2006**, *5*, 144-156.
(18) Doneanu, C. E.; Anderson, M.; Williams, B. J.; Lauber, M. A.; Chakraborty, A.; Chen, W. *Analytical chemistry* **2015**, *87*, 10283-10291.
(19) Zhang, Q.; Goetze, A. M.; Cui, H.; Wylie, J.; Tillotson, B.; Hewig, A.; Hall, M. P.; Flynn, G. C. *Biotechnol Prog* **2016**, *32*, 708-717.
(20) Farrell, A.; Mittermayr, S.; Morrissey, B.; McLoughlin, N.; Navas Iglesias, N.; Marison, I. W.; Bones, J. *Analytical chemistry* **2015**.
(21) Doneanu, C. E.; Xenopoulos, A.; Fadgen, K.; Murphy, J.; Skilton, S. J.; Prentice, H.; Stapels, M.; Chen, W. *mAbs* **2012**, *4*, 24-44.
(22) Schenauer, M. R.; Flynn, G. C.; Goetze, A. M. *Anal Biochem* **2012**, *428*, 150-157.
(23) Zhang, Q.; Goetze, A. M.; Cui, H.; Wylie, J.; Trimble, S.; Hewig, A.; Flynn, G. C. *mAbs* **2014**, *6*, 659-670.
(24) Walker, D. E.; Yang, F.; Carver, J.; Joe, K.; Michels, D. A.; Yu, X. C. *mAbs* **2017**, *9*, 654-663.
(25) Liu, H. F.; Ma, J.; Winter, C.; Bayer, R. *mAbs* **2010**, *2*, 480-499.
(26) Shukla, A. A.; Hinckley, P. *Biotechnol Prog* **2008**, *24*, 1115-1121.
(27) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. *Bioinformatics* **2010**, *26*, 966-968.
(28) Navarro, P.; Kuharev, J.; Gillet, L. C.; Bernhardt, O. M.; MacLean, B.; Rost, H. L.; Tate, S. A.; Tsou, C. C.; Reiter, L.; Distler, U.; Rosenberger, G.; Perez-Riverol, Y.; Nesvizhskii, A. I.; Aebersold, R.; Tenzer, S. *Nat Biotechnol* **2016**.
(29) Tait, A. S.; Hogwood, C. E.; Smales, C. M.; Bracewell, D. G. *Biotechnol Bioeng* **2012**, *109*, 971-982.
(30) Grzeskowiak, J. K.; Tscheliessnig, A.; Toh, P. C.; Chusainow, J.; Lee, Y. Y.; Wong, N.; Jungbauer, A. *Protein Expr Purif* **2009**, *66*, 58-65.

(31) Hogwood, C. E.; Ahmad, S. S.; Tarrant, R. D.; Bracewell, D. G.; Smales, C. M. *Biotechnol J* **2016**, *11*, 415-424.
(32) Lau, E. C.; Kong, S.; McNulty, S.; Entwisle, C.; McIlgorm, A.; Dalton, K. A.; Hoare, M. *Biotechnol Bioeng* **2013**, *110*, 1973-1983.
(33) Boychyn, M.; Yim, S. S.; Bulmer, M.; More, J.; Bracewell, D. G.; Hoare, M. *Bioprocess Biosyst Eng* **2004**, *26*, 385-391.
(34) Elias, J. E.; Gygi, S. P. *Nat Methods* **2007**, *4*, 207-214.
(35) FDA. *US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research: Rockville, MD* **2013**.
(36) Kreimer, S.; Gao, Y.; Ray, S.; Jin, M.; Tan, Z.; Mussa, N. A.; Tao, L.; Li, Z.; Ivanov, A. R.; Karger, B. L. *Analytical chemistry* **2017**.
(37) Farrell, A.; Mittermayr, S.; Morrissey, B.; Mc Loughlin, N.; Navas Iglesias, N.; Marison, I. W.; Bones, J. *Analytical chemistry* **2015**.
(38) Henry, S. M.; Sutlief, E.; Salas-Solano, O.; Valliere-Douglass, J. *mAbs* **2017**, 1-11.
(39) Carter-Franklin, J. N.; Victa, C.; McDonald, P.; Fahrner, R. *J Chromatogr A* **2007**, *1163*, 105-111.
(40) Tsou, C. C.; Avtonomov, D.; Larsen, B.; Tucholska, M.; Choi, H.; Gingras, A. C.; Nesvizhskii, A. I. *Nat Methods* **2015**, *12*, 258-264, 257 p following 264.
(41) Meleady, P.; Hoffrogge, R.; Henry, M.; Rupp, O.; Bort, J. H.; Clarke, C.; Brinkrolf, K.; Kelly, S.; Muller, B.; Doolan, P.; Hackl, M.; Beckmann, T. F.; Noll, T.; Grillari, J.; Barron, N.; Puhler, A.; Clynes, M.; Borth, N. *Biotechnol Bioeng* **2012**, *109*, 1386-1394.
(42) Tran, B.; Grosskopf, V.; Wang, X.; Yang, J.; Walker, D., Jr.; Yu, C.; McDonald, P. *J Chromatogr A* **2016**, *1438*, 31-38.
(43) Levy, N. E.; Valente, K. N.; Choe, L. H.; Lee, K. H.; Lenhoff, A. M. *Biotechnol Bioeng* **2014**, *111*, 904-912.
(44) Tarrant, R. D.; Velez-Suberbie, M. L.; Tait, A. S.; Smales, C. M.; Bracewell, D. G. *Biotechnol Prog* **2012**, *28*, 1037-1044.
(45) Sisodiya, V. N.; Lequieu, J.; Rodriguez, M.; McDonald, P.; Lazzareschi, K. P. *Biotechnology Journal* **2012**, *7*, 1233-1241.
(46) Nogal, B.; Chhiba, K.; Emery, J. C. *Biotechnol Prog* **2012**, *28*, 454-458.
(47) Nian, R.; Zhang, W.; Tan, L.; Lee, J.; Bi, X.; Yang, Y.; Gan, H. T.; Gagnon, P. *J Chromatogr A* **2016**, *1431*, 1-7.
(48) Gagnon, P.; Nian, R.; Lee, J.; Tan, L.; Latiff, S. M.; Lim, C. L.; Chuah, C.; Bi, X.; Yang, Y.; Zhang, W.; Gan, H. T. *J Chromatogr A* **2014**, *1340*, 68-78.

Figure legends

Figure 1: Overview of MS-based quantification strategies developed for HCP monitoring.

Figure 2: Estimations of global HCP contents using Top 3-ID-DIA (and ELISA quantification for PPA fractions). Quantification was performed using the 3 best responding peptides per protein relative to a standard spiked protein (20 fmol PYGM), deriving ppm values and summing all proteins amounts to obtain a total HCP content in ppm. ELISA quantification was obtained using a generic ELISA kit according to the manufacturer's protocol.

Figure 3: Quantification results obtained for Phospholipase B-like 2 protein in PPA fractions. PPA 1, 2, 3, 5, 6 and 7 fractions were obtained with the standard purification protocol, while PPA 4 and 8 fractions were obtained with the modified purification protocol.

Figure 1

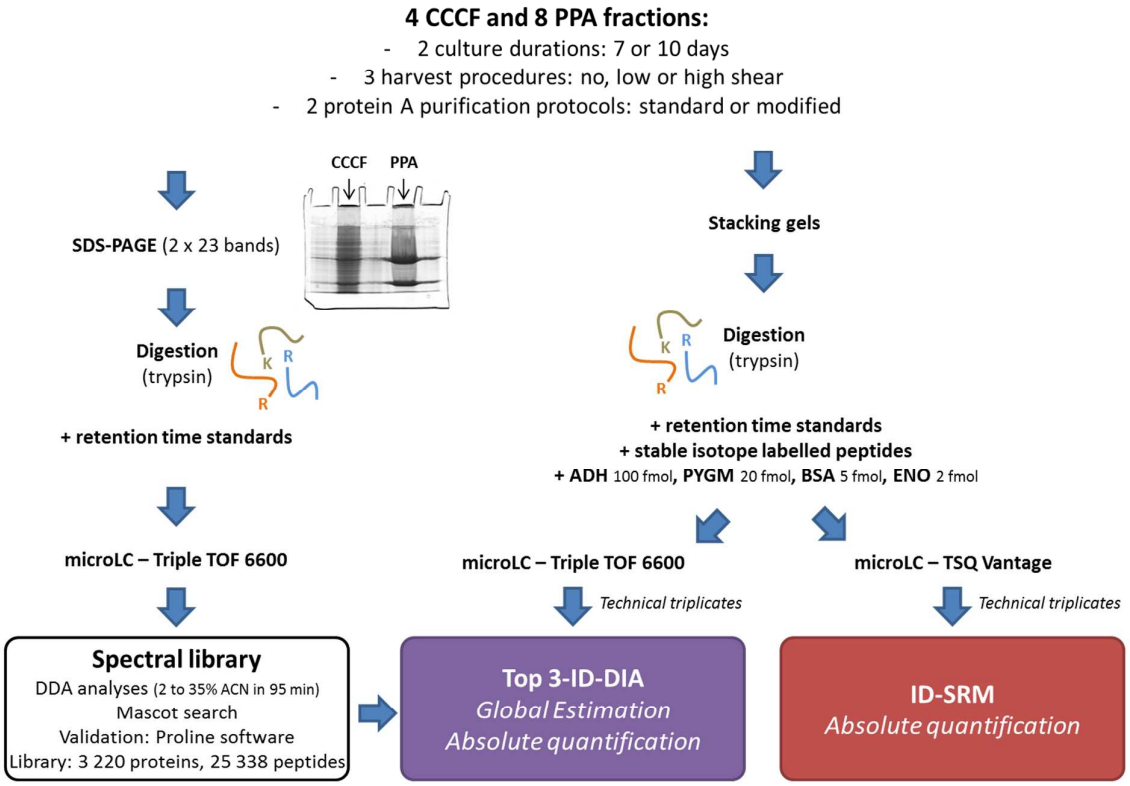

4 CCCF and 8 PPA fractions:
- 2 culture durations: 7 or 10 days
- 3 harvest procedures: no, low or high shear
- 2 protein A purification protocols: standard or modified
CCCF
PPA
SDS-PAGE (2 x 23 bands)
Digestion (trypsin)
+ retention time standards
microLC – Triple TOF 6600
Spectral library
DDA analyses (2 to 35% ACN in 95 min)
Mascot search
Validation: Proline software
Library: 3 220 proteins, 25 338 peptides
Stacking gels
Digestion (trypsin)
+ retention time standards
+ stable isotope labelled peptides
+ ADH 100 fmol, PYGM 20 fmol, BSA 5 fmol, ENO 2 fmol
microLC – Triple TOF 6600
Technical triplicates
Top 3-ID-DIA
Global Estimation
Absolute quantification
microLC – TSQ Vantage
Technical triplicates
ID-SRM
Absolute quantification

Figure 2

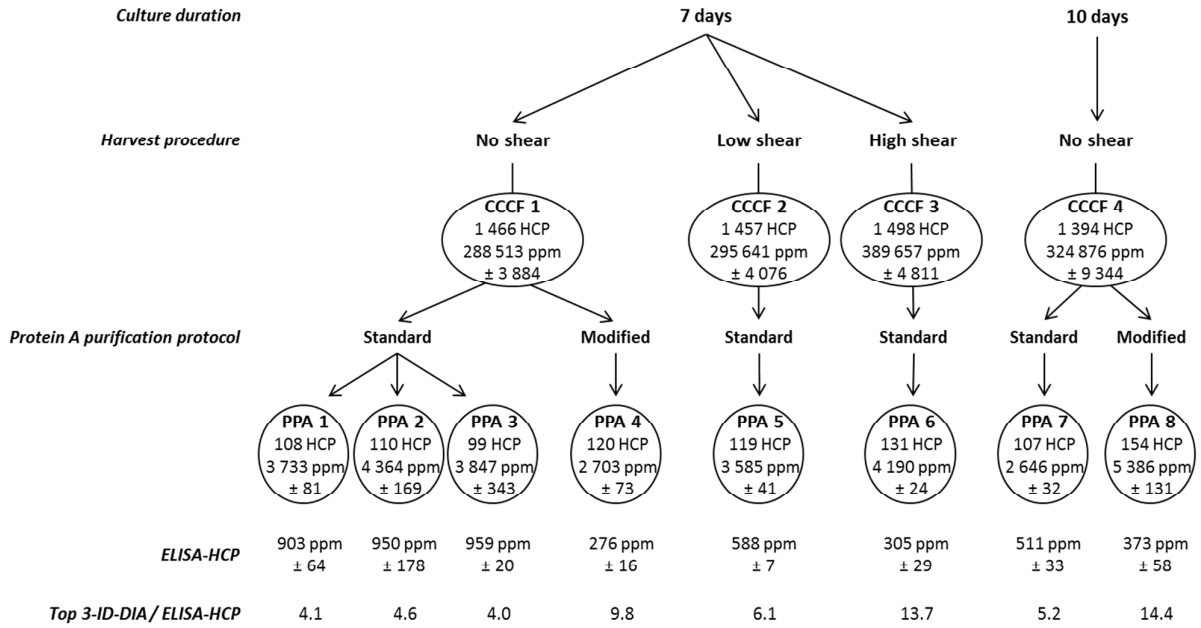
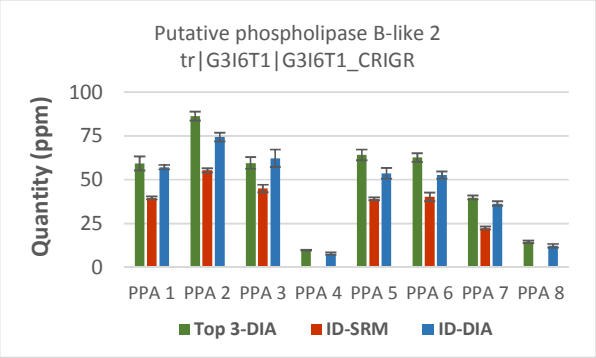
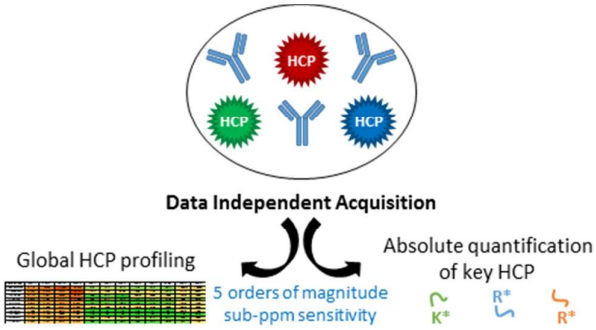

Culture duration
7 days
10 days
Harvest procedure
No shear
Low shear
High shear
No shear
CCCF 1
1 466 HCP
288 513 ppm
± 3 884
CCCF 2
1 457 HCP
295 641 ppm
± 4 076
CCCF 3
1 498 HCP
389 657 ppm
± 4 811
CCCF 4
1 394 HCP
324 876 ppm
± 9 344
Protein A purification protocol
Standard
Modified
Standard
Standard
Standard
Modified
PPA 1
108 HCP
3 733 ppm
± 81
PPA 2
110 HCP
4 364 ppm
± 169
PPA 3
99 HCP
3 847 ppm
± 343
PPA 4
120 HCP
2 703 ppm
± 73
PPA 5
119 HCP
3 585 ppm
± 41
PPA 6
131 HCP
4 190 ppm
± 24
PPA 7
107 HCP
2 646 ppm
± 32
PPA 8
154 HCP
5 386 ppm
± 131
ELISA-HCP
903 ppm ± 64
950 ppm ± 178
959 ppm ± 20
276 ppm ± 16
588 ppm ± 7
305 ppm ± 29
511 ppm ± 33
373 ppm ± 58
Top 3-ID-DIA / ELISA-HCP
4.1
4.6
4.0
9.8
6.1
13.7
5.2
14.4

Figure 3


Putative phospholipase B-like 2
tr|G3I6T1|G3I6T1_CRIGR
Quantity (ppm)
100
75
50
25
0
PPA 1
PPA 2
PPA 3
PPA 4
PPA 5
PPA 6
PPA 7
PPA 8
Top 3-DIA
ID-SRM
ID-DIA

For TOC only


HCP
HCP
HCP
Data Independent Acquisition
Global HCP profiling
5 orders of magnitude
sub-ppm sensitivity
Absolute quantification
of key HCP
K*
R*
R*