Ana Rebeca Oliveira Ramos

Degree in Forensic and Criminal Sciences

Development of an Analytical Toolbox for the Assessment of Complex Biologics

Dissertation to obtain a Master's degree in
Biochemistry for Health

Supervisor: Patrícia Gomes Alves, PhD, Senior Scientist, iBET/ITQB-NOVA
Co-supervisor: Ricardo Gomes, PhD, Senior Scientist, iBET/ITQB-NOVA

March 2022

Ana Rebeca Oliveira Ramos

Degree in Forensic and Criminal Sciences

Development of an Analytical Toolbox for the Assessment of Complex Biologics

Dissertation to obtain a Master's degree in
Biochemistry for Health

Supervisor: Patrícia Gomes Alves, PhD, Senior Scientist, iBET/ITQB-NOVA
Co-supervisor: Ricardo A. Gomes, PhD, Senior Scientist, iBET/ITQB-NOVA

Jury:

President: Dr Pedro Manuel Henriques Marques Matias
Arguer: Dr Romana Lopes Almeida dos Santos
Jury member: Dr Patrícia Isabel Gomes Alves
Dr Maria Teresa Nunes Mangas Catarino

Instituto de Tecnologia Química e Biológica António Xavier

March 2022

Copyright

Development of an Analytical Toolbox for the Assessment of Complex Biologics

Ana Rebeca Oliveira Ramos, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa

Acknowledgments

Em primeiro lugar quero agradecer à minha família, o meu mais profundo obrigada por me proporcionarem a continuidade dos meus estudos, pela dedicação ao longo destes anos e por me mostrarem que eu conseguia sempre ir mais longe, com o vosso apoio. Em especial quero agradecer à minha mãe, por nunca me permitir desistir dos meus sonhos e acreditar nas minhas capacidades. Obrigada por segurares o “barco” sempre que existem tempestades e por seres um exemplo para mim.

Aos meus amigos Sofia, Bruno e João, pela amizade e incentivo, nos momentos de incerteza, assim como nos momentos de vitória, acreditaram sempre em mim, nas minhas capacidades e motivaram-me sempre a continuar.

Às minhas colegas e amigas do curso, em especial à Rafaela e à Maria Beatriz, pela amizade, pela partilha e encorajamento mútuo face às adversidades e pela constante motivação e suporte nos momentos difíceis.

I would like to acknowledge all the people that directly or indirectly supported and contributed to the work developed and presented in this MSc thesis.

To Professor Paula Alves and Professor Cláudio Soares for giving me the opportunity to do my master thesis at the Mass Spectrometry Unit at iBET/ITQB.

In particular to Professor Paula Alves for offering the best circumstances to work and for being an example of leadership and professionalism.

To my supervisors Dr. Patrícia Alves and Dr. Ricardo Gomes, for the scientific and technical guidance, precious support, motivation and encouragement. I am thankful for that.

To my colleagues at the UNIMS for the support, encouragement and recommendations.

To the Sanofi satellite laboratory members for all that I have learned and the disposition to help, especially to Dr. Sofia Carvalho for always being there to help whenever I needed, for her confidence, the support and friendship.

To all the professors that lecture in the MSc in Biochemistry for Health at UNL for the precious knowledge they shared with their students.

Obrigada!

Preface

The following work was developed at the UniMS, the mass spectrometry unit of iBET and ITQB-UNL within the scope of internal projects and iNOVA4Health – UIDB/04462/2020 and UIDP/04462/2020, a program financially supported by Fundação para a Ciência e Tecnologia / Ministério da Ciência, Tecnologia e Ensino Superior, through national funds.

Part of this work has been included in two abstracts submitted for the 27th ESACT Meeting "Advanced Cell Technologies: Making Protein, Cell, and Gene Therapies a Reality":

- Ricardo A. Gomes, Bruno M. Alexandre, Rebeca Ramos, Sofia B. Carvalho and Patrícia Gomes-Alves, ***Mass spectrometry analytical toolbox for the comprehensive characterization of biotherapeutics*** *(submitted for oral presentation);*

- Sofia B. Carvalho, Ricardo A. Gomes, Bruno M. Alexandre, Rebeca Ramos, Anja Pfenninger, Martina Fischer, Michael Hoffman, Kevin Brower, Nihal Tugçu and Patrícia Gomes-Alves, ***SWATH-MS as a strategy for host cell protein identification and quantification to support bioprocess development*** *(submitted for poster presentation).*

Abstract

The biopharmaceutical industry is expanding rapidly and currently accounts for about 20% of the pharmaceutical market. Clinical and manufacturing processes for biologics, such as recombinant proteins or therapeutic monoclonal antibodies (mAbs), are lengthy, complex, and must meet safety and efficacy standards. As the complexity of the therapeutic products increases, the characterization becomes even more challenging, requiring the development of advanced analytical methodologies with increased sensitivity, throughput, and time/cost-effectiveness. Product heterogeneity (such as glycoforms, post-translational modifications, and sequence variants), and process-related impurities such as host cell proteins (HCP), are key product attributes that need to be monitored. ELISA-based HCP assays, commonly used for total HCP quantification, present several drawbacks, namely, not identifying individual HCPs, nor distinguishing their relative abundances.

This work aimed at developing reliable and high-throughput MS-based tools for HCP characterization of in-process mAb samples. Three steps were envisioned to accomplish this work: (1) total HCP quantification by ELISA (the current gold standard method); (2) development of an absolute quantification LC/MS-based method for 9 target HCP, using the TripleTOF 6600 platform (Sequential Window Acquisition of All Theorical Mass Spectra – Mass Spectrometry, SWATH-MS) and evaluate the applicability of the average curve generated to calculate the concentration of other HCP (untargeted approach); (3) implementation of an alternative targeted quantification method using another mass spectrometer system (QTRAP 6500+, the gold standard for MS-based absolute quantification of proteins), for results confirmation and using HCP A as a proof of concept.

In this thesis, we explored SWATH-MS as an orthogonal approach to traditional ELISA, providing critical information on individual HCP identity and quantity. Overall, we have developed a LC-MS approach for absolute quantification of HCP, either using target HCP's specific heavy peptides or an average curve with several heavy peptides, enabling comprehensive characterization of different HCP profiles in complex mAb samples.

Keywords: monoclonal antibodies; host cell protein; SWATH-MS; absolute quantification; average curve.

Resumo

A indústria biofarmacêutica está a expandir-se rapidamente e representa atualmente cerca de 20% do mercado farmacêutico. Os processos clínicos e de fabrico de produtos biológicos, como proteínas recombinantes ou anticorpos monoclonais terapêuticos (mAbs), são longos e complexos e têm que cumprir os padrões estritos de segurança e eficácia. À medida que a complexidade dos produtos terapêuticos aumenta, a caracterização torna-se ainda mais desafiante, exigindo o desenvolvimento de metodologias analíticas avançadas com maior sensibilidade, produtividade e rendimento (em termos de tempo e/ou custos). A heterogeneidade dos produtos (como glicoformas, modificações pós-traducionais e variantes de sequência) e impurezas relacionadas com o processo, como as proteínas provenientes da célula hospedeira (HCP, do inglês *host cell proteins*) são atributos-chave do produto que precisam de monitorização. A análise de HCPs baseada no método de ELISA, usado habitualmente para a quantificação de HCPs totais, apresenta várias desvantagens, não identificando HCPs individualmente, nem distinguindo a abundância relativa das diferentes HCPs.

O objetivo deste trabalho foi desenvolver ferramentas analíticas robustas e *high-throughput* baseadas em MS para a caracterização de HCPs em amostras de processo de mAbs. A realização deste trabalho foi repartida em três passos: (1) quantificação de HCP totais por ELISA (o método de referência atual); (2) desenvolvimento de um método de quantificação absoluta baseado em LC-MS para 9 HCPs alvo, utilizando a plataforma TripleTOF 6600 (SWATH-MS do inglês *Sequential Window Acquisition of All Theorical Mass Spectra – Mass Spectrometry*) e avaliação da aplicabilidade de uma curva de calibração média para calcular a concentração de outras HCP (abordagem não direcionada); (3) implementação de um método alternativo de quantificação absoluta (direcionada para uma proteína alvo) usando outro espetrómetro de massa (QTRAP 6500+, método de referência para quantificação absoluta de proteínas), para confirmação dos resultados e usando a proteína HCP A como caso de estudo.

Nesta tese, explorámos o SWATH-MS como uma abordagem ortogonal ao ELISA, fornecendo informações críticas sobre a identidade e a quantidade de cada HCP individual. No geral, desenvolvemos um método LC-MS para a quantificação absoluta de HCP, usando curvas de calibração individuais com péptidos marcados específicos da HCP alvo ou uma curva de calibração média (com todos os péptidos), permitindo uma caracterização mais abrangente de diferentes HCPs em amostras complexas.

Palavras-chave: anticorpos monoclonais; proteína da célula hospedeira; SWATH-MS; quantificação absoluta; curva média.

List of Contents

List of Figures

List of Tables

List of Appendix

Abbreviations

4PL	Four-parameter logistic
ACN	Acetonitrile
AmBic	Ammonium bicarbonate
CE	Collision energy
CHO	Chinese hamster ovary
CSH	Charged surface hybrid
CV	Coefficient of variation
CXP	Collision cell exit potential
DIA	Data-independent acquisition
DSP	Downstream process
DTT	Dithiothreitol
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
FA	Formic acid
FDR	False discovery rate
HCP	Host Cell Protein
HRP	Horseradish peroxidase
IAA	Iodoacetamide
IDA	Information-dependent acquisition
IPA	2-Propanol
IS	Internal Standard
IT	Ion trap
LC	Liquid chromatography
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LLOQ	Lower limit of quantification
m/z	Mass-to-charge ratio
mAb	Monoclonal antibody
MRD	Minimum required dilution
MRM	Multiple reaction monitoring
MS	Mass spectrometry
PPI	Protein-protein interactions
PRM	Parallel reaction monitoring
PTFE	Polytetrafluoroethylene
PTMs	Post-translational modifications
Q	Quadrupole
Q1/Q3	First/third quadrupole
R^2	Coefficient of determination
rcf	Relative centrifugal force

RS	Reference standard
RT	Room temperature
S/N	Signal-to-noise ratio
SDB	Sample diluent buffer
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SF	Surfactant
SIL	Stable-isotope-labeled
SRM	Single reaction monitoring
SST	System suitability test
SWATH-MS	Sequential Window Acquisition of All Theorical Mass Spectra – Mass Spectrometry
T	Triple
TFA	Trifluoroacetic acid
TIC	Total ion chromatogram
TMB	Tetramethylbenzidine
TOF	Time-of-flight
UPLC	Ultra-performance liquid chromatography
WB	Western blot
XIC	Extracted-ion chromatogram

1. Introduction

1.1. Monoclonal antibodies characterization

Monoclonal antibodies are one of the most important products in the biopharmaceutical industry and the most popular class of biopharmaceuticals. mAbs are highly selective and successful in treating a variety of diseases, including cancer and infectious diseases, and in modulating the immune response [1]. mAbs are used to recognize antigens specific to cancerous tissue, to direct the action of immune cells or complement, to deliver cytotoxic drugs or radioisotopes, or to block the activity of tumor-promoting agents [2].

Nowadays, the production of mAbs is based on bioreactor culture of producer clones derived mainly from Chinese Hamster Ovary cells (CHO). The clones are genetically engineered to produce mAbs with a specific antigen specificity. There is an increasing pressure to achieve high-quality mAb products and, at the same time, increased cost-effectiveness, and high productivity processes [2,3]. To achieve this, manufacturers must demonstrate not only product quality but also clearance of host cell impurities and contaminants during process development and production, to ensure drug purity and efficacy, manufacturing process consistency, and ultimately patient safety [4].

Given this, a detailed characterization of mAb products is of utmost importance, not only prior to use in clinical trials but also during product development. However, the comprehensive characterization of mAbs is still an analytical challenge because these products are inherently heterogeneous due to the presence of post-translational modifications (PTMs), incomplete processing, susceptibility to degradation, and disulfide shuffling [5]. PTMs, including oxidation, deamidation, glycosylation, glycation, proteolysis, and process-related impurities such as host cell proteins (HCP), can significantly alter the efficacy and safety of the therapeutic product. Therefore, regulatory agencies require a complete characterization of the structural, biological, and chemical heterogeneity of any biopharmaceutical [6].

A collection of orthogonal analytical methods and separation techniques are routinely used to monitor the quality of mAbs and to identify product variants and impurities. These include enzyme-linked immunosorbent assay (ELISA), sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE), western blot (WB), capillary electrophoresis and chromatographic separation techniques such as ion exchange, reverse phase, size exclusion, or hydrophobic interaction [5]. Most of these methodologies are only able to evaluate one of the product's characteristics and may not present the necessary sensitivity and specificity for such a complex and heterogeneous product. Depending on the method or conjugation of methods applied, the characterization workflow can be time-consuming, require instrument adjustments or specific sample preparation and specialized data analysis. All these variables make direct and rapid real-time product quality measurements extremely difficult. Moreover, these methods are only able to monitor product quality attributes (PQA) at the intact protein level and not at the amino acid level. In this context, mass spectrometry (MS) emerged as one of the most powerful analytical tools

available, and several MS-based workflows have been used to monitor critical quality attributes (CQA) of mAb products [7,8]. The selectivity, specificity, sensitivity, dynamic range, mass accuracy, and resolving power of modern MS instruments, as well as the ability to be coupled with various separation methods, make this analytical tool an invaluable asset for the qualitative and quantitative analysis of biologics. By combining powerful and robust separation technologies with exceptional selectivity based on precise mass accuracy and high resolution, MS can provide an assessment of multiple attributes in a short-time frame, enabling rapid and direct measurement of product quality attributes in complex biologics.

1.2. Host Cell Proteins

Recombinant proteins and other biopharmaceuticals are produced by living organisms. Therefore, the biologic product itself must be purified from any HCP and culture media impurities during the manufacturing process, as referred above [9]. HCPs are proteins of the host cells, unrelated to the target recombinant product, that are co-purified with it. They can compromise patient safety and product purity and quality in several ways, such as potential immunogenicity, catalytic activity for product fragmentation, and involvement in product aggregation. The protease activity of some HCPs can influence the protein sequence/structure in the culture supernatant, affect the purification process and the stability/efficacy of the product. In addition, the presence of several HCP leads to different biochemical properties and different reactions in the human body, as some HCPs may act as adjuvants and induce an immune response against the product itself or cause allergic reactions, including anaphylactic shock [10]. Hence, HCPs are identified as a CQA [9] and the identification and levels of residual HCP in the final mAb formulations is a significant area of interest for mAbs' bioprocessing. Despite this, there are no universally accepted unbiased methods for the characterization and quantification of HCPs. Moreover, the limits of residual HCPs are also not properly defined, 100ng/mg therapeutic protein being the "conventionally" acceptable value [10]. HCPs are a heterogeneous, complex group of proteins that differ significantly in molecular mass, structure, hydrophobic properties, and isoelectric point. Like any other proteins, HCPs can also undergo PTMs, which complicates their quantification and characterization [11]. Additionally, the amount and identity of HCPs are affected by the expression system (intracellular/secreted protein, host, ...) of the recombinant product [12]. In some cases (when the recombinant product is not secreted), product recovery requires cell disruption, resulting in significantly increased HCP content that requires an even more effective HCP removal technique to ensure product quality and safety. All the above-mentioned highlights the importance of HCPs monitoring during the mAb bioprocessing, from the upstream (production of the mAb) to downstream processes (purification of the mAb) [10].

The selection of appropriate methods for the identification and/or quantification of HCPs is of utmost importance in process development and manufacturing. When developing and characterizing a process, the ideal method (or combination of methods) should detect all HCP species and provide a quantitative response, in a short analysis time, with high throughput.

1.2.1. HCP Identification and Quantification

Currently, HCPs can be detected and characterized by several methods that can be classified as either immunospecific, for example, ELISA, western blotting and microarrays, or non-immunospecific, such as electrophoresis (1 or 2-dimensional gel separation followed by protein staining) or liquid chromatography (LC) followed by MS [10,13]. The most common and accepted method is the ELISA. However, more recently, MS-based approaches are gaining significant relevance as a key orthogonal method able to provide unbiased identification and quantification of HCPs. Over the last decades, liquid chromatography in combination with tandem MS (LC-MS/MS) has become the preferred technology for the high-throughput characterization of proteins and proteomes (proteome is a set of proteins produced in an organism, system, or biological context) [13,14].

1.2.1.1. ELISA

ELISA is an analytical biochemistry assay (described for the first time in 1971 by Engvall and Perlmann) that is traditionally used for the detection and quantification of a specific protein in a complex mixture [15]. The use of high-affinity antibodies and the possibility to wash away non-specific bound materials (antibodies are immobilized together with the targeted antigen in the well) makes ELISA a powerful tool for measuring specific analytes within a crude preparation.

There are several formats of ELISA assay: direct, indirect, sandwich or competitive ELISA [16]. The key step is the immobilization of the antigen of interest, which is obtained by direct adsorption to the assay plate or indirectly via a capture antibody that has been attached to the plate. The antigen is then detected directly (labeled primary antibody) or indirectly (labeled secondary antibody). The most used ELISA assay format is the sandwich ELISA. This format is highly used because of its sensitivity and specificity [16]. In this format, a target protein-specific antibody is applied to the surface of the plate wells and incubated first with the target protein and then with another target protein-specific antibody labeled with an enzyme. The protein to be measured is bound between the two antibodies (capture and detection antibodies). Different epitopes of the target protein must be recognized by the immobilized antibody (orange) and the enzyme-labeled antibody (green) (**Figure 1.1**). The activity of the enzyme bound to the "detection antibody" is assayed after adding the adequate substrate. The light absorption of the product after the addition of the substrate is measured and converted to a number (correlating with the target protein concentration) [17]

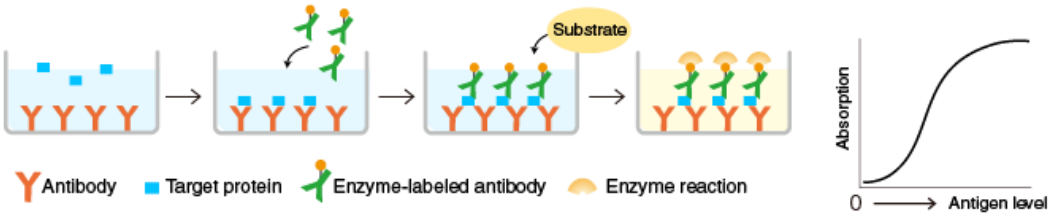


Figure 1.1 – Sandwish ELISA. Adapted from: [16].

In particular, ELISA formats to detect HCPs use polyclonal antibodies usually generated in goats or rabbits from immunization of the host animal with an HCP pool (cell lysate or partially purified material) [18]. This method is considered the gold standard for ”total” HCP quantification due to its ease to use, semi-quantitative nature and high sensitivity (1-100 ppm). Nevertheless, it presents several limitations: i) HCP detection is not unbiased, since it does not cover all process/host-specific HCP, namely the non-immunoreactive or weakly immunoreactive proteins present in the animal used to generate the polyclonal antibodies. Even with the most recent “generic” ELISA kits, focused on ensuring that low molecular weight low abundance proteins and less immunogenic impurities are represented, process-specific HCPs ELISA often needs to be developed. ii) ELISA provides quantitative data of ”total” HCP amount but does not define the identity of specific HCPs. This is a major drawback for two main reasons: identification is key to define the purification strategy to further remove a particular and critical HCP (or set of HCPs), and the product-related risks associated with the presence of certain HCPs (and HCP levels) are already known/or need to be investigated. Given this, different HCPs are not considered to be of equal importance and the desired quantification range limit required varies accordingly to HCPs effect on product quality and patient safety [19,20].

1.2.1.2. Mass spectrometry

Alternative and more advanced orthogonal methods that adopt a proteomic approach such as LC coupled with MS are capable of both identification and quantification of specific HCPs. This will probably provide the most powerful method for HCPs profiling. MS is emerging as the leading analytical technology for the detection and quantification of low abundance HCPs with high confidence in the presence of the recombinant product (highly abundant). Immunogenic and problematic HCPs can also be targeted in highly purified samples with high sensitivity once the MS protocol is established. Nevertheless, the adoption of this approach may also be impaired by some inherent drawbacks: highly skilled personnel are required to operate LC-MS/MS; the equipment is expensive, and absolute quantification of individual HCPs requires the use of synthetic peptides, which are costly and labor-intensive [21,22]

MS-based proteomics takes advantage of the availability of annotated genomes and protein sequence databases and also the technological advances. MS consists of ionization and vaporization of molecules by an ion source, forming ions, which are separated according to their mass-to-charge ratio (m/z) by a mass analyzer. Then, a detector registers the number of ions at each m/z value, resulting in a spectrum of m/z and intensity pairs [23].

There are two primary techniques for using MS to characterize proteins. Top-down proteomics examines intact proteins without requiring proteolytic digestion, whereas bottom-up proteomics requires cleavage of proteins into peptides prior to MS analysis [23]. Over the last decades, liquid chromatography separation of peptides coupled online with tandem mass spectrometry (LC-MS/MS) has become the preferred technology for high-throughput characterization of proteins and proteomes (**Figure 1.2**) [22].

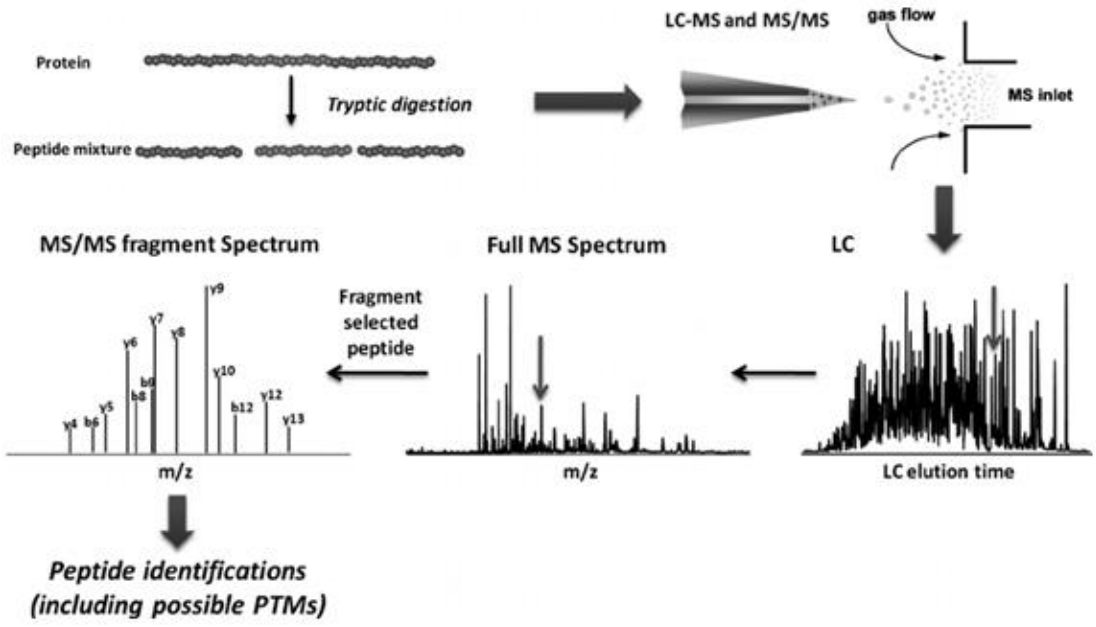


Figure 1.2 – General workflow for LC-MS -based global proteomics: proteins in complex biological samples are first converted to peptides by proteolytic digestion (e.g., tryptic digestion). The resulting peptide mixture is then separated by LC and ionized by electrospray before entering the mass spectrometer. In a typical data-dependent acquisition mode, a full MS spectrum is acquired for the peptides eluting from the LC column at a given time. One of the most intense ion species (i.e., peptides) is then isolated and fragmented to obtain the MS pattern of its fragments (i.e., MS /MS spectrum). Since peptide bonds tend to fragment under conditions of collision-induced dissociation in MS /MS analysis and generate predominantly b- or y-type ions (N- or C-terminal fragments carrying charge, respectively), the peptide sequence can be deduced from the MS /MS spectrum. This process is fully automated by matching the MS /MS spectra with protein sequence databases. Possible post-translational modifications can also be identified by including dynamic modifications at specific amino acid residues (e.g. Ser, Thr or Tyr for phosphorylation) in the database search. Quantification of individual peptides is usually performed using extracted ion chromatogram level. PTMs , post-translational modifications. Adapted from [24].

In bottom-up proteomics, LC-MS /MS analysis begins with the proteolytic cleavage of proteins into peptides, which are then separated by liquid chromatography prior to MS analysis (**Figure 1.2**). For peptide and protein identification (i.e., Discovery/Shotgun proteomics), MS acquisition is performed in information-dependent acquisition mode (IDA, or data-dependent acquisition, DDA), where a subset of precursor ions generated in the first stage of MS (MS) is selected and fragmented, resulting in fragment ions (MS/MS) that are generated from a single precursor ion. Subsequently, the MS/MS spectra are analyzed using a sequence database search. This MS workflow is performed in high-resolution MS systems as quadrupole time-of-flight (QTOF) or orbitraps and is the gold standard for protein identification. Although some MS-based quantification methods have been applied to IDA workflows (such as MS or MS/MS-based quantification) they suffer from several limitations. These include a low specificity and reproducibility and also a low degree of repeatability, specifically for low-abundant proteins such as HCP. For accurate MS-based protein quantification, Multiple Reaction Monitoring (MRM) is used in target proteomics. In this method, the first mass analyzer is set to filter ions of a certain m/z. The isolated peptide is fragmented, and one of the resulting fragment ions is monitored with another mass analyzer set to filter a concrete m/z value. The double selection of a peptide precursor ion and peptide fragment ion is called a transition [23,25,26]. This MRM-targeted MS

method is commonly performed in low-resolution instruments such as triple-quadrupoles or QTRAP. To overcome the limitations of IDA and MRM, a powerful data-independent acquisition technique (DIA) has recently emerged: Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra (SWATH-MS) [25]. SWATH must be done in high resolution and high-speed MS as the case of the triple time-of-flight (TripleTOF) systems [26].

1.2.1.2.1. TripleTOF 6600 mass spectrometer

Mass spectrometry measures the mass-to-charge ratio of ions to identify unknown compounds, quantify known compounds, and provide information about the structural and chemical properties of molecules [27].

The TripleTOF® 6600+ system features a series of quadrupole filters that pass ions according to their mass-to-charge (m/z) value. The first quadrupole in this series is the QJet ion conductor, located between the orifice and the Q0 region. The QJet ion guide does not filter the ions but focuses them before they enter the Q0 region. By pre-focusing the larger ion flux created by the wider aperture, the QJet ion guide increases the sensitivity of the instrument and improves the signal-to-noise ratio. In the Q0 region, the ions are refocused before entering the Q1 quadrupole [27].

The Q1 quadrupole sorts the ions before they enter the Q2 collision cell. The Q1 quadrupole operates in two modes:

- forwards all ions within a certain m/z range to the Q2 collision cell. This is a time-of-flight (TOF) MS scan. All ions are analyzed by the TOF system.
- passes an ion with a specific m/z ratio to the Q2 collision cell. This is a TOF MS /MS scan. Only the selected ion is analyzed.

In the Q2 collision cell, the internal energy of the ions is increased by collisions with gas molecules to the point where molecular bonds break and fragmentation occurs. With this technique, it is possible to design experiments to obtain information about the structural and chemical properties of the molecules [27].

After the ions pass through the Q2 collision cell, they enter the TOF region for further mass analysis. Contrary to quadrupoles, the TOF is a high resolution (40.000 full width at half maximum) mass analyzer and, therefore, enables high-resolution MS experiments. Ions reach the detector at different times depending on their m/z ratio. In the detector, the ions generate a current which is converted into a voltage pulse. These voltage pulses are counted and the number of pulses is directly proportional to the amount of ions entering the detector. The mass spectrometer converts the voltage pulses into a signal and then correlates the signal with the time each ion takes to reach the detector. The signal represents the ion intensity and the time to reach the detector for a given m/z value. The mass spectrometer displays this data as a mass spectrum [27].

1.2.1.2.2. QTRAP mass spectrometer

A QTRAP instrument is a triple quadrupole mass spectrometer with a linear ion trap in Q3 (**Figure 1.3**). A QTRAP system is a very versatile instrument because it can be used in a variety of ways. If needed, it can be used as a simple triple quadrupole instrument or in QTRAP mode only, or a combination of both technologies. In QTRAP mode, a MRM3 and MS/MS confirmation information can be obtained. QTRAP systems are advantageous as they offer higher sensitivity and better resolution compared to triple-quadrupole instruments [28].

A series of quadrupole filters in the mass spectrometer transfers ions based on their mass-to-charge ratio (m/z). The IonDrive™ QJet ion guide, located between the aperture and the Q0 region, is the first quadrupole in this series. The ions are focused before entering the Q0 zone by the IonDrive™ Qjet ion guide, which does not filter them. The IonDrive™ Qjet ion guide improves the sensitivity of the system and the signal-to-noise ratio by pre-focusing the higher ion flux that results from the larger aperture. The ions are refocused in the Q0 region before entering the Q1 quadrupole [28].

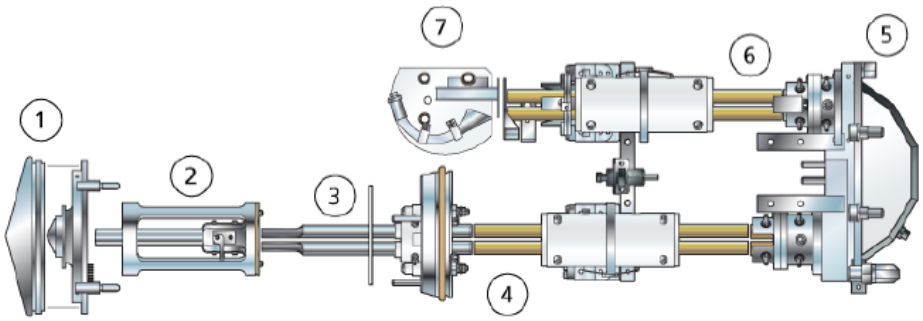


Figure 1.3 – A diagram of the internal structure of the 6500+ QTRAP systems: (1) aperture; (2) Qjet ion guide; (3) Q0 region; (4) Q1 quadrupole; (5) Q2 collision cell; (6) Q3 quadrupole; and (7) detector. Adapted from [28].

The Q1 quadrupole acts as a filter and sorts the ions before they enter the Q2 collision cell (**Figure 1.4**). In the Q2 collision cell, the internal energy of an ion is increased by colliding with gas molecules to the point that molecular bonds break making product ions [28].

After the ions are fragmented in the Q2 collision cell, fragment ions enter the Q3 quadrupole, where they are further filtered. The ions generate a current in the detector, which is then converted into a voltage pulse. The amount of ions entering the detector is directly proportional to the voltage pulses leaving the detector. These voltage pulses are monitored by the system, which then converts the data into a signal. The signal represents the ionic strength for a given m/z value, which is displayed by the system as a mass spectrum [28].

As described above, QTRAP systems are mainly used to performed MRM-targeted quantification, using the signal of selected MS/MS fragment ions for quantification. Briefly, in this methodology, Q1 filters a specific peptide ion, Q2 fragments the peptide, and Q3 selects a specific fragment ion for the detector, as shown below in **Figure 1.4** [29].

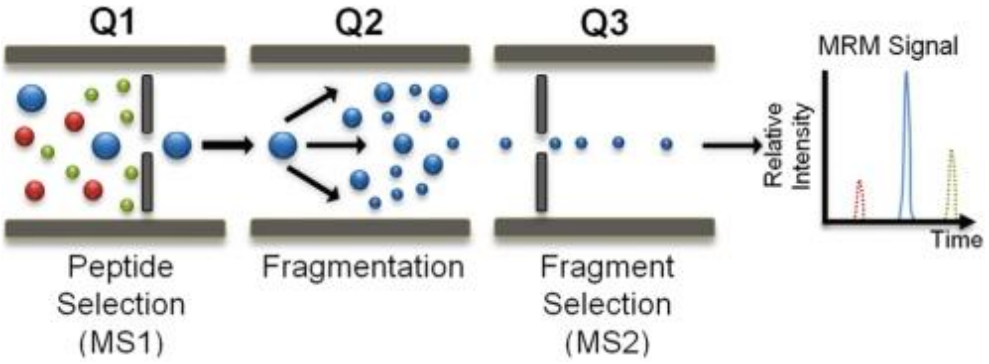


Figure 1.4 – Mass spectrometry with multiple reaction monitoring (MRM-MS). Schematic diagram of a triple quadrupole mass spectrometer (QQQ-MS) commonly used for MRM-MS: Q1 and Q3 represent two mass filters for precursor and fragment ion selection, respectively, while Q2 (collision cell) generates fragment ions by collision-induced dissociation (CID). In this case, one of the three peptide precursor ions (colored blue) is selected in Q1, fragmented in Q2, and quantified using one of its fragment ions (transition), which is selected in Q3 based on the relative intensity of its peak area. Q1 filters outselects the m/z of a given HCP peptide of interest, while Q3 selects the fragment ion of interest used for peptide quantification. Adapted from [29].

1.2.1.2.3. MS-based protein quantification

Proteome quantification using liquid chromatography-mass spectrometry (LC-MS) is a common method for studying biological processes. In global proteome studies, quantitative mass spectrometry approaches can be applied to perform an absolute and relative quantification, each with its own set of advantages in terms of quantification accuracy, proteome coverage, complexity, and robustness [30].

Relative quantification, which compares full proteomes or the relative amount of a large number of proteins, and absolute quantification, which determines the absolute concentration of specific proteins within a sample, are two common methodologies [31].

MS-based methods for absolute quantification uses peptides labeled with stable, non-radioactive isotopes. Isotope–labeled, or 'heavy' peptides, are derived from natural amino acids by substitution of certain atoms (N, C, H) with their 'heavy isotope' variant. The most frequent substitutions are ^{12}C by ^{13}C (carbon-13), ^{14}N by ^{15}N (nitrogen-15), and ^{1}H by ^{2}H (deuterium) [32]. Most commonly, peptides are labelled at arginine or lysine residues, leading to a mass increase of +10 Da and +8 Da, respectively. Stable-isotope-labeled (SIL) peptides display identical physiochemical properties and chemical reactivity as their non-labeled counterparts with a defined mass difference. This constitutes the basis for using stable isotope labeling peptides in a variety of absolute quantification applications such as quantitative proteomics, the quantification of complex protein mixtures at very low concentration, or NMR studies [32].These SIL peptides are used to build a calibration curve in the sample matrix. Additionally, heavy labeled peptides are used as internal standards since these peptides can be spiked into the sample. This methodology is often performed in MRM-targeted proteomics workflow using triple quadrupoles or QTRAP MS systems. The peptides used for absolute quantification are called proteotypic peptides, which are defined as the peptides that uniquely identify each protein and are consistently observed when a sample mixture is interrogated by a (tandem) mass spectrometer [33].

Global proteome quantification may involve chemical or metabolic stable isotope labeling (label-based) or quantification without the use of stable isotopes (label-free) [30]. There are several MS labeled-based methods that have been successfully used for large-scale quantitative proteomic studies resulting in high quantification accuracy, precision and also allowing sample multiplexing. These include chemical tagging methods such as isobaric tags for relative and absolute quantification (iTRAQ), isobaric labelling with tandem mass tags (TMT) and isotope-coded affinity tags (ICAT), and metabolic labelling, namely stable isotope labelling with amino acids (SILAC) [22,34]. In essence, a chemical reaction *in vitro* between the reagent and the peptides of interest results in a heavy-labeled peptide mixture. For relative quantification after LC-MS/MS, a mass shift between two distinct isotopic-labeled samples is used. As a result, comparing the MS intensity of a sample of interest to that of a labeled peptide standard present in the same sample can be used to perform relative and absolute quantification [30].

In the field of quantitative proteomics, label-free relative quantification is a high-throughput approach. Sample preparation is accomplished without the need to label reagents, making such methods easy and cost-effective. In this work, we used SWATH-MS as a data-independent MS acquisition method [32].

1.2.1.2.4. SWATH-MS

SWATH-MS, on the other hand, is a revolutionary method for MS -based quantification of peptide ions in a sample that does not require labeling (label-free quantification) [35]. In a SWATH-MS measurement, all ionized peptides of a given sample that fall within a certain mass range are fragmented systematically and unbiasedly, using relatively large windows for precursor isolation (**Figure 1.5**) [22]. SWATH-MS combines the advantages of the high throughput of shotgun proteomics with the uniformity and repeatability of SRM. The high-resolution MS data are generated by analyzing successive windows (called swaths) across the chromatographic elution range with a mass analyzer combining a quadrupole and TOF instrument (QTOF) [26]. The operations of tagging methods typically result in protein loss, which is prevented in SWATH-MS [36]. Being label-free eliminates quantification errors due to insufficient labelling [37].

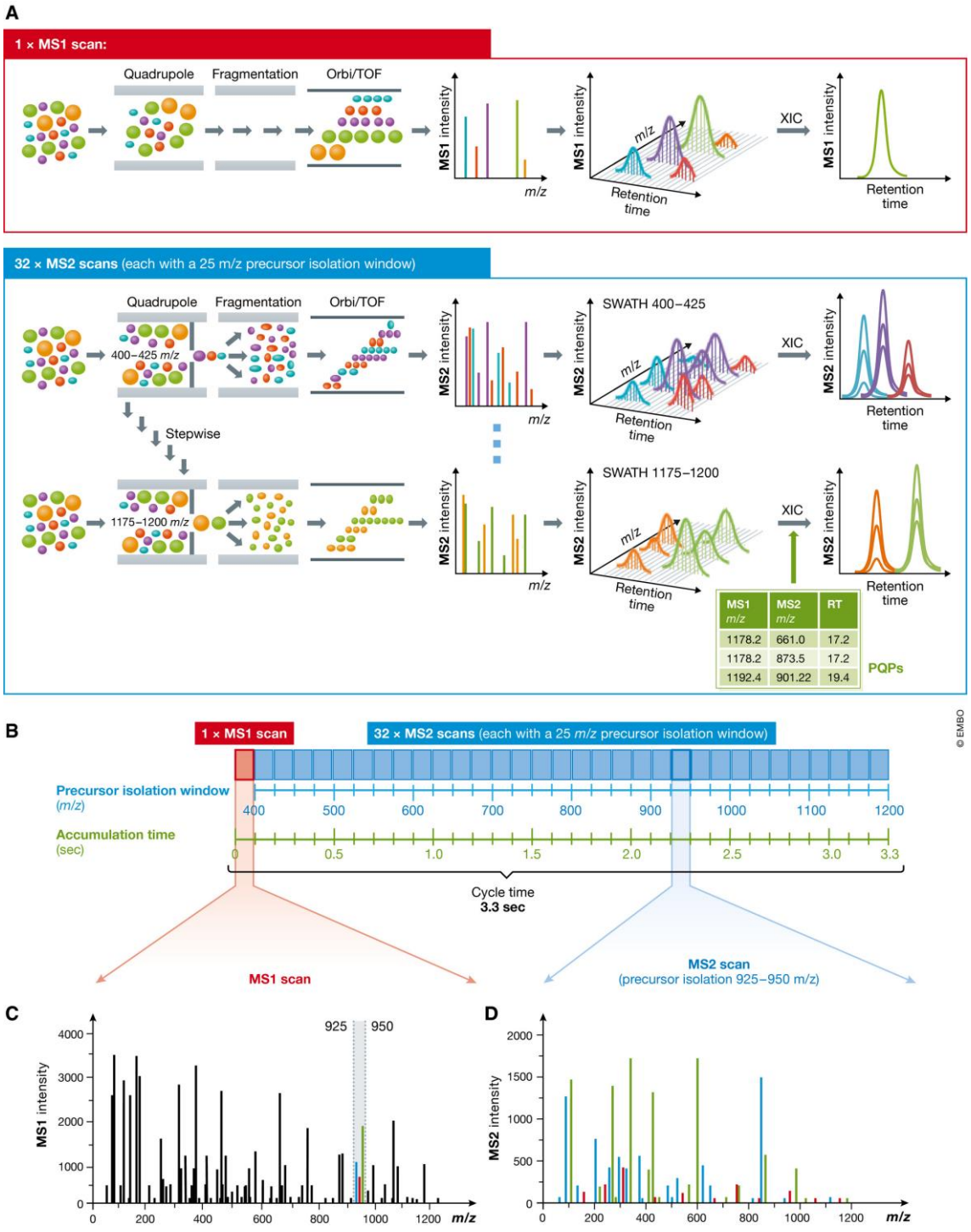


Figure 1.5 – Principle of sequentially windowed, data-independent acquisition in SWATH-MS: (A) SWATH-MS measurements are performed with fast-scanning hybrid mass spectrometers, typically using a quadrupole as the first mass analyzer and a TOF or Orbitrap as the second mass analyzer. In SWATH-MS mode, a single precursor ion spectrum (MS) is typically recorded, followed by a series of fragment ion spectra (MS/MS) with wide precursor isolation windows (e.g., 25 m/z). By repeatedly cycling through successive precursor isolation windows over a defined mass range, a comprehensive data set is recorded that contains continuous information on all detectable fragments and precursor ions. In this way, extracted ion chromatograms can be generated at both MS/MS and MS levels. A peptide-centric scoring strategy can be used to analyse the SWATH-MS data, which requires prior knowledge of the chromatographic and mass spectrometric behaviour of all interrogated peptides in terms of peptide interrogation parameters . (B) The data acquisition scheme SWATH-MS for a Q-TOF mass spectrometer uses 32 MS/MS scans with defined increments of 25 m/z, starting at 400 m/z and ending at 1,200 m/z. A full MS scan is recorded at the beginning. With a recording time of 100 ms per scan, the total cycle time is ≈3.3 s. (C) The full MS scan detects all peptide precursors eluting at a given time point. For example, in the mass range of 925 to 950

m/z, three co-eluting peptide species are detected (green, red, and blue). (D) The corresponding MS/MS scan with a precursor isolation window of 925-950 m/z represents a mixed MS/MS spectrum with fragments of all three peptide species. Adapted from: [22].

The major advantage of SWATH-MS is that it allows quantitative studies of peptides covering thousands of proteins with a high degree of quantitative consistency and precision (**Table 1.1**). It is suitable for studies with large numbers of samples that require precise and repeatable quantification of the expressed proteome or peptidome in each sample. To perform SWATH-MS analysis, unlabeled protein samples are digested with trypsin and the peptides are then analyzed using liquid chromatography and a tandem mass spectrometer in DIA mode. In this mode, all ionized chemicals in a sample that fall within a certain mass range are systematically and unbiasedly fragmented [22].

Table 1.1 – Advantages and limitations of SWATH-MS compared to data-dependent (DDA) and targeted (SRM, PRM) proteomics. Adapted from: [22].

	Data independent acquisition-based SWATH-MS	**Data-dependent acquisition (DDA)**	**Targeted acquisition (SRM and PRM)**
Ease of data acquisition	** Easy, requires definition of mass range to cover, precursor isolation window width and number of MS2 scans per cycle	*** Easiest, default setup on most mass spectrometers, requires definition of TopN method, MS2 trigger threshold and dynamic exclusion time	* Hardest, requires generation and optimization of targeted peptide assays for data acquisition
Ease of data analysis	* Currently hardest, requires peptide query parameters, sophisticated software tools, large informatics resources	*** Currently easiest, multitude of pipelines available	** Currently easy, several software tools developed for manual and automated analysis
Breadth of protein and peptide detection/multiplexing	*** 10,000s of peptides per MS injection quantifiable	*** 10,000s of peptides per MS injection quantifiable	* 10s - 100s peptides per MS injection quantifiable
Selectivity/sensitivity/dynamic quantification range	** 4 orders of magnitude per MS injection	** 4 orders of magnitude per MS injection	*** 4–5 orders of magnitude per MS injection
Reproducibility/data consistency	*** High, due to peptide-centric scoring analysis	* Low, due to stochastic sampling in DDA	*** High, due to targeted data acquisition
Retrospective targeting (using chromatogram extraction)	*** Possible on MS1 and MS2 level	** Possible on MS1 level only	* Not possible due to targeted data acquisition

*Least optimal performance.
**Medium performance.
***Best performance.

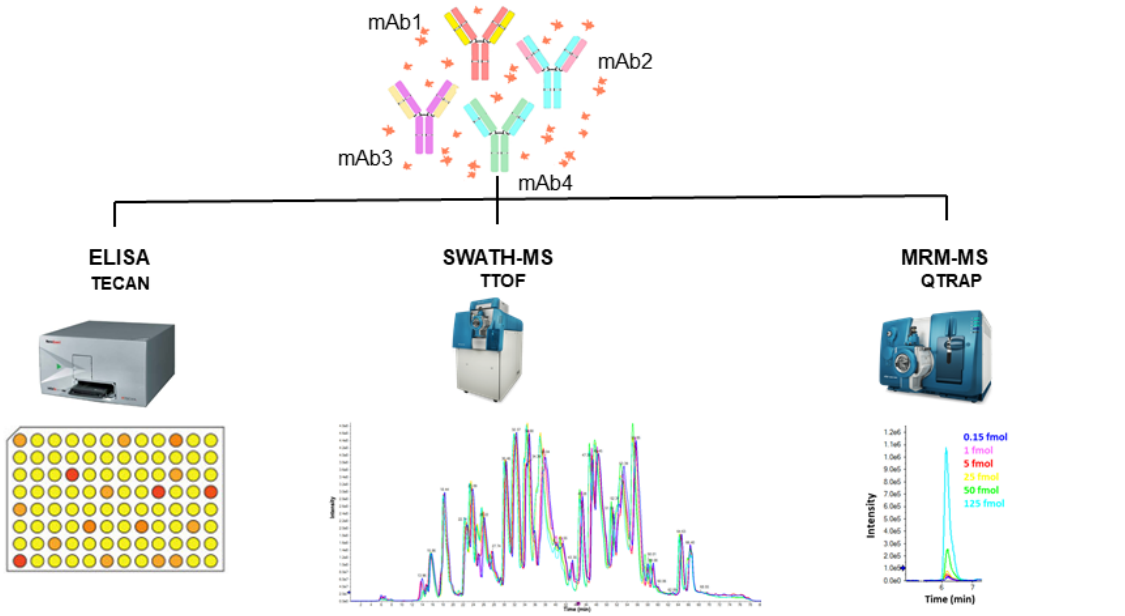
The work developed during this thesis aimed at combining the great potential of the SWATH approach (described above) with the specificity and accuracy of using labeled peptides for quantification, enabling not only the identification but also relative and semi-absolute quantification (using a generic average curve of different heavy peptides) of hundreds of proteins in different samples, and particularly, the absolute quantification of targeted proteins (using their specific heavy peptides).

2. Aim of the Thesis

This work was focused on the development of an accurate quantitative methodology for the in-process characterization of mAb samples using MS-based technologies. This analytical approach was used to detect and quantify nine target HCPs across different downstream process (DSP) steps (polishing platform). In addition, untargeted analysis with identification and semi-absolute quantification of HCPs was also established, further enabling a deeper characterization of HCP profile in the polishing platform studied.

The work developed in this dissertation is divided into three main parts:

- Quantification of total HCP by ELISA;
- Development of a SWATH-DIA method enabling the absolute quantification of 9 target HCPs and identification with relative/semi-absolute quantification of over 200 HCPs;
- Implementation of a method for the absolute quantification of a selected HCP (HCP_A) by MRM-MS (confirmation of the SWATH results and demonstration of method development).



3. Materials and Methods

3.1. Samples

The purified monoclonal antibody samples were generated and provided by one of iBET's research groups. Four different molecules (mAb1, mAb2, mAb3 and mAb4), purified by protein A affinity chromatography, were processed in the screening polishing platform using polishing resin 1 followed by polishing resin 2A or 2B (alternative). Ten different DSP conditions of pH and salt concentration (NaCl) were tested per resin (hereinafter referred as "Loads"), and in each of these chromatographic runs several fractions were obtained and combined into a pool (hereinafter referred to as "Pools"). A total of 31 samples per mAb were analysed by SWATH-MS and QTRAP-MS (protein A eluate - PAE and 10 pools per resin) and a total of 61 samples were analysed by ELISA (PAE, 10 loads and 10 pool samples per resin).

3.2. Chemicals

LC-MS grade acetonitrile 0.1% formic acid (solvent B), LC-MS grade water 0.1% formic acid (solvent A) and trifluoracetic acid (TFA) were purchased from Fisher Chemicals. Sequencing Grade Modified Trypsin was purchased from Promega; RapiGest Surfactant (SF) was purchased from Waters; dithiothreitol (DTT), iodoacetamide (IAA), and ammonium bicarbonate (Ambic) were purchased from Sigma. Tuning solutions were purchased from SCIEX.

3.3. Sample Preparation for LC-MS methods

Each sample to be tested on LC-MS was prepared in duplicate (A and B). The protocol is divided into four phases: (1) denaturation and reduction; (2) alkylation; (3) digestion; and (4) degradation of Rapigest SF.

In phase (1), sample concentration was adjusted to 0.8 µg/µL using 50 mM Ambic. In a new microtube, 125 µL of sample (100 µg protein), 6.5 µL Ambic (1 M), 7 µL Rapigest SF solution (2%), 1.5 µL DTT solution (500 mM) were added and incubated for 60 min at 60 ºC and 500 rpm.

In phase (2), the reaction mixture was cooled to room temperature and 4.5 µL IAA solution (500 mM) were added and incubated for 30 minutes at room temperature, 500 rpm.

In phase (3), 25 µL of trypsin solution (20 µg/ µL) were added to the reaction and incubated for 16-18 hours at 37 ºC, 500 rpm (protected from light).

Finally, in phase (4), 1.7 µL of TFA were added to the digested sample and incubated for 45 minutes at 37 ºC, 500 rpm. Then samples were centrifuged at 30000 rpm for 10 minutes to remove the precipitate. 150 µL of the clear supernatant were removed without disturbing the precipitate and transferred to a LC vial. Samples were stored at -80 ºC until further analysis.

3.4. Analytical Methods

3.4.1. HCPs Quantification by ELISA

Quantification of residual HCP from CHO cells was performed using a commercial sandwich ELISA kit (Cygnus technologies, F550-1) according to the manufacturer's recommended protocol.

In summary, we added to each anti-CHO coated well 100 µL of anti-CHO:horseradish peroxidase (HRP) conjugate plus 50 µL of: standard (CHO HCPs in bovine serum albumin with preservative, at 0, 1, 3, 6, 12, 20, 40 and 100 ng/mL), positive control (CHO HCP Antigen Concentrate F553H-1, Cygnus, final concentration 20 ng/ml) or samples with unknown HCP concentration. Plates were incubated at 500 rpm for 2 hours. Each well was washed four times with 350 µL of 1x wash buffer. After the wash step, 100 µL of tetramethylbenzidine substrate were added and plates were incubated for 30 min without agitation and protected from light. Finally, 100 µL of stop solution were added to each well. The absorbance was then measured at 450 nm using a fluorescent plate reader and read in duplicate (TECAN Infinite 200 Pro NanoQuant).

3.4.2. HCPs Quantification by Mass Spectrometry Analysis

Mass spectrometry analysis was performed with the assistance of expert technicians at UniMS, the Mass Spectrometry Facility at iBET/ITQB-NOVA.

3.4.2.1. HCP analysis using the Sciex TripleTOF 6600

Monoclonal antibody samples were analyzed by NanoLC-MS/MS using an ekspert™ NanoLC 425 cHiPLC system coupled to a TripleTOF® 6600 with a NanoSpray® III source (SCIEX).

3.4.2.1.1. Spectral library generation

To build a spectral library for the SWATH analysis, each sample (1.6 µg protein) was used for information-dependent acquisition (IDA) analysis by reverse-phase chromatography (RP-LC) with a trap-and-elute configuration. Trapping was performed at 2 µL/min with solvent A (Water 0.1% Formic Acid), for 10 min, on a Nano cHiPLC Trap column (SCIEX 200 µm x 0.5 mm, ChromXP C18-CL, 3 µm, 120 Å). Peptide separation was performed on a Nano cHiPLC column (SCIEX 75 µm x 15 cm ChromXP C18-CL 3 µm 120 Å) at a flow rate of 300 µL/min applying a 60 min linear gradient of 5% to 30% (v/v) of solvent B (Acetonitrile 0.1% Formic acid).

Peptides were sprayed into the MS through an uncoated fused-silica PicoTip™ emitter (360 µm O.D., 20 µm I.D., 10 ± 1.0 µm tip I.D., New Objective). The source parameters were set as follows: 15 GS1, 0 GS2, 30 CUR, 2.5 keV ISVF and 100ºC IHT. For IDA acquisition, the mass spectrometer was set for IDA scanning full spectra (400-2000 m/z) for 250 ms and the top 50 ions were selected for subsequent MS/MS scans (150-1800 m/z for 40 ms each) with a total cycle time of 2.3 s. The selection criteria for parent ions included a charge state between +2 and +5 and counts above a minimum threshold of 125 counts per second. Ions were excluded from further MS/MS analysis for 12 s. Fragmentation was performed using rolling collision energy with a collision energy spread of 5. For protein identification and spectral library generation, all IDA raw files were combined using ProteinPilot™ (version 5.0 (SCIEX)) search engine with the Paragon algorithm. The following search parameters were used: search against the CHO database (UP000001075, 23885 proteins from Swiss-Prot & TrEMBL); trypsin digestion; iodoacetamide cysteine alkylation; through identification efforts. A false discovery rate analysis was also performed. The output of these searches, in the form of a group file, was used as the reference

spectral library. Only the proteins within a false discovery rate (FDR) <1% were considered (236 proteins)

3.4.2.1.2. Selection of proteotypic peptides

For this work, 9 different HCPs were targeted (HCP_A; HCP_B; HCP_C; HCP_D; HCP_E; HCP_F; HCP_G; HCP_H; HCP_I). For each HCP, proteotypic peptides were initially chosen using the Skyline software based on the following criteria: 0 missed cleavages; 8-45 amino acids length; excluding peptides containing Cys, Met, His and RP/KP and NXT/NXS motifs. Additionally, a peptide library (built based on IDA runs of PAE samples and CHO supernatants) was used to confirm the detection/identification of the proteotypic peptides. The final list of peptides was ranked by intensity levels observed on PAE samples, the TOP 2 peptides of each protein were selected for synthesis of heavy labelled peptides (denominated for each HCP peptide 1 and peptide 2). For HCP_I just one peptide was selected according to the criteria defined. All 17 heavy peptides (labeling on R or K, depending on the peptide - ^{13}C, ^{15}N) were purchased from PEPSCAN.

3.4.2.1.3. SWATH-MS

For identification and quantification of HCP by MS, 1.6 µg of protein from each sample was analysed by SWATH-MS acquisition in triplicate runs. Each sample was spiked with 5 fmol of each of the 17 heavy peptide standards. The amount of heavy peptide standards was determined from the MS signal observed in the IDA runs. Chromatographic conditions were similar to the previously described IDA run. The mass spectrometer was operated with a cyclic product ion data-independent acquisition (DIA). A variable window calculator (SWATH Variable Window Calculator_V1.0, SCIEX) and an acquisition method editor SWATH (SCIEX) were used to set up the SWATH acquisition. A set of 64 overlapping windows (1 m/z for window overlap) was created, covering the precursor mass range of 350 - 1800 m/z. A survey scan of 40 ms was performed at the beginning of each cycle and the SWATH-MS/MS spectra were collected for 50 ms (150-1800 m/z), resulting in a cycle time of 3.3 s. A rolling collision energy with a collision energy spread of 15 was used. Spectral alignment and targeted data extraction of the DIA samples were performed using PeakView v.2.1 (SCIEX) with the reference spectra library. The following parameters were used for data extraction: six peptides/protein, six transitions/peptide, peptide confidence level of > 99%, FDR threshold of 1%, common peptide exclusion, and extracted ion chromatogram (XIC) window of 10 min and width of 20 ppm.

For the calibration curve, the mAb1 PAE sample was used as a matrix. This sample was prepared as described in section 3.4 and spiked with the 17 heavy peptides chosen from the 9 target HCPs. Seven calibration points were used for each curve: 0.25, 0.5, 1, 2, 5, 10, and 50 fmol of each heavy peptide. Quality criteria applied to evaluate each peptide/transition calibration curve were: coefficient of determination (R^2) for the linear regression higher than 0.99 and the coefficient of variation (CV) between replicates of each concentration lower than 25% (according to the preliminary data obtained during method development and previous knowledge on the nanoLC-MS workflow general performance).

3.4.2.2. QTRAP – Targeted MRM-MS

This analytical procedure was used for the targeted detection and quantification of HCP_A in different samples from different mAbs. Samples were analysed using a system ExionLC™ AC (SCIEX) coupled to a triple quadrupole mass spectrometer QTRAP® 6500+ with a Turbo Spray IonDrive source (SCIEX). For the absolute quantification, the same two HCP_A heavy peptides selected for SWATH analysis were used. Again, for each peptide, three transitions were followed, however only the most intense transition was used for quantification: transition y5 for peptide 1 and transition y8 for peptide 2.

For each sample, 16 µg of total protein were injected. Each sample was spiked with 10 fmol of the two HCP_A heavy peptides. Peptide separation was performed on an Acquity UPLC peptide column (Waters CSH C18 column, 1.7 µm, 1 mm x 150 mm, 130 Å) at 0.1 mL/min using a 37 min linear gradient of 12% solvent A (Water 0.1% formic acid) to 95% (v/v) of solvent B (Acetonitrile 0.1% formic acid). Peptides were sprayed into MS operated in low-mass mode. The following additional parameters were applied (**Table 3.1**): Multiple Reaction Monitoring (MRM) transitions, where a different collision energy (CE) and collision cell exit potential (CXP) were defined for each transition; Schedule MRM scanning with a detection window of 192 s; and a target scan time of 0.5 s. For the calibration curve, HCP_A heavy peptides were spiked in mAb1 PAE and four calibration points were used: 1, 5, 10 and 50 fmol. The quality criteria applied to evaluate each peptide/transition calibration curve were: coefficient of determination (R^2) for the linear regression has to be higher than 0.995 and the coefficient of variation (CV) between replicates of each concentration must be lower than 15% (according to the protocol used as reference for method implementation).

Table 3.1 - Multiple reaction monitoring (MRM) transitions for peptide 1 and peptide 2 (light and heavy) of HCP_A.

Q1 Mass [Da]	Q3 Mass [Da]	Time [min]	ID	CE [V]	CXP [V]
414.781	501.340	10.20	HCP A_ peptide 1 y4.light	22.5	33.0
417.781	614.424	10.20	HCP A _ peptide 1 y5.light	20.0	40.0
414.781	715.471	10.20	HCP A _ peptide 1 y6.light	21.0	48.0
418.788	509.354	10.20	HCP A _ peptide 1 y4.heavy	22.5	33.0
418.788	622.438	10.20	HCP A _ peptide 1 y5.heavy	20.0	40.0
418.788	723.485	10.20	HCP A _ peptide 1 y6.heavy	21.0	48.0
615.346	702.389	6.90	HCP A _ peptide 2 y7.light	39.0	45.0
615.346	817.416	6.90	HCP A _ peptide 2 y8.light	32.0	54.0
615.346	930.500	6.90	HCP A _ peptide 2 y9.light	32.0	54.0
620.350	712.398	6.90	HCP A _ peptide 2 y7.heavy	39.0	45.0
620.350	827.424	6.90	HCP A _ peptide 2 y8.heavy	32.0	54.0
620.350	940.509	6.90	HCP A _ peptide 2 y9.heavy	32.0	54.0

Note: CE (collision energy); CXP (collision cell exit potential)

3.5. Data analysis

3.5.1. ELISA

Samples' HCP concentration were determined using the calibration curve of the standards provided in the kit. All the calculations were done using a nonlinear fitting sigmoidal four-parameter logistic (4PL), as recommended by the manufacturer and using GraphPad Prism (version 8.2). HCP concentrations were expressed in ng/mL.

3.5.2. Mass spectrometry data

The SWATH-MS DIA data were analyzed for untargeted, relative quantification and absolute quantification. The relative quantification analysis of the SWATH data (analysis of non-targeted HCP) was performed using the MSMS All SWATH micro APP with PeakView (version 2.3, SCIEX) with the reference spectral library. For data extraction the following parameters were used: 6 peptides/protein, 6 transitions/peptide, peptide confidence level of >95%, FDR threshold of 1%, excluded shared peptides, and XIC extraction window of 6 min and width set at 20 ppm. For absolute quantification, SWATH-MS data were analyzed using Skyline software (version 21.1) to extract peak areas from the endogenous light peptide and heavy labeled standard. Data were manually curated and peak areas were determined automatically by the software.

The targeted MRM QTRAP-MS data was analyzed using the Analytics workspace of Sciex-OS (version 1.7, SCIEX). The processing method was built based on the MRM-QTRAQ acquisition method. The peak areas for the absolute quantification (SWATH-MS targeted data analysis and targeted MRM-QTRAP analysis) were exported as an Excel file for further calculations and analysis. The calculation details were not detailed in this thesis as the data analysis protocol was transferred from the partner where the targeted QTRAP quantification method was developed. The same equations were used for SWATH-MS absolute quantification. In summary, the peak areas were used to generate external calibration curves for the heavy-labeled peptides of each HCP. The quantifier peptide and transition were selected after the analysis of the calibration curves (i.e., three transitions per peptide and two peptides per HCP). Calibration curves $y = a_0 + a_1 x$ were obtained using a linear regression model and the coefficient of variation (CV) was calculated for each concentration level. The amount and concentration of endogenous HCP in the samples were determined using the external calibration curves normalized by the corresponding internal standard (IS) (heavy peptide spiked in each sample at a defined amount; 5 fmol for SWATH-MS or 10 fmol for targeted MRM-MS). The concentration of each HCP was expressed in ppm taking into consideration the molecular weight of the HCP target protein (conversion table: **Appendix 3.1)**. For the skyline analysis, data were manually revised, and the correct endogenous peptide was selected based on the heavy peptide signal.

4. Results and Discussion

The project where this thesis is inserted is aimed at understanding the clearance capacity of each resin of the chromatographic purification process implemented for mAb products produced in CHO cells. To address this, 10 different pH and conductivity (salt) conditions were investigated per chromatographic resin of the downstream process (DSP). The less purified sample (protein A eluate, PAE) was processed in two sequential polishing steps and with alternative resins (step 1 - DSP polishing resin 1, step 2 - DSP polishing resins 2A or 2B) (**Figure 4.1**). In particular, the work plan of this thesis was focused on the development of a quantitative LC-MS method that enables the characterization of the HCP profiles in different steps of the DSP. Loads (before a chromatographic step) and pools (after a chromatographic step) from 4 different mAb molecules (mAb1, mAb2, mAb3 and mAb4) were analysed by ELISA, while for SWATH-MS and MRM-QTRAP only pools (essential samples to evaluate clearance capacity of each resin) were used.

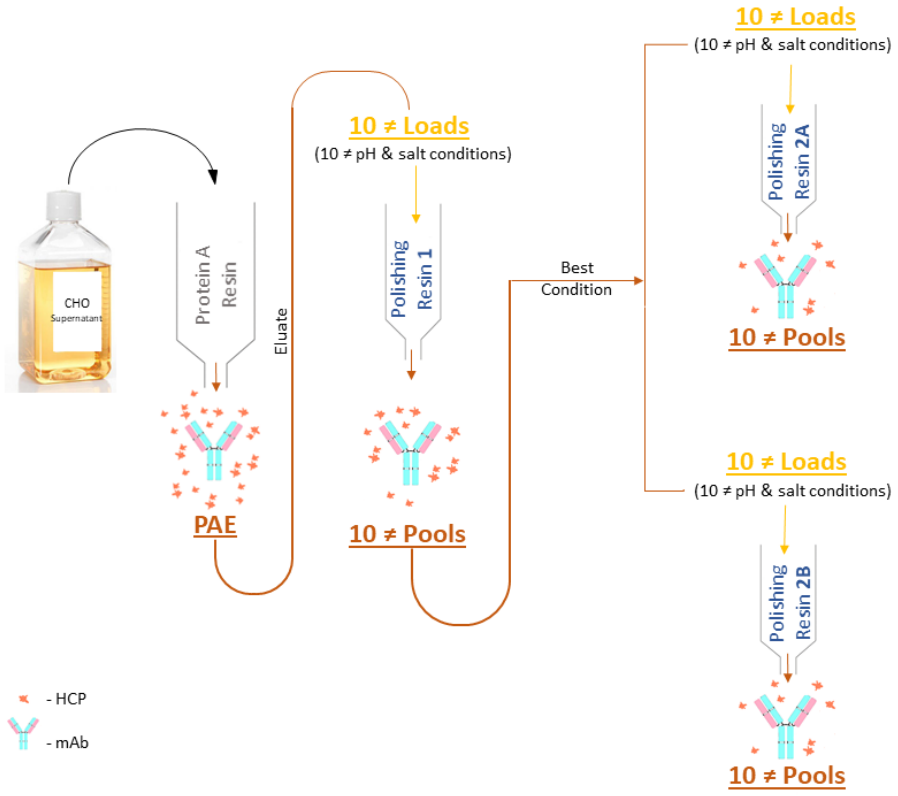


Figure 4.1 – Schematic representation of the different operations of the chromatographic purification process under study for each of the 4 mAbs studied, with sequential polishing steps and alternative resins where samples analysed in this thesis were generated (Protein A Eluate – PAE, Loads and Pools).

4.1. HCP determination by ELISA

Firstly, we used a commercial ELISA kit to confirm the presence and measure the total concentration of CHO HCPs in the mAbs samples. For that, a calibration curve (**Figure 4.2**) was constructed with 8 concentrations points (0; 1; 3; 6; 12; 20; 40 and 100 ng/mL).

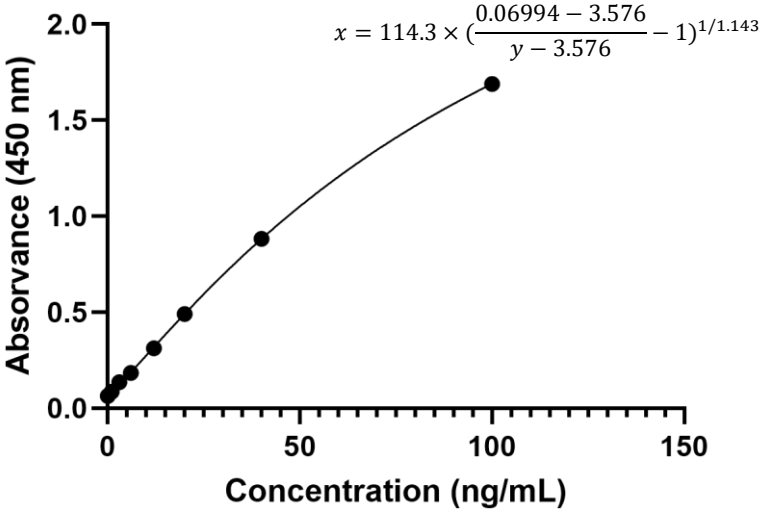


Figure 4.2 – Representative calibration curve using the CHO HCP ELISA kit. Measurements were performed in duplicate. A four-parameter logistic (4PL) nonlinear regression was applied to the data according to the manufacturer's instructions and using GraphPad software.

The established assay was used to quantify CHO HCP by interpolation in the calibration curve. Samples were analysed in duplicate (unspiked and spiked with the positive control for recovery calculation) and the dilutions used ranged from 1:5 to 1:500, depending on the samples (step of the polishing platform). To guarantee the adequate performance of the assay and the accuracy of the concentrations obtained (no matrix interference), for every sample the recovery of the spiked control was calculated (defined between 80%-120%, according to the kit acceptance criteria)

As shown in **Figures 4.3** to **4.6**, with this ELISA assay we confirmed the presence of CHO HCP in all mAb samples analysed and estimated the total HCP concentration in each of them. The calculated amount of CHO HCP in the 4 mAbs PAE samples was: mAb1 PAE ≈2000 ng/mL; mAb2 PAE ≈980 ng/mL; mAb3 PAE ≈23000 ng/mL and mAb4 PAE ≈1800 ng/mL. As expected, those concentrations corresponded to the highest amount of HCPs in all samples analysed for each mAb (less purified sample). With these data, we can conclude that mAb3 is the one that presents the highest concentration of total HCP, followed by mAb1 and mAb4 which are relatively similar and then mAb2, the one with the lowest amount of HCPs. While for mAb3 we were able to detect HCPs on all the different polishing steps/conditions, for mAb2 on the second polishing step there are almost no HCP quantified/detected (the limit of detection – LOD is ≈0.3 ng/mL, the limit of quantification – LOQ is ≈1 ng/mL according to kit specifications). Additionally, we can observe that even within the same polishing resin, some pH/salt conditions improved the clearance of HCP when compared to other conditions. Data show that for almost all conditions the amount of HCP is higher in the loads than in the pools (exceptions observed when working near the LOQ, where small discrepancies are within method variation), and is also higher in DSP polishing resin 1 (first step of the polishing platform) than in resins 2A or 2B (alternatives for the second step of the polishing platform).

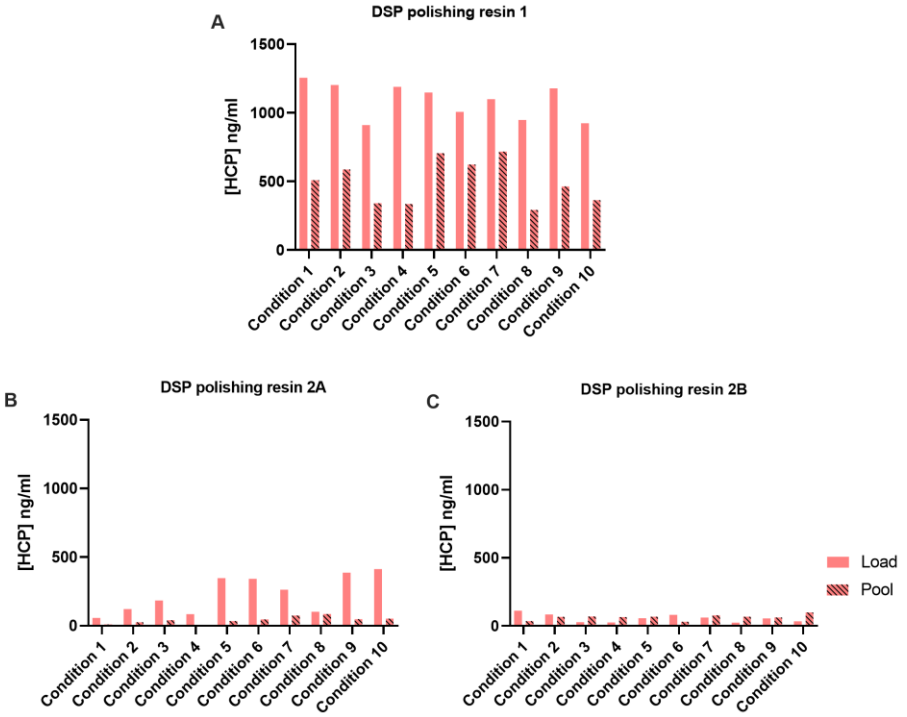


Figure 4.3 – Measurement of CHO HCPs by ELISA in mAb1 samples for the 10 conditions tested in each of the three DSP polishing resins under evaluation.

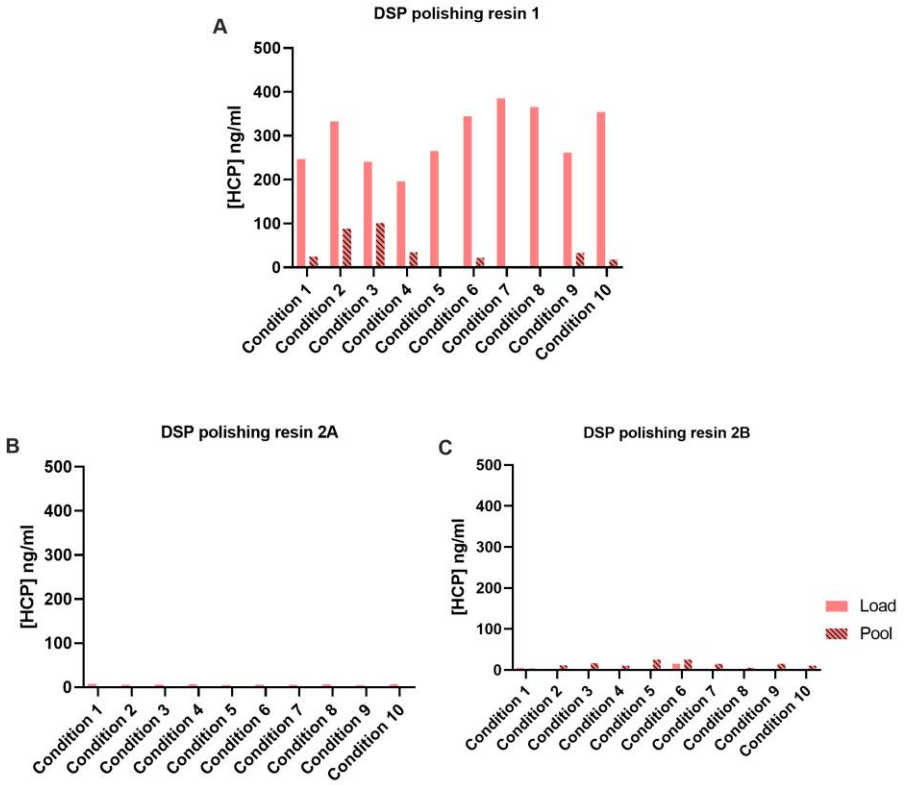


Figure 4.4 – Measurement of CHO HCPs by ELISA in mAb2 samples for the 10 conditions tested in each of the three DSP polishing resins under evaluation.

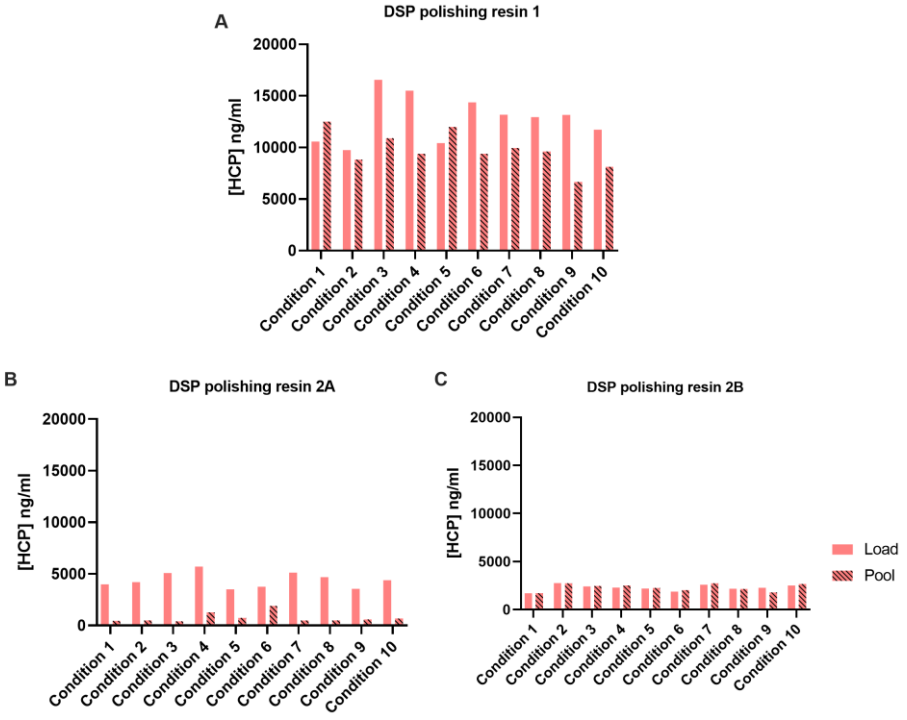


Figure 4.5 – Measurement of CHO HCPs by ELISA in mAb3 samples for the 10 conditions tested in each of the three DSP polishing resins under evaluation.

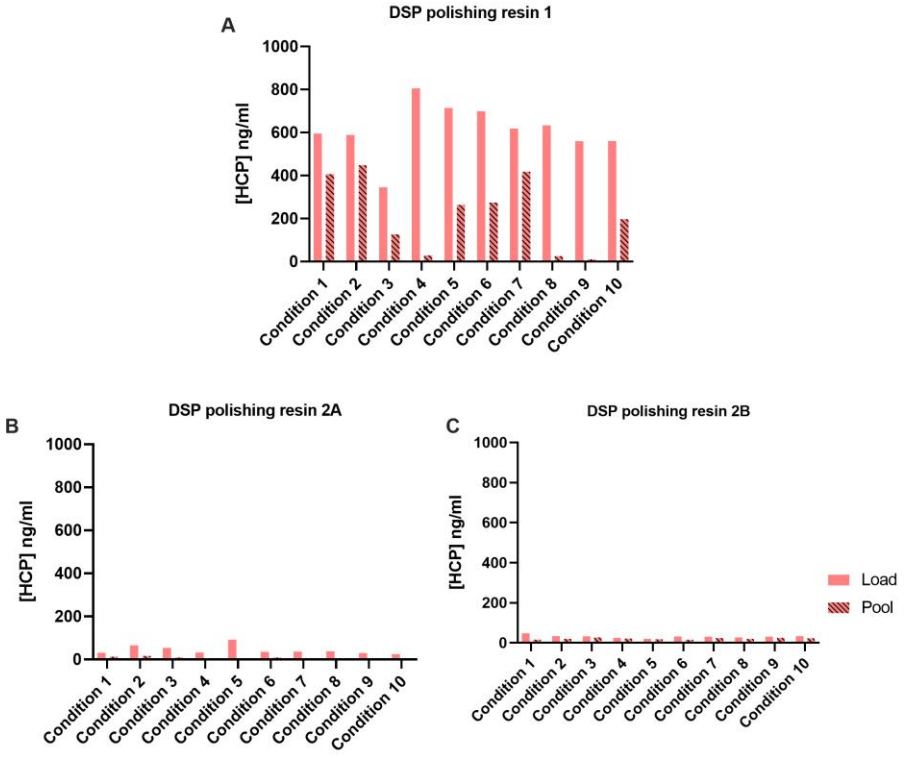


Figure 4.6 – Measurement of CHO HCPs by ELISA in mAb4 samples for the 10 conditions tested in each of the three DSP polishing resins under evaluation.

As expected, we observed a decrease in HCP concentration in the sequential chromatographic purification process under study, showing the resins capacity to clear HCP from the mAb samples (zoomed graphics are displayed in the Appendices section: **Appendix 7.2** to **7.5**). These data, although showing the expected trend of HCP clearance, do not identify which specific HCP are removed or which are still present in the purified samples. Additionally, we also cannot determine from the ELISA results if we still have dozens of different HCPs in the purified samples or just 1-2 residual proteins. This is key information, as certain HCPs, even if present in small amounts, can be detrimental to mAb quality, impacting final product efficacy and/or safety. To tackle these issues, we developed a quantitative MS-based method that allows us to determine which HCP and in which quantity is present in a particular sample.

4.2. Development of SWATH-MS quantification methods for HCPs

In this work, we developed a SWATH-MS workflow to perform, in a single MS assay; i) a targeted MRM-like absolute quantification of specific HCPs (using their specific standard curves), ii) untargeted semi-absolute quantification (for any HCP detected in the sample using a generic average standard curve with several peptides from different HCP) and/or iii) relative HCP quantification (between different samples, without a standard curve).

For the development of a targeted quantification method by SWATH-MS, 9 HCPs previously identified in biologics produced in CHO cells were selected according to the literature [19,20,38–41]. In this thesis, we generically refer to these 9 HCP as: HCP_A, HCP_B, HCP_C, HCP_D, HCP_E, HCP_F HCP_G, HCP_H and HCP_I. To proceed with absolute (and/or semi-absolute) quantification for each of these HCPs, 2 proteotypic peptides were selected according to the criteria described in section 3.4.2.1.2 (an exception is HCP_I for which only one peptide met the defined criteria). An ion library was built using IDA analysis not only to refine the selection of the proteotypic peptides (to confirm detection and intensity) for absolute quantification but also for the untargeted analysis of SWATH data. The MS data were analysed for peptide and protein identification using ProteinPilot. A total of 256 proteins were identified in this library. The mAbs light and heavy chains were identified as top-hits as expected, and some typical contaminants such as keratins and trypsin were also identified. The total number of HCPs (i.e., CHO proteins) identified were 236.

4.2.1. HCP absolute quantification using SWATH-MS

After the selection of the most suitable proteotypic peptides for detection/quantification of each targeted HCP, the SWATH runs were performed. Besides analysing samples spiked with the heavy labeled peptides as internal standards (IS), 7 different peptide amounts (0.25; 0.5; 1; 2; 5; 10 and 50 fmol) were spiked in mAb1 PAE and run individually to build the calibration curves for each peptide/HCP. In **Figure 4.7 A**, we present the total ion current (TIC) chromatogram of eight SWATH runs of mAb1 PAE (mab1 PAE was used as a matrix for all the calibration curves), each of them spiked with a different amount of the 17 heavy peptides. As can be observed, mAb1 PAE is a very complex sample, presenting high signals for mAb1 peptides (the therapeutic product is highly concentrated) and low-intensity signals for HCP peptides (some orders of

magnitude below the therapeutic product) (**Figure 4.7 B**). A XIC for the heavy labeled peptide 1 of HCP_A and the corresponding mass spectrum are shown in **Figure 4.7 C** and **D**, respectively. To account for technical variation (LC-MS system) triplicate injections of each calibration point were performed. As shown in **Figure 4.7 E** (TIC) and **F** (XIC of peptide 1 of HCP_A), highly reproducible technical replicates were obtained.

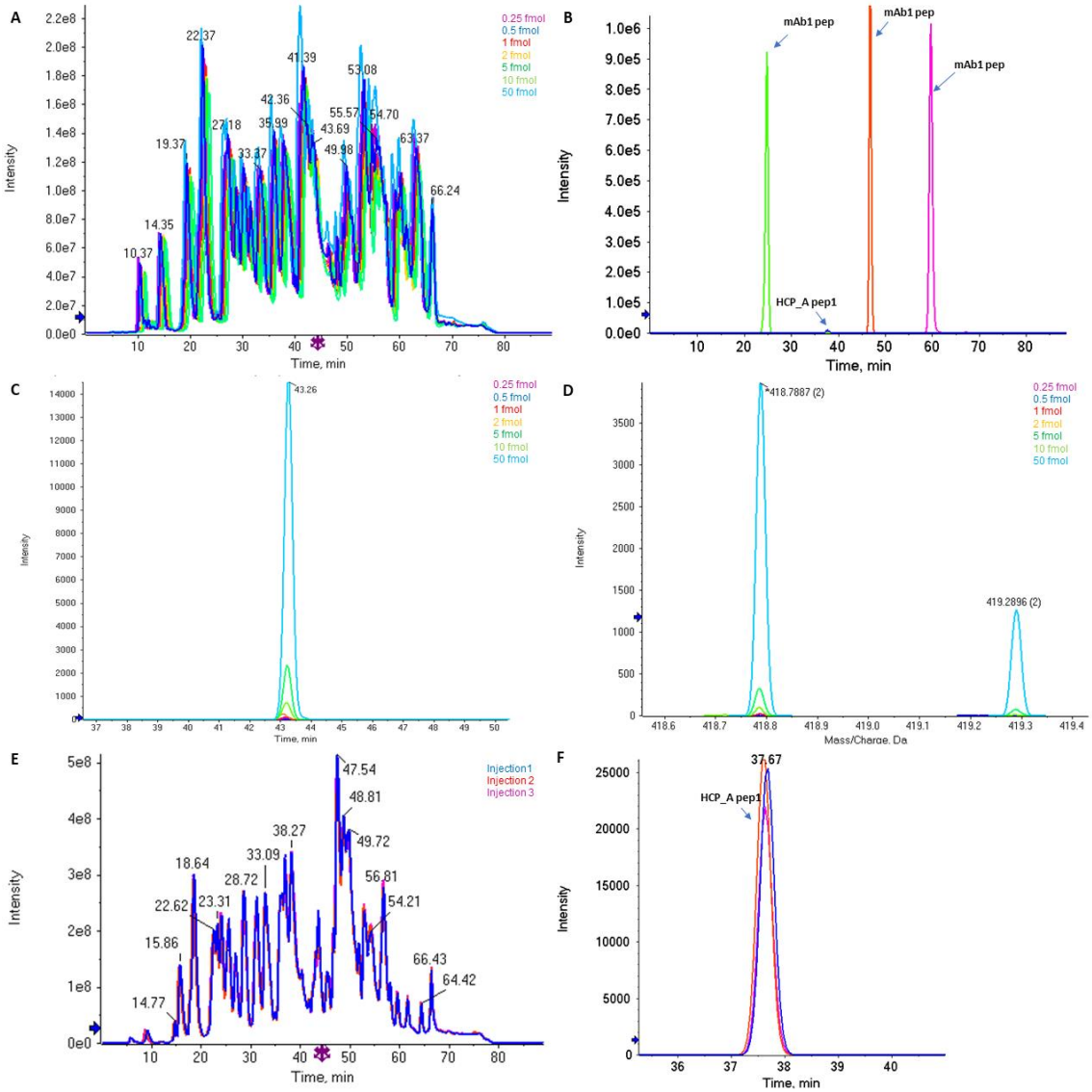


Figure 4.7 – (A) TIC from calibration curve runs (7 calibration points) using 17 heavy labeled proteotypic peptides of 9 HCPs; (B) XIC of mAb1 peptides and peptide 1 of HCP_A; (C) XIC of peptide 1 HCP_A and (D) Mass spectra of peptide 1 of HCP_A; (E) Triplicate injections of mAb1 PAE sample spiked with 5 fmol of peptide 1 of HCP_A; (F) XIC of peptide 1 of HCP_A.

To build the individual calibration curves (and later to obtain the quantification measurements of the endogenous HCP peptide – "light peptide" in the tested samples), the peak area for each transition (MS/MS) of the heavy labeled peptides at each concentration was determined using the Skyline software (as exemplified in **Figure 4.8**).

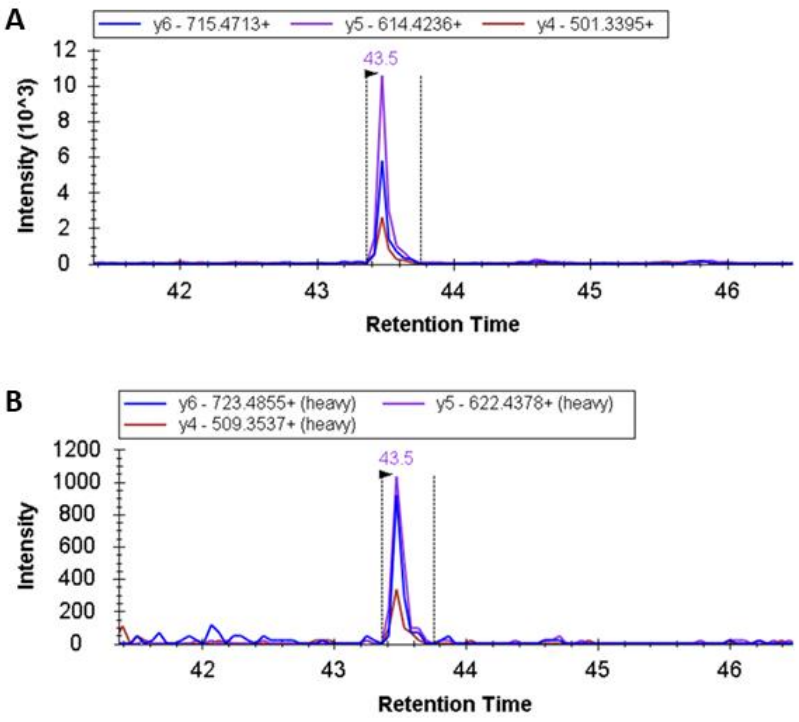


Figure 4.8 – Peak areas of each peptide transition were determined using Skyline software: (A) light peptide 1 and (B) heavy labeled peptide 1 for HCP_A.

The performance of each transition/peptide was evaluated according to the acceptance criteria defined for the calibration curve: CV (triplicate runs per point) < 25% and R^2 > 0.99. Peptide transitions that do not meet these criteria were excluded and not used for sample quantification (**Table 4.1** in red and **Figure 4.9**, referring to mAb1 experiment). The same evaluation was done for the other 3 mAb experiments (data not shown).

Table 4.**1** – List of HCP analysed and evaluation of their corresponding heavy peptides performance in the generated standard curves (mAb1 experiment) according to the acceptance criteria defined (CV triplicates < 25%, R^2>0.99).

HCP	Heavy Peptide	CV<25% (triplicates)	R^2>0.99	Remarks
HCP_A	Peptide 1	Yes	Yes	
	Peptide 2	Yes	Yes	
HCP_B	**Peptide 1**	**No**	**Yes**	CV > 25% in 1 and 10 fmol concentrations
	Peptide 2	Yes	Yes	Excluding 0.25 fmol
HCP_C	**Peptide 1**	**No**	**No**	CV > 25% in 1, 5 and 10 fmol concentrations
	Peptide 2	Yes	Yes	
HCP_D	Peptide 1	Yes	Yes	
	Peptide 2	Yes	Yes	Excluding 0.25 fmol

HCP_E	Peptide 1	Yes	Yes	Excluding 0.5 fmol
	Peptide 2	Yes	Yes	Excluding 0.25 fmol
HCP_F	Peptide 1	Yes	Yes	Excluding 0.25 fmol
	Peptide 2	Yes	Yes	Excluding 0.5 fmol
HCP_G	Peptide 1	Yes	Yes	Excluding 0.25 fmol
	Peptide 2	Yes	Yes	
HCP_H	Peptide 1	Yes	Yes	Excluding 0.25 fmol
	Peptide 2	Yes	Yes	
HCP_I	Peptide 1	Yes	Yes	

For each peptide, only one transition was chosen for quantification purposes (the final selection step of the best transition per peptide was based on signal intensity). Whenever possible, the quantification was also determined using a second peptide per HCP. Knowing that even from the same protein different peptides and transitions may present different ionization efficiency (varying their MS response intensity), it was important to guarantee that a minimum of 1 transition/1peptide was available for each HCP quantification.

After this thorough evaluation of peptides and their transitions in the 4 independent experiments performed, we proceeded for HCP quantification with the following number of heavy peptides: mAb1 experiment - 15 peptides, mAb2 experiment - 14 peptides, mAb3 experiment -16 peptides, mAb4 experiment - 16 peptides (data not shown). Some of the initially selected peptides, namely peptide 1 of both HCP_B and HCP_C appear to be particularly “problematic”, probably due to ionization issues, presenting a high variation between replicates and “non-linear behaviour” in their standard curves (**Figure 4.9**). Peptide 1 of HCP_B was rejected in all 4 independent experiments performed and peptide 1 of HCP_C in 2 of them.

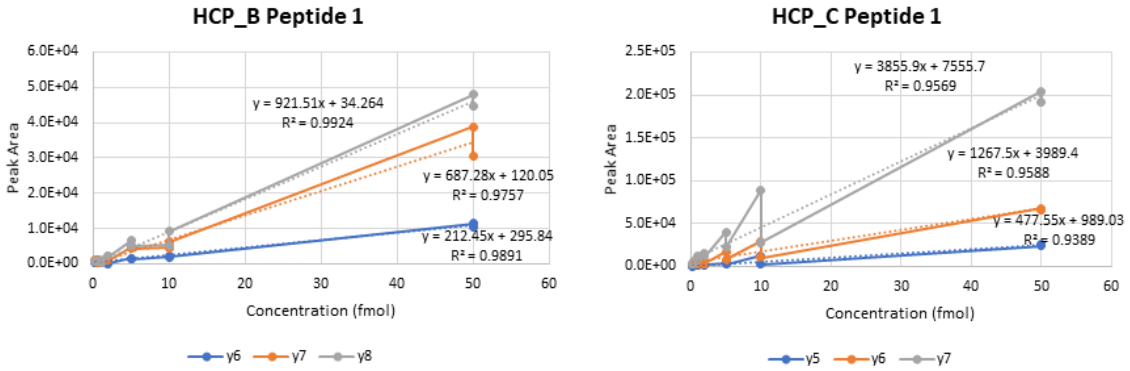


Figure 4.9 – Calibration curves of peptide 1 from both HCP_B and HCP_C. The 3 transitions analysed for each of the peptides are represented (data from mAb 1 experiment). For HCP_B peptide 1, only one transition meets the R^2 criterium, however, 2 calibration points (1 and 10 fmol) from this transition have a CV>25%. For HCP_C peptide 1, none of the 3 transitions passed the R^2 and CV criteria established.

From the standard calibration curves represented in **Figure 4.10** (mAb1 experiment), it is clear that some HCP for which 2 peptides were assessed, namely HCP_A, HCP_E and HCP_F, present very similar regression equations for both peptides under study. On the other hand, for

HCP_D, HCP_G and HCP_H, the differences in the behavior of the 2 peptides are quite obvious (very different slopes in the linear regression), consequently retrieving more variability in the quantification results. In this case, as we cannot guarantee which of the 2 peptides is more adequate for quantification, an average value may be considered to the respective HCP quantification. In an extended analysis, to further confirm the results obtained, if a third peptide is not available to validate which of the peptides retrieves a more accurate result, we may evaluate an extra quality criterium (as applied for MRM-QTRAP data), allowing a maximum ±30% variation between the quantification values obtained using peptide 1 and peptide 2.

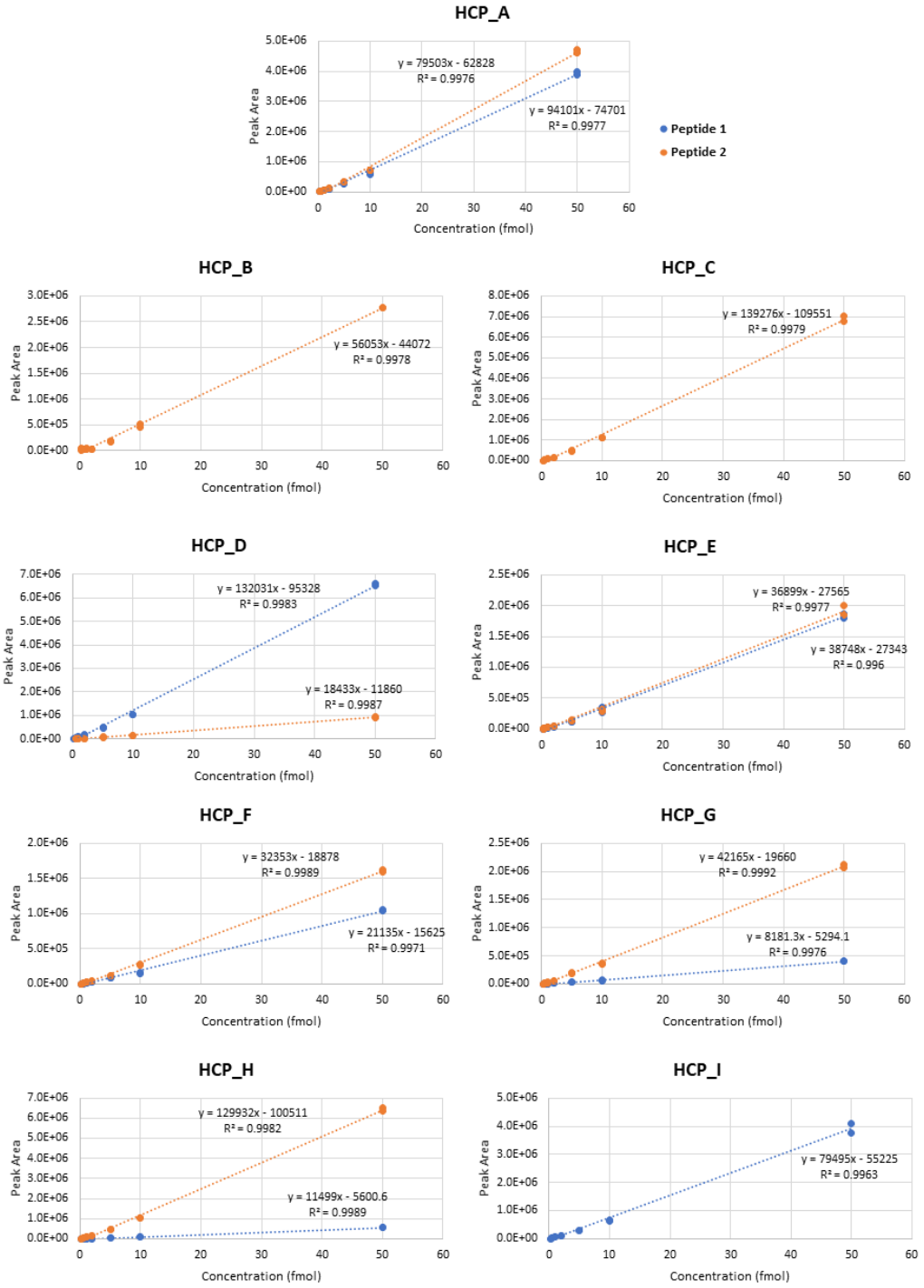


Figure 4.10 – Calibration curves for each HCP, using 1 or 2 proteotypic peptides selected according to the acceptance criteria described in **Table 4.1** (data from mAb 1 experiment).

4.2.2. Semi-absolute quantification using SWATH-MS

We also explored the SWATH-MS data with spiked heavy labeled peptides to establish a semi-absolute quantification method for untargeted HCP. For this, an MRM-like approach was applied after extraction of the MSMS data, using the information of the ion library previously generated. Therefore, with a single SWATH analysis, we are able to perform absolute quantification of targeted HCP and untargeted semi-absolute quantification of any HCP present in the ion library. For the latter approach, we evaluated the use of an average calibration curve using the peak areas of the previously selected heavy peptides (**Figure 4.11**). We hypothesized that, although the MS response intensity varies according to the peptide sequence, an average calibration curve with a large population of peptides would balance out the peptides' specific MS detection signal. A generic/"universal" calibration curve would be suitable for the quantification of any HCP in the samples under study, considering the risk of being more accurate for some proteins than for others, depending on the peptides used. Similar approaches using standard proteins (not related to the sample) spiked at defined concentrations in the sample preparation step, were already pursued by some authors showing some success [41]. This solution represents a great advantage since it allows to perform a semi-absolute quantification for those HCPs detected on the SWATH analysis without having their specific heavy peptides to perform quantification. To guarantee the suitability of this approach, we quantified target HCP using both the calibration curves with their specific heavy peptides and the average curve with previously selected peptides (14-16 peptides depending on the experiment, see section 4.2.1).

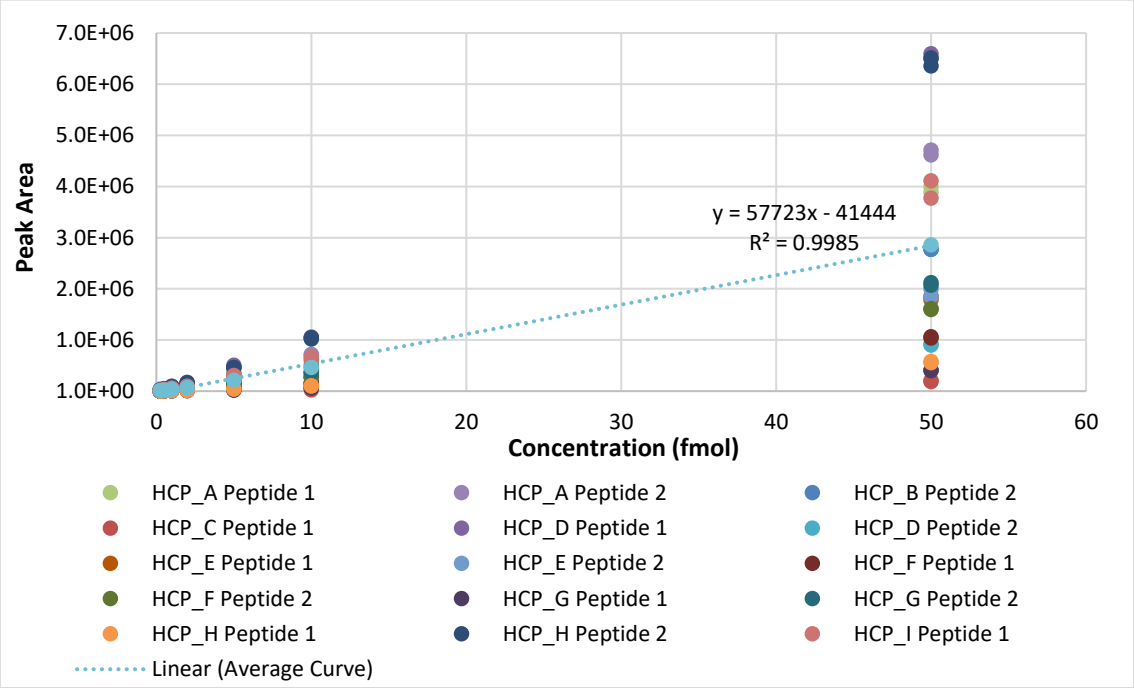


Figure 4.11 – Representative average calibration curve of mAb1 experiment with 15 peptides from 9 HCP. Linear regression represented refers to the average of each calibration point.

4.2.3. Selected Host Cell Proteins quantification

The quantification of HCP for the different mAbs under study, using either the specific heavy peptide calibration curve or the average curve generated with all heavy peptides (passing the quality criteria established), is presented and discussed in this section. The values obtained for PAE samples ("starting sample") using the absolute quantification method herein developed are shown in **Table 4.2**. We observed that the 4 PAE samples presented distinct levels of different HCPs from the start of the process. From the 9 HCP under focus in this analysis, mAb 3 PAE is where we identify more HCP (8 out of 9) and in higher quantity (in agreement with total HCP levels measured in the ELISA assay). Results also show that mAb2 and mAb4 samples seem to be the more "clean" ones, not only in terms of total HCP (ELISA results) but particularly in terms of the 9 targeted HCP, where we only detect HCP_A and HCP_H. Moreover, these two HCPs are the only targeted HCPs quantified in all the 4 mAbs. Contrarily, we were not able to quantify HCP_B in any of the studied samples. HCP_A, one of the most cited and recognized as problematic in mAb bioprocesses, seems to be the most abundant HCP in mAb1, mAb3 and mAb4 PAE samples in our study (**Table 4.2**).

Table 4.**2** – Absolute quantification of each HCP for the different PAE samples.

HCP	mAb 1 PAE		mAb 2 PAE		mAb 3 PAE		mAb 4 PAE	
	[Peptide 1] ppm	[Peptide 2] ppm	[Peptide 1] ppm	[Peptide 2] ppm	[Peptide 1] ppm	[Peptide 2] ppm	[Peptide 1] ppm	[Peptide 2] ppm
HCP_A	330.3	440.3	37.8	N/Q	1595.0	1306.1	200.1	196.30
HCP_B	-	N/Q	-	N/Q	-	N/Q	-	N/Q
HCP_C	-	N/Q	-	N/Q	190.6	72.8	N/Q	N/Q
HCP_D	N/Q	N/Q	N/Q	N/Q	124.3	118.1	N/Q	N/Q
HCP_E	31.8	26.3	N/Q	N/Q	518.8	393.7	N/Q	N/Q
HCP_F	48.0	N/Q	-	N/Q	407.4	168.8	N/Q	N/Q
HCP_G	41.2	19.6	N/Q	N/Q	OUT	882.2	N/Q	N/Q
HCP_H	47.4	N/Q	372.2	237.6	103.4	49.1	50.4	41.4
HCP_I	N/Q	-	N/Q	-	439.8	-	N/Q	-

Note: N/Q (not quantified); OUT (outside of range of calibration curve)

4.2.3.1. HCP_A

As mentioned above, HCP_A was detected and quantified in mAb1 (**Figure 4.12**), mAb2 (only in PAE sample, using peptide 1), mAb3 (**Figure 4.13**) and mAb4 (**Figure 4.14**) samples. For the 4 mAbs, we can observe that the absolute quantification of HCP_A is quite different already on PAE samples: mAb3 >>> mAb1 >> mAb4 > mAb2. Quantification results in each mAb are consistently comparable when using either the heavy peptide specific calibration curve or the

average curve. Main variations, as mentioned in the previous section, are observed between results from peptide 1 and peptide 2.

Overall, HCP_A concentration decreases from PAE to polishing step 1, and from polishing step 1 to polishing step 2, with polishing resin 2A showing the lowest amounts (**Figures 4.12**, **4.13** and **4.14**). In mAb1 and mAb4 samples, for most of the conditions, after polishing resin 2A, HCP_A is below quantification thresholds (< 53 ppm ≈0.8 fmol) or not detected (**Figures 4.12 B** and **4.14 B**). From the three resins, it is in polishing resin 1 that the different conditions tested seem to have a higher impact on HCP_A clearance (higher concentration variation between different pH & salt conditions).

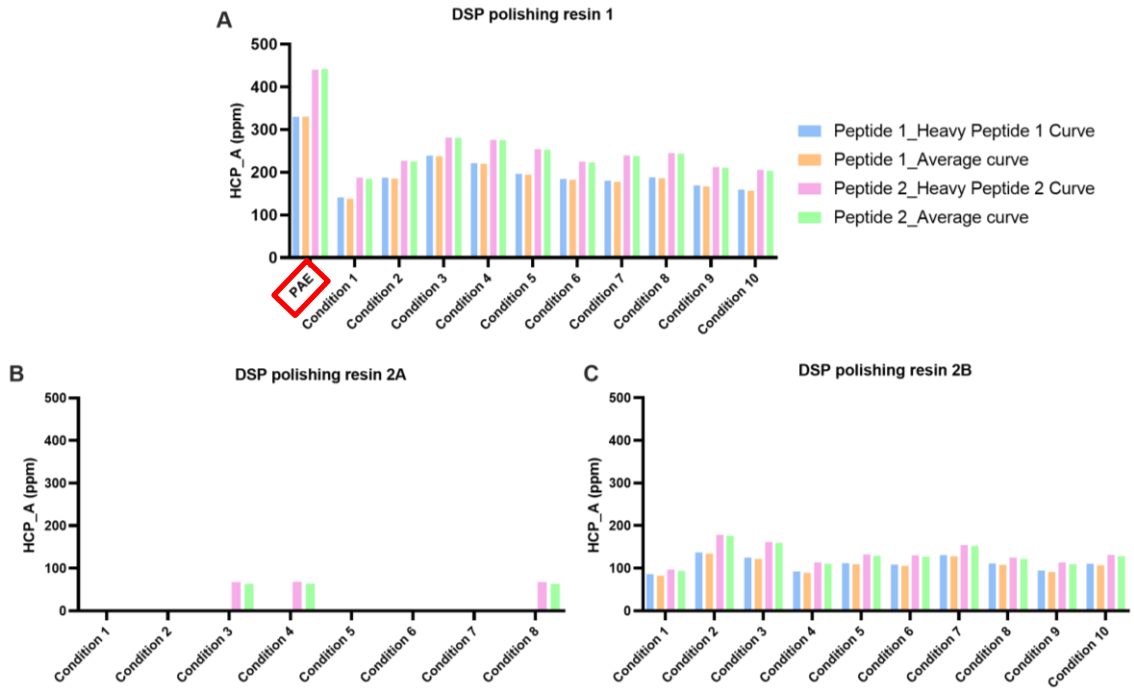


Figure 4.12 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_A in mAb1 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

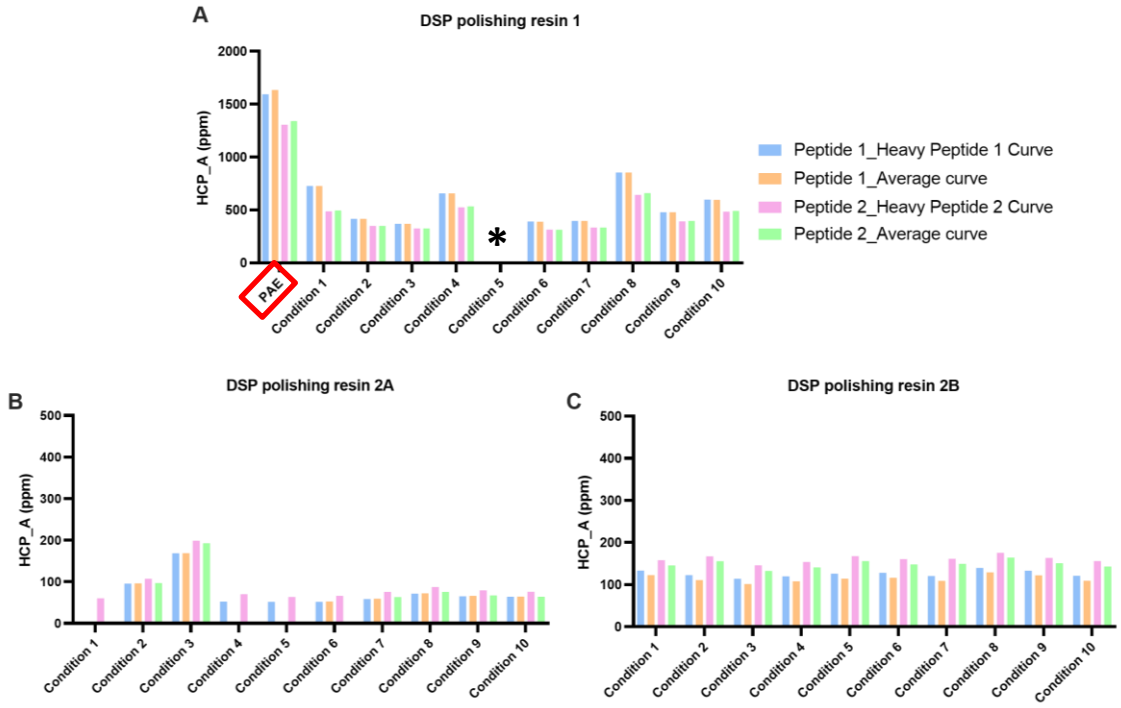


Figure 4.13 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_A in mAb3 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

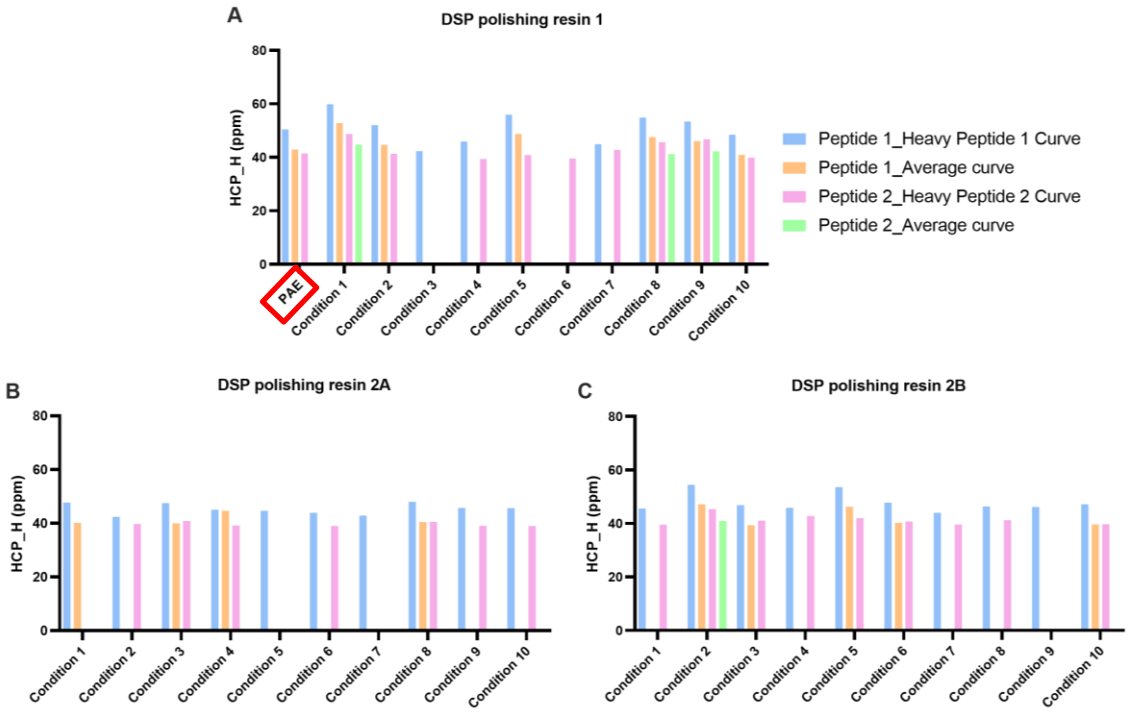


Figure 4.14 **–** SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_A in mAb4 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

4.2.3.2. HCP_C

HCP_C was only quantified in mAb3 samples. Contrarily to what we observed for HCP_A, HCP_C concentration in PAE sample is already quite low (≈200 ppm of HCP_C vs ≈1450 ppm of HCP_A). Moreover, we can also observe that after polishing resin 1 HCP_C values reduce to 100-50 ppm, and although no major improvements are obtained with resin 2B, with resin 2A HCP_C concentration decreases practically to undetectable levels (< 42 ppm ≈ 0.9 fmol) (**Figure 4.15**).

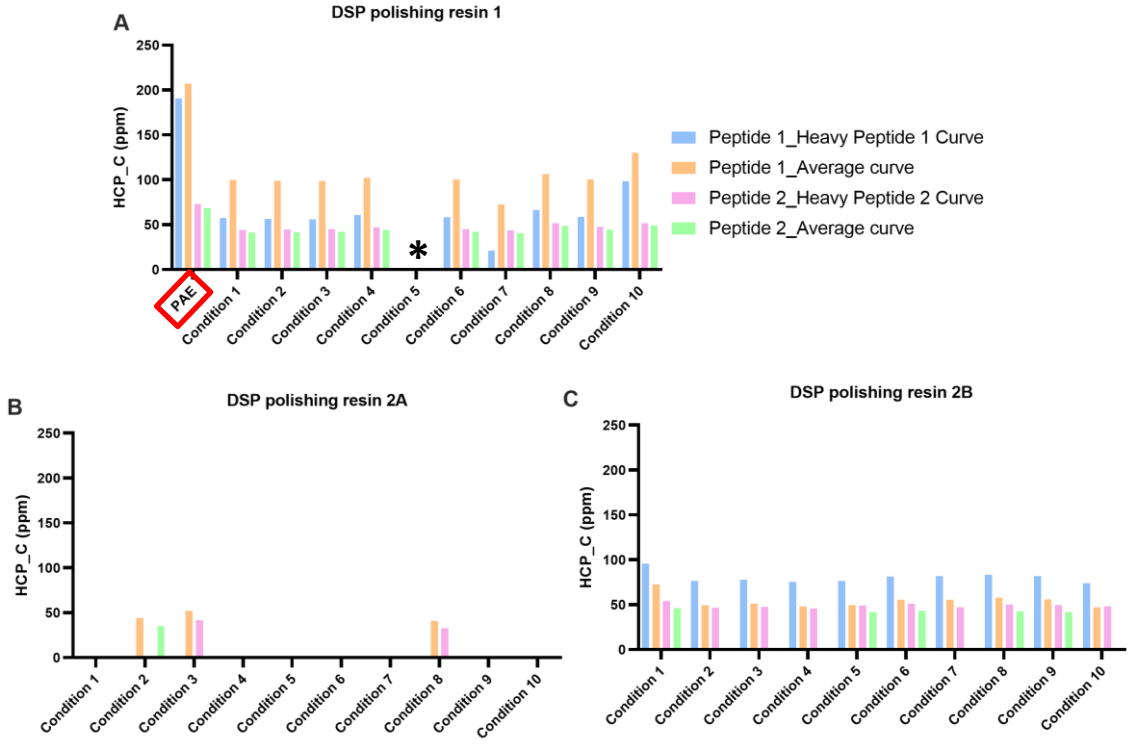


Figure 4.15 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_C in mAb3 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red). DSP polishing resin 1 condition 5 (*), sample lost due to injection issue.

4.2.3.3. HCP_D

HCP_D was only quantified in mAb3. The absolute quantification of HCP_D is in most of the conditions very similar to the quantification obtained by the average curve, especially in polishing resin 1 and peptide 1 in polishing resin 2B (**Figure 4.16**). HCP_D concentration decreases with sequential DSP steps. After polishing resin 2A, we can only detect both peptide 1 and peptide 2 through their specific averages curves (exception to condition 2 and 3) as we are close to the LOQ for this HCP (< 47 ppm ≈ 0.9 fmol). Therefore, small changes resulting from the average curve are enough to be below that limit.

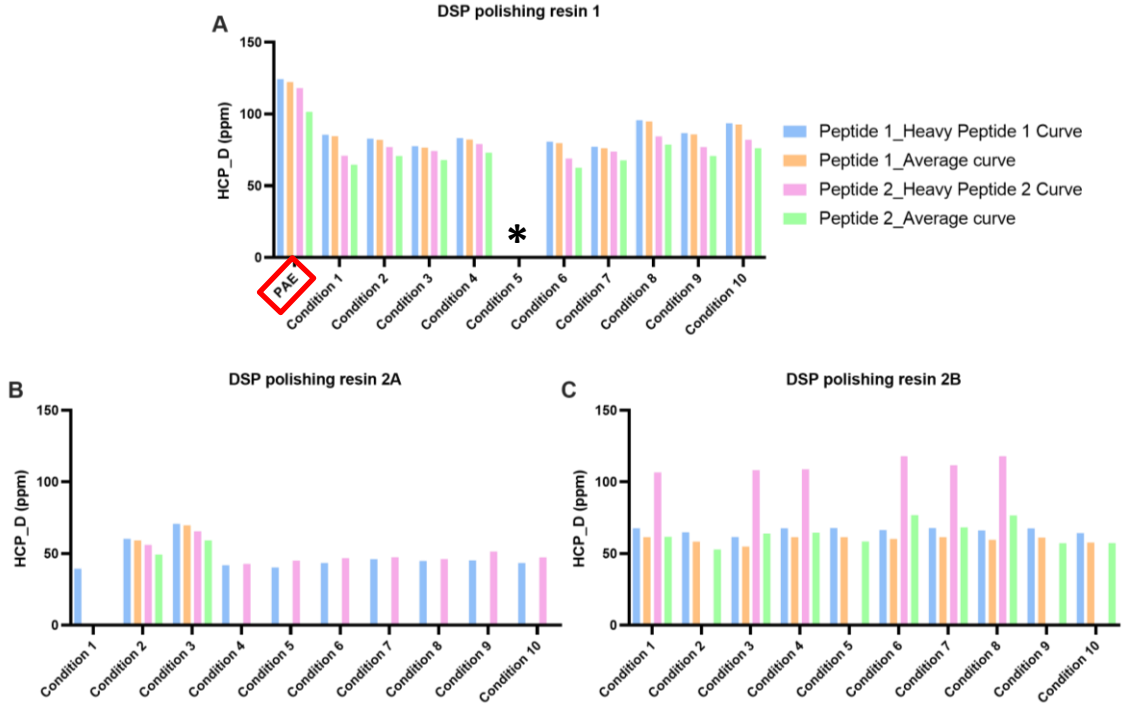


Figure 4.16 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_D in mAb3 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red). DSP polishing resin 1 condition 5 (*), sample lost due to injection issue.

4.2.3.4. HCP_E

The HCP_E was detected and quantified in mAb1 (**Figure 4.17**) and mAb3 (**Figure.4.18**), although in very different concentrations (more than 10x higher in mAb3). For the two mAbs, we can observe that the absolute quantification of HCP_E is very similar to the semi-quantification obtained by the average curve for both peptides. For mAb1, HCP_E concentration after both polishing resins 2A and 2B, for most of the pH and salt conditions tested, is below the LOQ (<19 ppm < 0.9 fmol). From mAb3 data, we observe a large variation between quantification values obtained with peptide 1 or peptide 2 (concentration ranges ≈500-100 ppm), which we did not observe for mAb1 (30-20 ppm range).

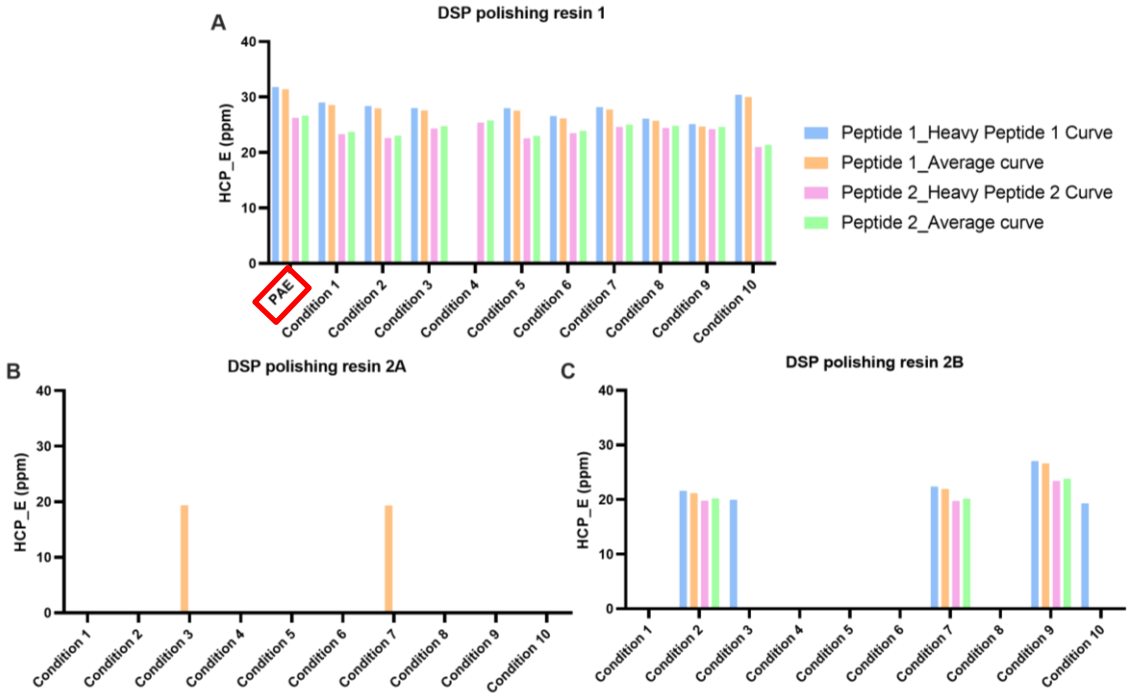


Figure 4.17 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_E in mAb1 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

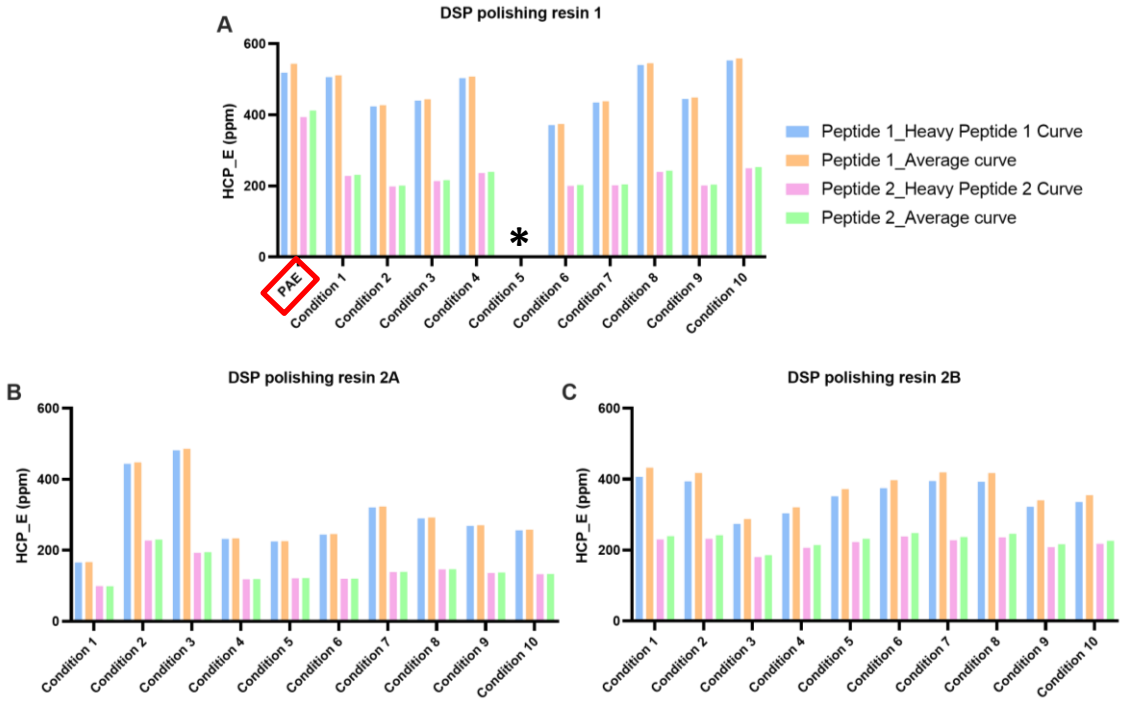


Figure 4.18 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_ E in mAb3 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red). DSP polishing resin 1 condition 5 (*), sample lost due to injection issue.

4.2.3.5. HCP_F

HCP_F was only quantified in mAb3. We can observe in **Figure 4.19** that the absolute quantification of HCP_F is very similar to the semi-absolute quantification obtained by the average curve for both peptides. Again, we observe major discrepancies between peptide 1 and peptide 2 quantification values at higher concentrations (1st polishing step, 400 – 100 ppm ≈ 7.4 - 1.8 fmol range) than in lower concentrations (2nd polishing step, <30 ppm ≈0.6 fmol range).

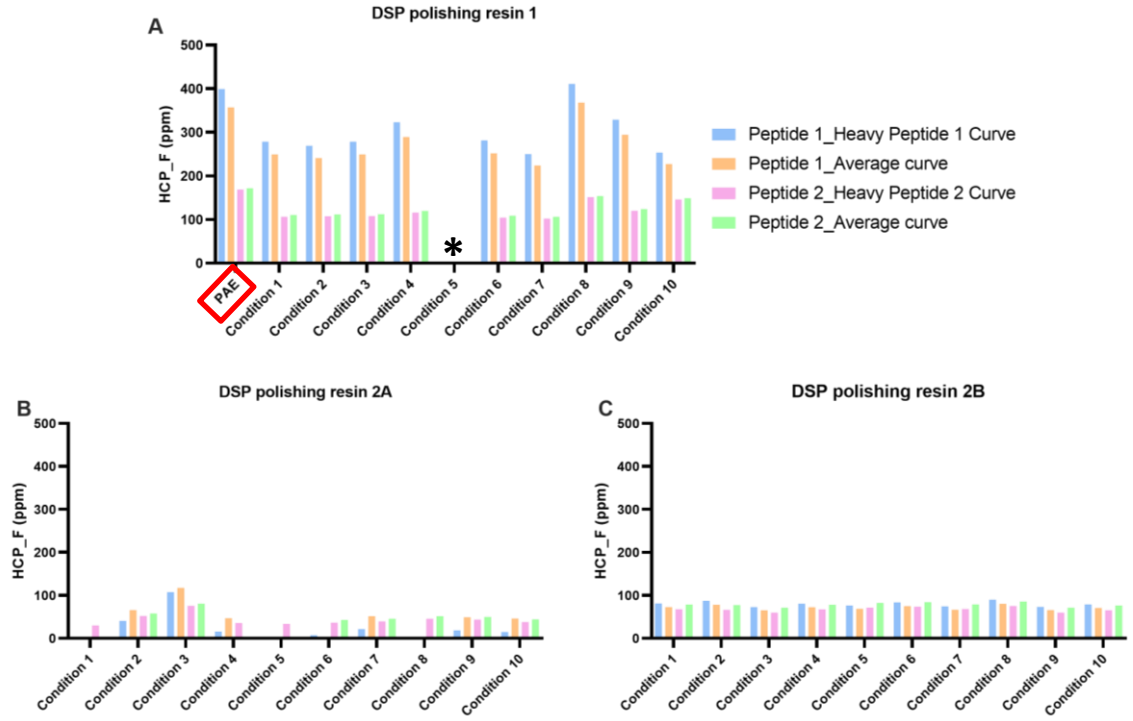


Figure 4.19 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_F in mAb3 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red). DSP polishing resin 1 condition 5 (*), sample lost due to injection issue.

4.2.3.6. HCP_G

The HCP_G was detected and quantified in mAb1 (**Figure 4.20**) and mAb3 (**Figure.4.21**) although in very different concentrations (around 25x higher in mAb3). Thus, in mAb1, after the 2nd polishing step, despite the resin or conditions, we cannot detect/quantify any HCP_G (< 18 ppm ≈0.6 fmol). On the contrary, for mAb3, in several conditions of polishing step 1 (and PAE), peptide 1 is outside the range of the calibration curve. Even after polishing step 2, HCP_G is still detected and presents higher concentrations in general after resin 2A. For the two mAbs, we can observe that the absolute quantification of HCP_G is very similar to the semi-quantification obtained by the average curve.

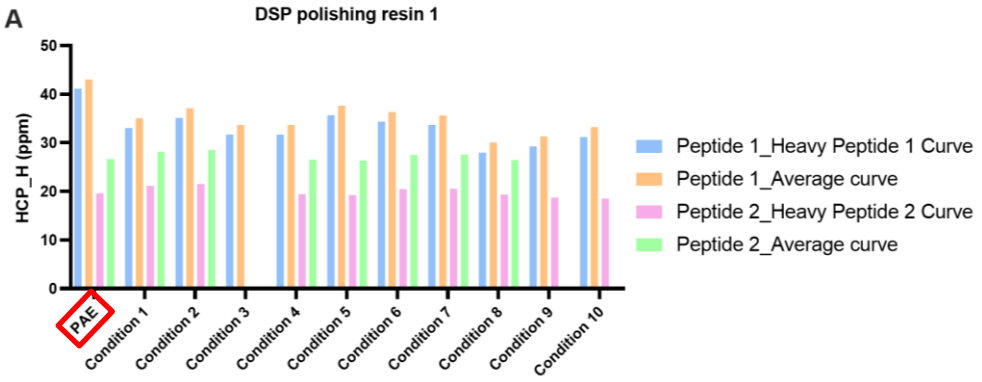


Figure 4.20 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_G in mAb1 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

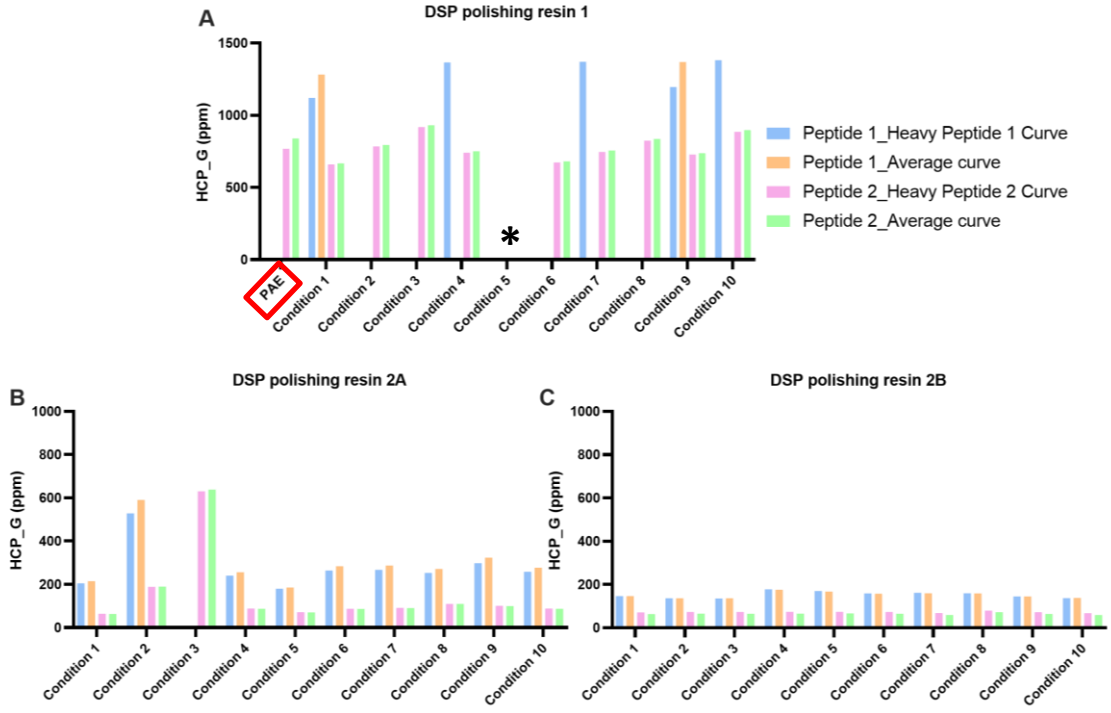


Figure 4.21 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_G in mAb3 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red). DSP polishing resin 1 condition 5 (*), sample lost due to injection issue. For DSP polishing resin 1, the quantification using peptide 1 (either the heavy peptide 1 curve and average curve) for sample PAE and condition 2,3,6 and 8 were outside of range of calibration curve.

4.2.3.7. HCP_H

HCP_H was detected and quantified in all mabs: mAb1 (**Figure 4.22**), mAb2 (**Figure 4.23**), mAb3 (**Figure 4.24**) and mAb4 (**Figure 4.25**). For the 4 mAbs, we can observe that the absolute quantification of HCP_H is quite different already on PAE samples: mAb2 >> mAb3 > mAb1 ≈ mAb4. From the 9 targeted HCP, HCP_H is the most abundant in mAb2 samples.

Quantification results in each mAb are consistently comparable when using either the heavy peptide specific calibration curve or the average curve. Main variations are observed between results from peptide 1 and peptide 2 (these results were expected from what we observed in HCP_H calibration curves of mAb1 experiment in **Figure 4.10**). Peptide 1 of HCP_H consistently interpolates higher concentration values than peptide 2. Surprisingly, for mAb4 results from peptide 1 and peptide 2 are very similar. HCP_H concentration in mAb1 and mAb4 samples is mostly stable across all DSP steps evaluated (variations in the range 40 - 30 ppm ≈ 1 - 0.6 fmol, close to LOQ, indeed in some conditions of polishing step 2 we were not able to quantify HCP_H),

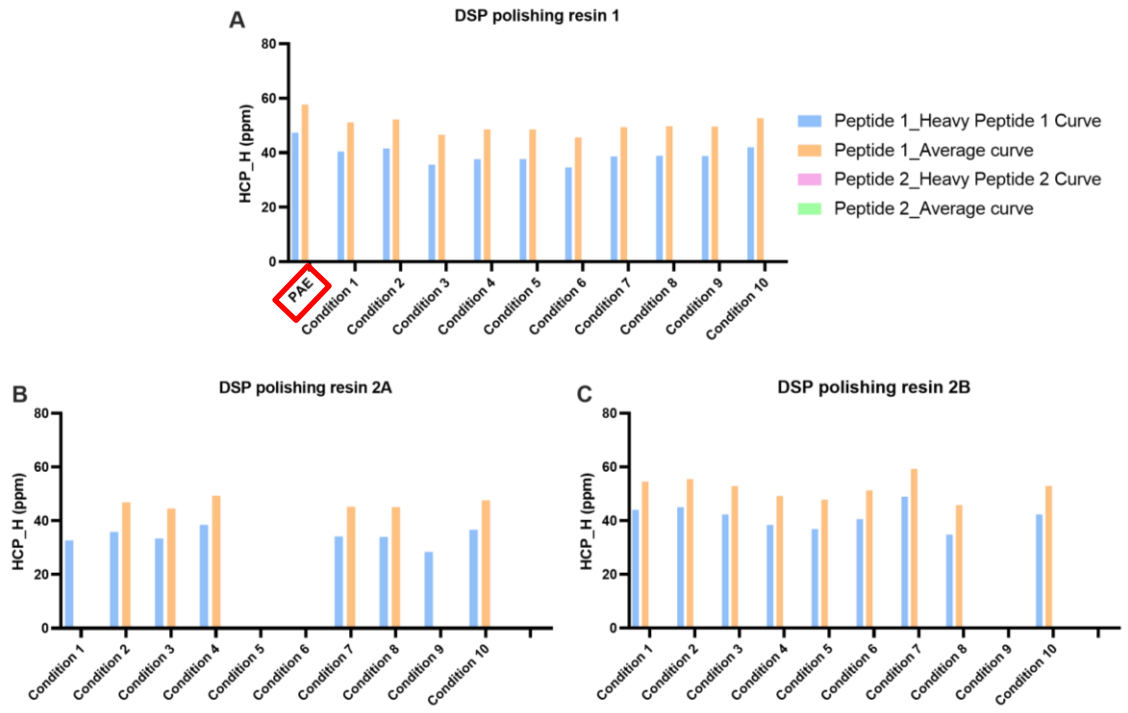


Figure 4.22 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_H in mAb1 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

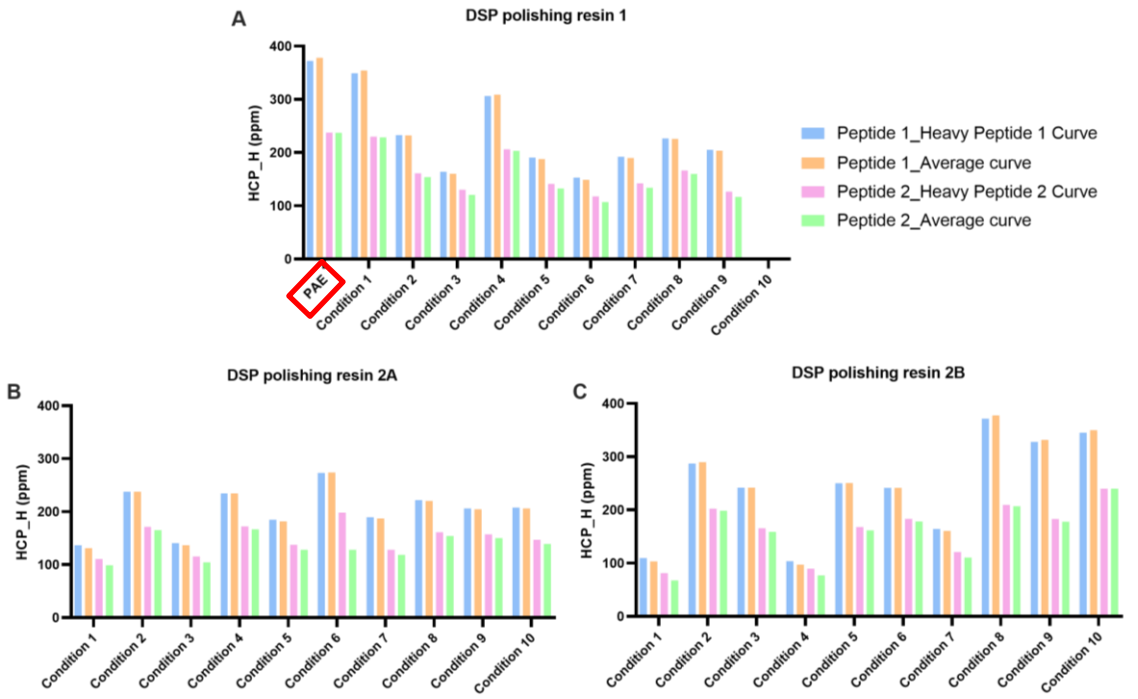


Figure 4.23 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_H in mAb2 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

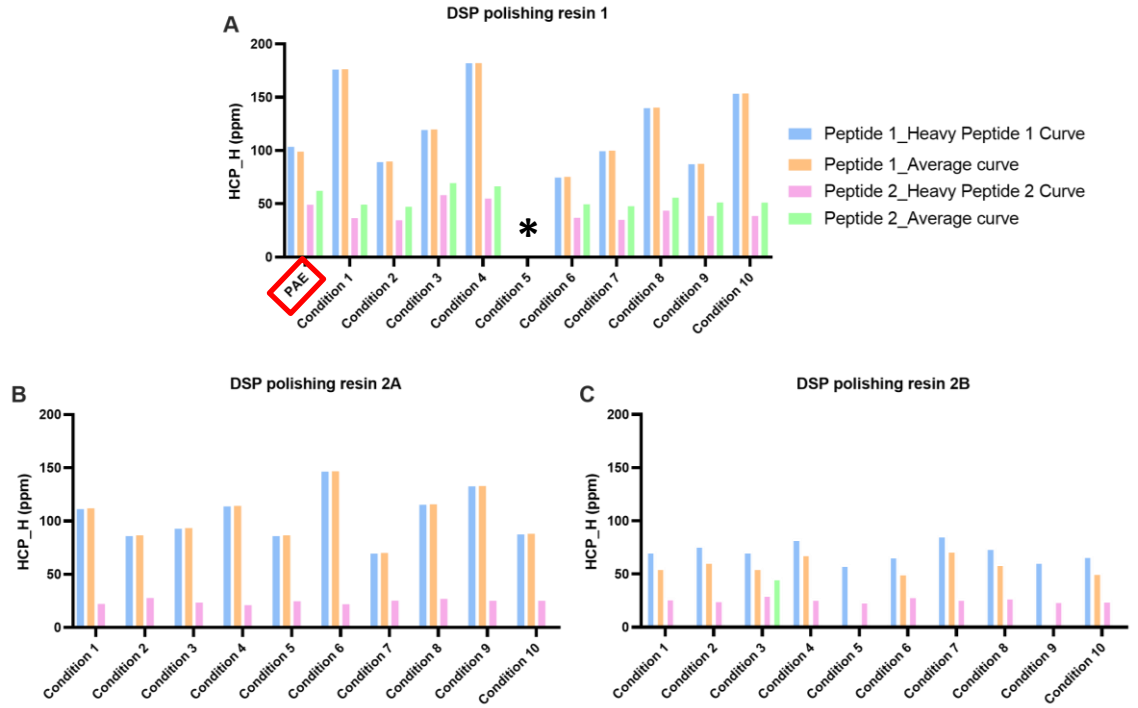


Figure 4.24 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_H in mAb3 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).). DSP polishing resin 1 condition 5 (*), sample lost due to injection issue.

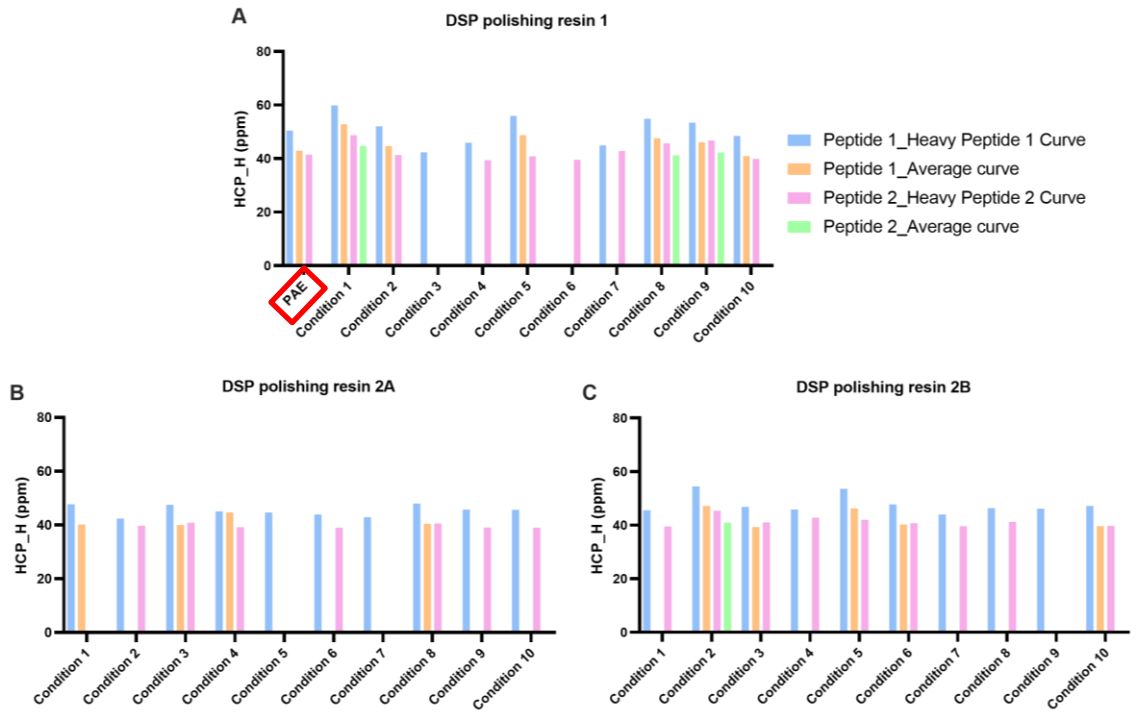


Figure 4.25 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_H in mAb4 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

4.2.3.8. HCP_I

HCP_I was only quantified in mAb3. For this HCP we only had available 1 heavy proteotypic peptide and, as can be observed in **Figure 4.26**, the absolute quantification of HCP_I is very similar to the quantification obtained by the average curve. HCP_I concentration after polishing resin 1 is quite variable depending on the conditions applied, condition 10 being the worst in terms of HCP_I clearance. The results of polishing resin 2B are homogeneous between conditions and similar to those obtained with resin 2A (around ≈200 ppm or lower), except for condition 1 (≈100 ppm ≈6.8 fmol) and condition 4 (< 11 ppm ≈ 0.7 fmol).

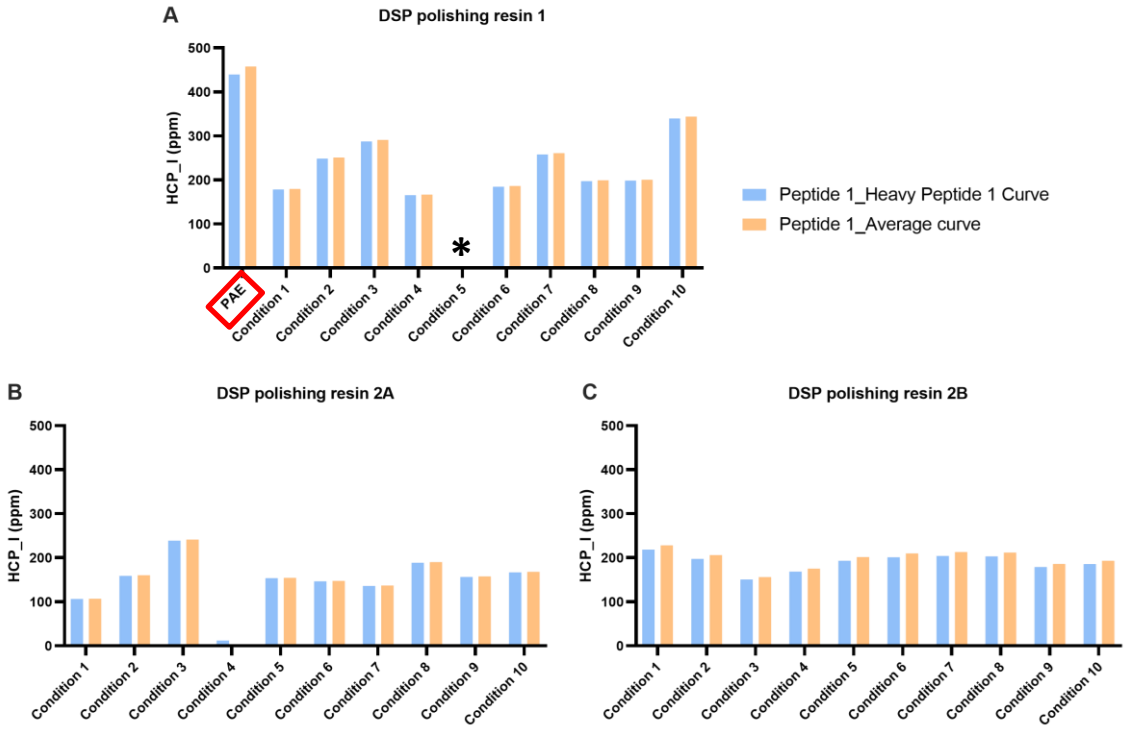


Figure 4.26 – SWATH absolute (heavy peptide specific calibration curves) and semi-absolute (average curve) quantification of HCP_I in mAb3 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).). DSP polishing resin 1 condition 5 (*), sample lost due to injection issue.

4.3. HCP_A quantification: Proof of concept

As referred above, HCP_A is one of the most common HCP found in mAb bioprocesses. In our study, this protein was also found in all PAE samples, being one of the most abundant of the 9 HCP studied. Given this, we used this HCP as our “proof of concept” to evaluate not only SWATH quantification precision, but also to implement an alternative targeted quantification method using the QTRAP 6500+ mass spectrometer system (the gold standard for MS-based absolute quantification of proteins), which enabled confirmation of our results.

To verify the precision of the SWATH absolute quantification method, the inter-assay standard deviation and coefficient of variation (from 4 independent digestions and 4 independent MS experiments run between June and October) were determined using data from one of the reference points of HCP_A standard curve (5 fmol, peptide 1). As shown in **Table 4.3**, for this specific peptide/HCP the variation between runs is quite small, with a CV lower than 4%, showing great robustness of the method and stability of the nanoLC-MS system.

Table 4.3 – Inter-assay (n=4) precision evaluation (HCP_A concentration was calculated with peptide 1 standard curve).

PAE mAb1 sample	[HCP_A] ppm
SWATH Run 1	322.4
SWATH Run 2	319.4
SWATH Run 3	339.3
SWATH Run 4	345.0
Mean	331.5
Standard deviation	12.6
CV (%)	3.8

To further evaluate the quantification results obtained from the SWATH-MS analysis, we implemented an alternative targeted quantification workflow for HCP_A. This procedure was based on LC-MS analysis using an MRM-workflow in a QTRAP-MS system. Currently, this is considered the most reliable analytic workflow for MS-based protein quantification. The same HCP_A peptides were used for the SWATH analysis, peptide1_transition y5 and peptide 2_transition y8. A calibration curve was constructed with 4 calibration points (1; 5; 10 and 50 fmol).

Figure 4.27 A shows the TIC of peptide 1 and peptide 2 from 7 different injections (unspiked mAb1 PAE sample and calibration curve points). As the method filters specific m/z ions (peptide ion in Q1 and fragments ions in Q3), only two signals were observed in the chromatogram corresponding to the two peptides at different retention times (RT). **Figure 4.27 B** represents the XIC for the light and heavy labeled peptides for HCP_A in the mAb1 PAE sample (heavy peptides spiked at 10 fmol). To account for technical variation (LC-MS system) triplicate injections were performed. As shown in **Figure 4.27 C**, highly reproducible technical replicates were obtained.

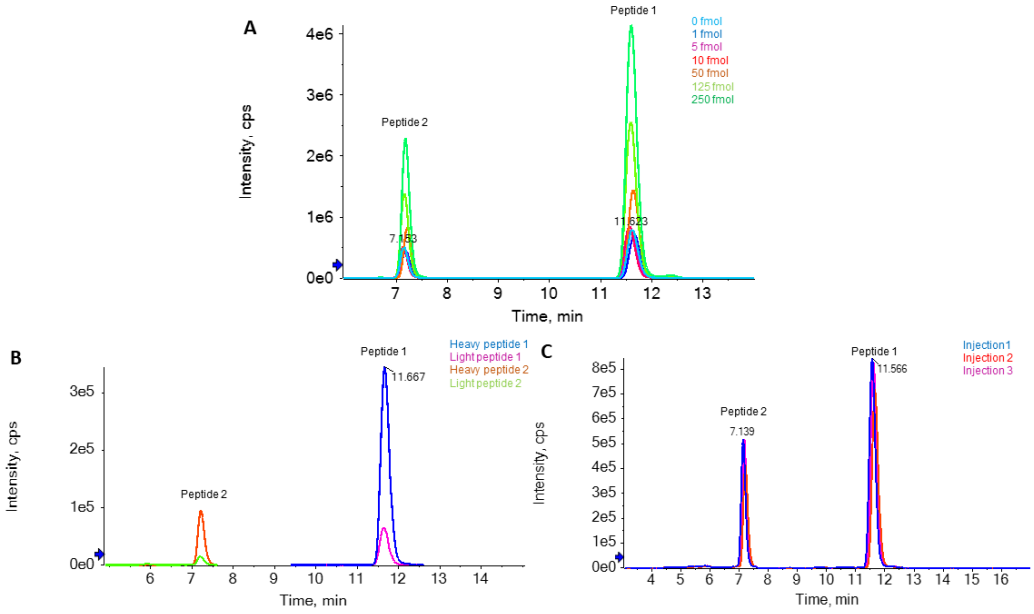


Figure 4.27 – MRM runs in the QTRAP system. (A) TIC of 7 different injection spiked with different amounts of both heavy proteotypic peptides of HCP_A (calibration curve points and unspiked mab1 PAE sample); (B) XIC for both light and heavy peptides of HCP_A; (C) Triplicates of injection of mAb1 PAE sample spiked with 10 fmol of the two heavy peptides.

As in the SWATH absolute quantification workflow, we evaluated the calibration curve of both peptides and different transitions to check CV thresholds (below 15%), R^2 >0.995 and signal intensity. As referred previously, to guarantee the accuracy of the results obtained from the two peptides evaluated we also considered a maximum of ±30% variation between the quantification using peptide 1 and peptide 2. Exceptions were accepted when the amount of HCP_A is less than 1 fmol (65.5 ppm) which is normally outside the linear range of peptide 2 in this system.

Similarly to the results obtained with the SWATH quantification, HCP_A was also quantified in all mAbs using the targeted MRM approach. Results for mAb1 (**Figure 4.28**), mAb2 (**Figure 4.29**), mAb3 (**Figure 4.30**) and mAb4 (**Figure 4.31**) are plotted below. We observe again a decrease in concentration in sequential DSP steps, and after polishing resin 2A the levels of HCP_A are very close or below quantification levels. Major variations in HCP_A clearance levels are observed within the different conditions tested in polishing resin 1. All results presented from both peptides 1 and 2 are within the 30% variation interval established.

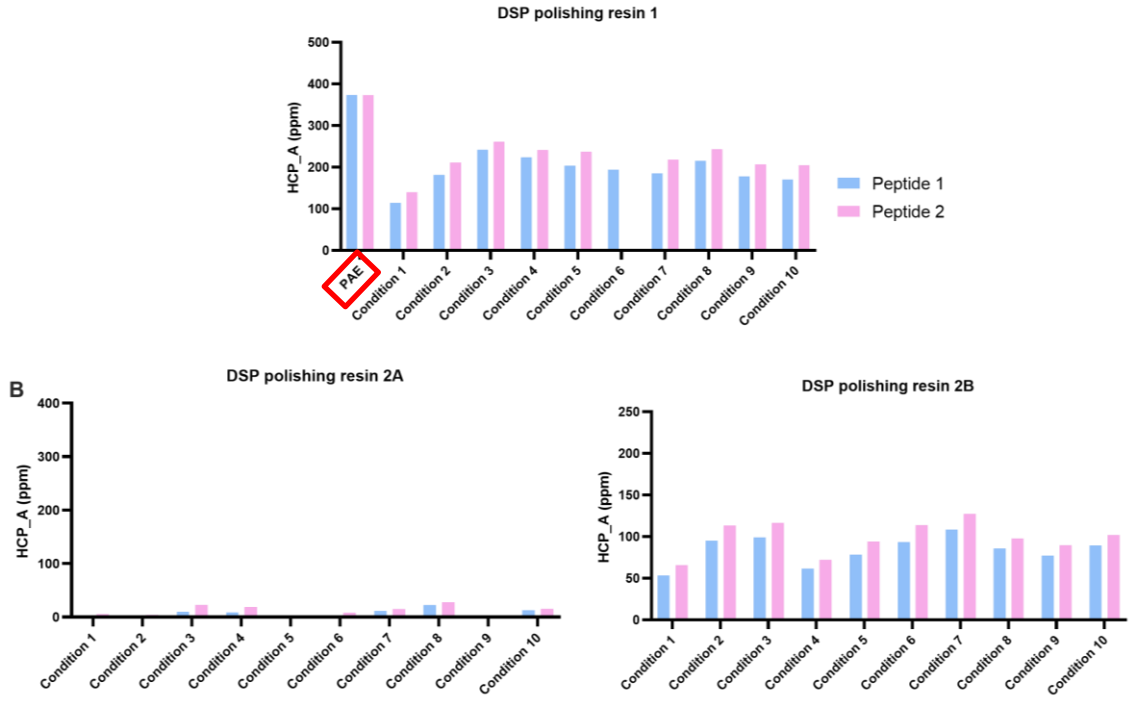


Figure 4.28 – QTRAP-MRM absolute quantification of HCP_A in mAb1 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

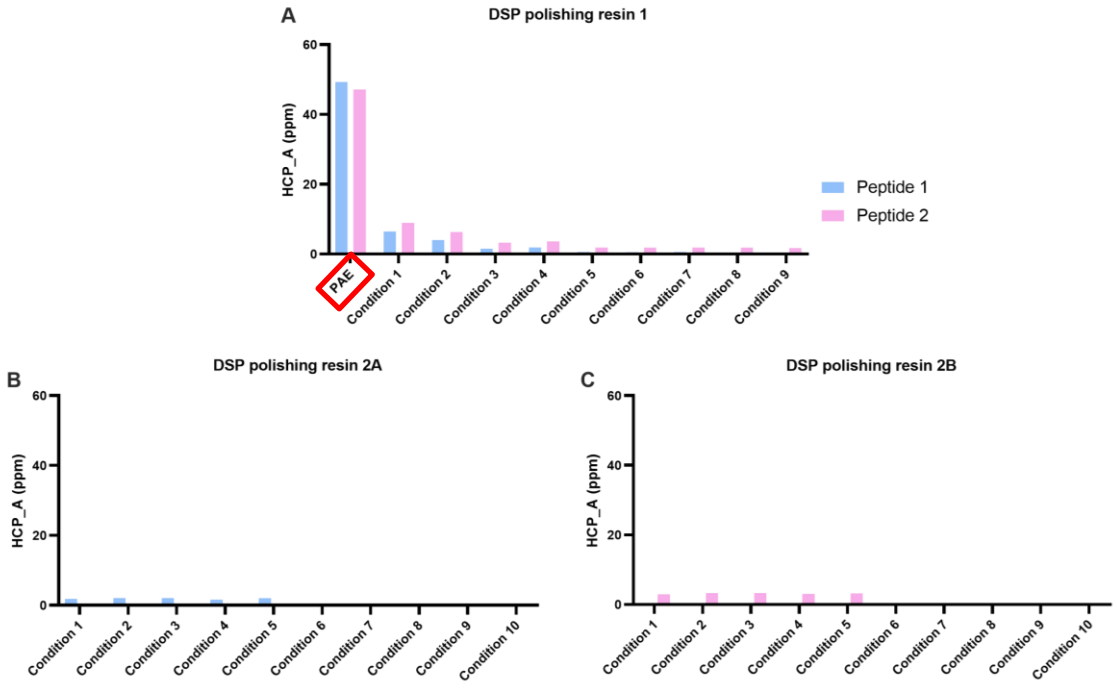


Figure 4.29 – QTRAP-MRM absolute quantification of HCP_A in mAb2 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

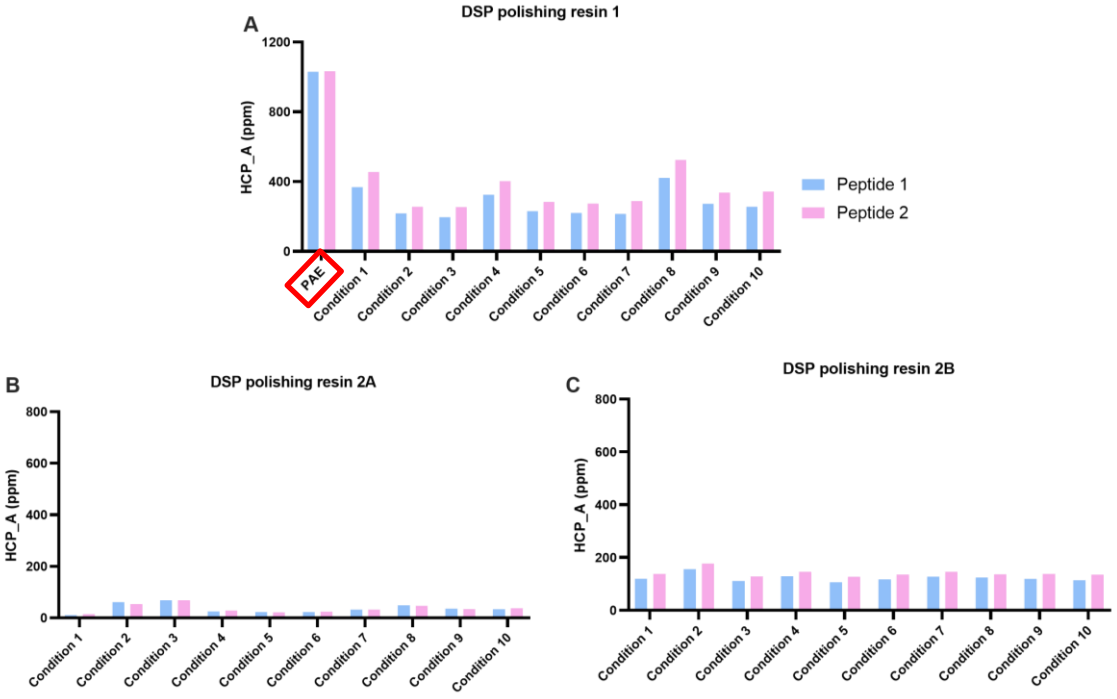


Figure 4.30 – QTRAP-MRM absolute quantification of HCP_A in mAb3 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

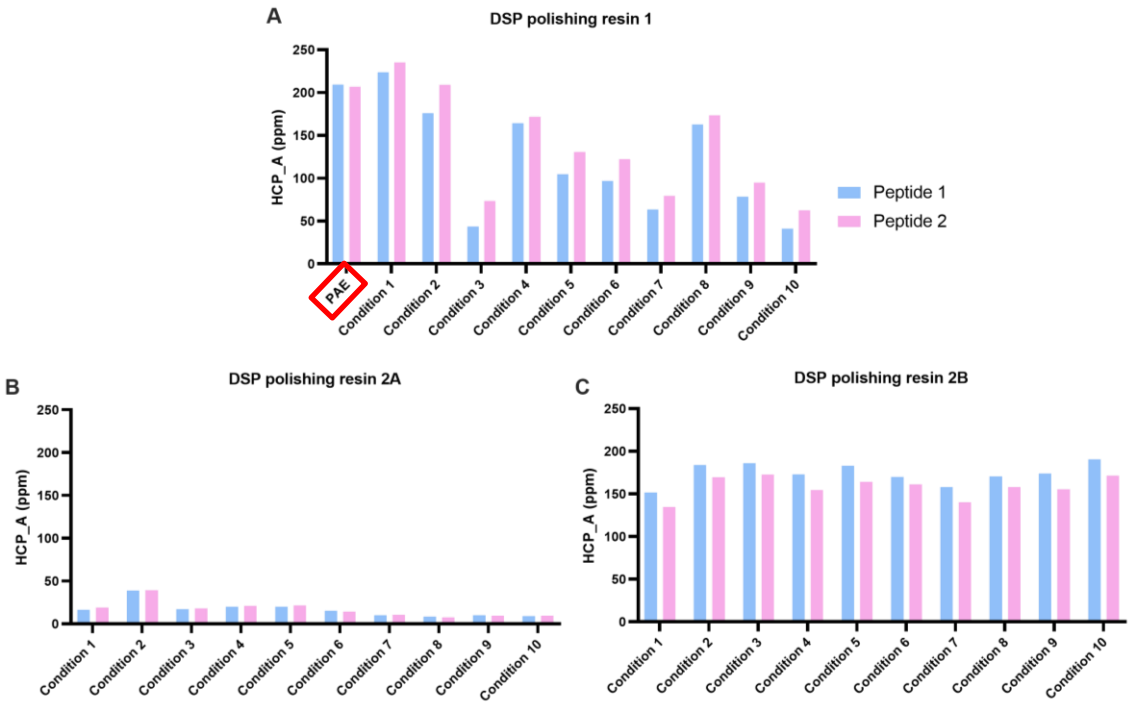


Figure 4.31 – QTRAP-MRM absolute quantification of HCP_A in mAb4 samples from three different DSP polishing steps (10 different pH & salt conditions per resin). PAE sample results are also represented for reference (highlighted in red).

4.4. SWATH quantification vs MRM absolute quantification of HCP_A

Comparing the results of **Figure 4.12** to **Figure 4.14** (HCP_A absolute and semi-absolute quantification by SWATH-MS) and **Figure 4.28** to **Figure 4.31** (targeted absolute quantification by MRM-QTRAP), we observe that similar trends in HCP removal were obtained with both methods, although some specific cases show high variations in absolute values calculated. Several reasons may contribute to the variations observed and we already identified some of those: a) the range of the calibration curves used in the two methods are very different, with SWATH starting at 0.25 or 0.5 fmol until 50 fmol, and MRM-QTRAP starting at 1 fmol up to 50 fmol), which may impact differently the accuracy of the results in certain concentration ranges and affecting precision between methods; b) mAb3 SWATH experiment was chronologically the last one to be performed. After careful data analysis we detected some signal decrease (lower peak areas, probably due to degradation) of the synthetic heavy peptides spiked in the samples (as IS) and used to build the calibration curves; this of course is impacting SWATH quantification values of mAb3 samples which are the ones presenting more variation.

In particular, when we compare the quantification of HCP_A in PAE samples of the 4 mAbs, obtained by the two acquisition methods explored plus the analysis using the average curve, we can see that for this HCP the values reported are very similar (**Table 4.4**), presenting a CV <25% (the threshold established for triplicate analysis in SWATH analysis). Nevertheless,

and in accordance with what we have just discussed, it was for the mAb3 sample that we observed the higher variation between methods.

Table 4.4 – Comparison of HCP A quantification in PAE samples by MRM-QTRAP and SWATH method (standard curves with specific heavy peptides 1 & 2 and average curve).

Sample	[HCP_A] ppm						CV (%)	
	MRM-QTRAP		SWATH approach Specific heavy peptide		SWATH approach Average curve			
	Peptide 1	Peptide 2	Peptide 1	Peptide 2	Peptide 1	Peptide 2	Peptide 1	Peptide 2
mAb 1 PAE	373.6	373.3	330.3	440.3	330.3	442.1	**7.2**	**9.4**
mAb 2 PAE	49.3	47.2	37.8	N/Q	N/Q	N/Q	**18.7**	-
mAb 3 PAE	1029.9	1032.3	1595	1306.1	1634.1	1340	**23.8**	**13.8**
mAb 4 PAE	209.4	206.7	200.1	193.3	179.5	178.5	**8.1**	**7.4**

Note: N/Q (not quantified)

Additionally, for a more complete comparison between methods, we should bear in mind that the sample prep used in both methods is the same, but the amount injected in SWATH-MS is 10x lower than in the QTRAP (1.6 µg vs 16 µg), demonstrating the higher sensitivity of the nanoLC system explored for SWATH. However, the data acquisition and data analysis time in SWATH is higher when compared to QTRAP (90 min gradient vs 20 min gradient). Nevertheless, SWATH-MS has a great advantage: while in QTRAP a limit number of only one (or a few) targeted HCP can be quantified, in SWATH-MS we showed the possibility to target several HCP (at least 9) in a single experiment, being able to perform either an absolute quantification (using specific heavy proteotypic peptides for the targeted HCP) and/or semi-absolute quantification using a "generic" average curve (using multiple heavy peptides from different HCPs). Additionally, there is the possibility of performing relative quantification of all HCP identified in the samples by IDA and present in the spectral library. This major advantage allows us to understand specific HCP abundance profile in a set of samples, even if that HCP was not targeted from the beginning as problematic/interesting to follow (and we do not have specific heavy peptides to quantify it). Moreover, we have shown that in those specific cases, the average standard curve may be applied to obtain a semi-absolute value.

4.5. Relative and semi-absolute quantification using SWATH-MS

As stated above, one major advantage of SWATH-MS is the possibility of performing the relative quantification of all HCP identified in the samples by IDA and present in the spectral library. In this work, we clearly detected one HCP (not targeted originally; hereinafter referred as HCP_X) that is present in all mAbs (in the Top 12 of HCP list) and whose relative abundance is highly increased in mAb3 samples (**Figure 4.32 A**). This (and others even in small amounts) may represent problematic HCP that would be undetectable if only a targeted MRM HCP profiling was performed. In addition, using the "generic" average curve generated, a semi-absolute

quantification of HCP_X was also performed (**Figure 4.32 B**). Similar HCP_X concentration values were obtained for mAb1, mAb2 and mAb4 (44.8, 20.1, 95.5 ppm, respectively), and a higher value, 1697.6 ppm, was determined for mAb3.

Further studies regarding the impact of this HCP in mAb quality, safety and efficacy should now be carried on to assess the criticality of this co-eluting protein. Moreover, understanding which HCP are co-eluting with the therapeutic molecule enables a more rational design of the DSP operations aiming at obtaining a better clearance of these process impurities, decreasing costs and timelines, and delivering a higher quality product.

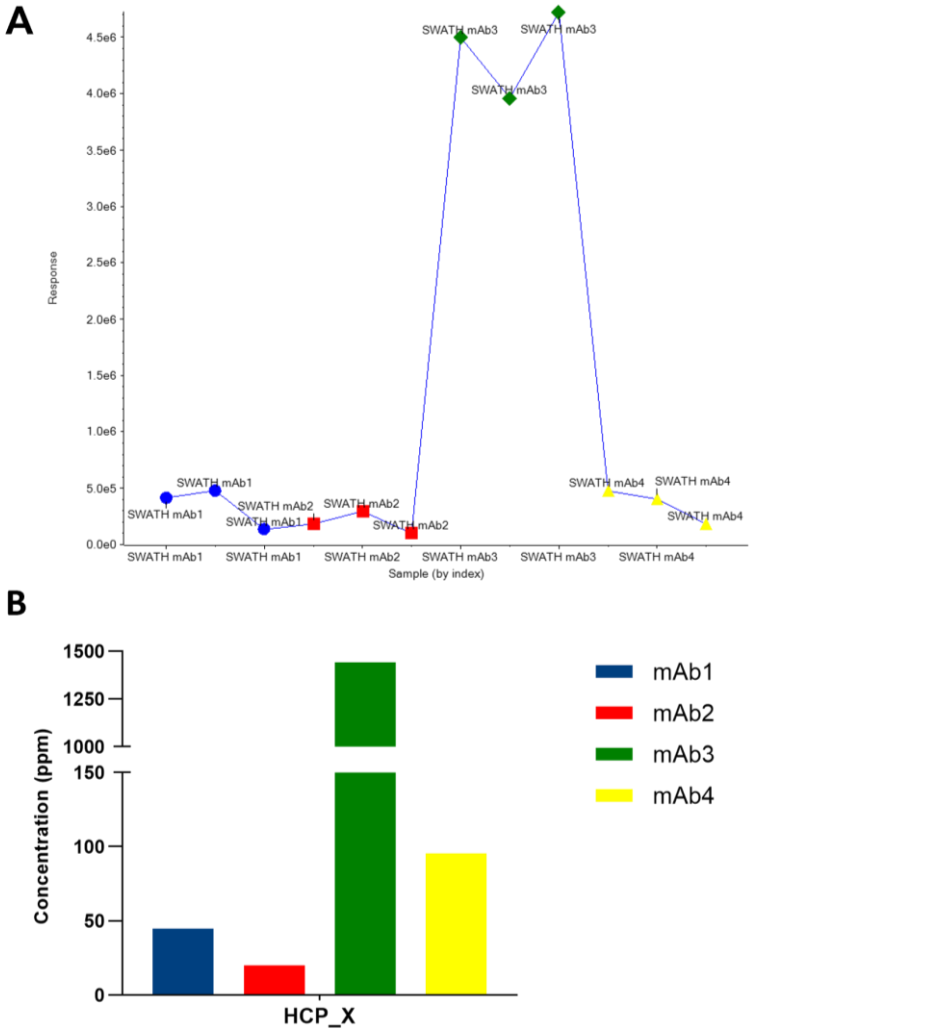


Figure 4.32 – Relative quantification (A) and semi-absolute quantification using the "generic" average curve (B) of HCP_X using SWATH-MS data (PAE samples of the 4 mAbs, run in triplicates).

5. Conclusion

The main objective of this thesis was to develop an analytical MS-based toolbox for the evaluation of HCP profiles in complex samples (using different mAb molecules at different DSP polishing steps - 1, 2A and 2B).

During this work, we firstly confirmed by ELISA that the samples under study contained HCP from CHO cells (the host cells where the mAbs were produced). Then, a method for absolute quantification of 9 selected HCP was developed exploring SWATH-MS technology and calibration curves using their specific heavy peptides. Additionally, we showed that we could also perform semi-absolute quantification for any other HCP present in the samples, without using their specific heavy peptides, by applying an average curve defined with the peptides of the selected 9 HCP. Moreover, the SWATH-MS spectral library generated with samples of the 4 mAbs, contained more than 200 HCPs derived from the CHO expression system, which we could relatively quantify across the different samples analysed and define the HCP removal profile along the DSP process.

To validate the results of the method developed and the different quantification approaches explored, we also implemented an MRM-based targeted method in the QTRAP system for quantification of one of the most common HCPs reported in the literature for this expression system, the HCP_A. Our data showed that the results obtained in QTRAP were similar to those obtained by SWATH quantification, either using HCP_A specific heavy peptides or the heavy peptides (from 9 HCP) average curve. The differences observed are acceptable and within the inherent variability of the analytical method.

When comparing the MRM-QTRAP MS workflow versus SWATH-MS, we found a greater advantage in using SWATH. Even though the data collection and data analysis are more time consuming (much more data is generated and consequently needs to be curated and analysed), using the same sample we can quantify (relative, semi-absolute and/or absolute quantification) several HCPs (target HCP or others found to be critical within SWATH data) instead of just one target HCP evaluated on the QTRAP.

ELISA-based HCP assays have been traditionally used for total HCP quantification. However, this method presents several drawbacks. SWATH strategy is an orthogonal approach to ELISA, as it can provide critical information on individual HCP identity and quantity, which cannot be obtained with the traditional method.

The SWATH quantification approach has the potential to become a powerful tool to overcome the challenges in HCP analysis during biologics' development and manufacturing stages. It enables higher throughput for monitoring and quantification of individual HCPs, improving process understanding of the prevalence of key HCPs and their clearance strategy.

Overall, we developed a successful methodology for absolute, semi-absolute (and relative) quantification of different HCPs in complex mAb samples (broad dynamic ranges and containing potentially hundreds of proteins), that can have an important impact on the development and optimization of mAb's bioprocessing.

6. References

[1] S. Ahmadi *et al.*, “Monoclonal antibodies expression improvement in CHO cells by PiggyBac transposition regarding vectors ratios and design,” *PLoS ONE*, vol. 12, no. 6, Jun. 2017, doi: 10.1371/JOURNAL.PONE.0179902.

[2] M. Grom, M. Kozorog, S. Caserman, A. Pohar, and B. Likozar, “Protein A affinity chromatography of Chinese hamster ovary (CHO) cell culture broths containing biopharmaceutical monoclonal antibody (mAb): Experiments and mechanistic transport, binding and equilibrium modeling,” *Journal of Chromatography B*, vol. 1083, pp. 44–56, Apr. 2018, doi: 10.1016/J.JCHROMB.2018.02.032.

[3] S. B. Carvalho *et al.*, “Multi attribute method implementation using a High Resolution Mass Spectrometry platform: From sample preparation to batch analysis,” *PLOS ONE*, vol. 17, no. 1, p. e0262711, Jan. 2022, doi: 10.1371/journal.pone.0262711.

[4] K. Pilely *et al.*, “Monitoring process-related impurities in biologics-host cell protein analysis”, doi: 10.1007/s00216-021-03648-2/Published.

[5] A. Beck, E. Wagner-Rousset, D. Ayoub, A. van Dorsselaer, and S. Sanglier-Cianférani, “Characterization of Therapeutic Antibodies and Related Products,” *Analytical Chemistry*, vol. 85, no. 2, pp. 715–736, Jan. 2012, doi: 10.1021/AC3032355.

[6] X. Wang, Z. An, W. Luo, N. Xia, and Q. Zhao, “Molecular and functional analysis of monoclonal antibodies in support of biologics development.,” *Protein Cell*, vol. 9, no. 1, pp. 74–85, Jan. 2018, doi: 10.1007/s13238-017-0447-x.

[7] Y. Lyubarskaya, K. Kobayashi, and P. Swann, “Application of mass spectrometry to facilitate advanced process controls of biopharmaceutical manufacture,” *undefined*, vol. 3, no. 4, pp. 313–321, Jun. 2015, doi: 10.4155/PBP.15.10.

[8] J. Sjögren, F. Olsson, and A. Beck, “Rapid and improved characterization of therapeutic antibodies and antibody related products using IdeS digestion and subunit analysis,” *Cite this: Analyst*, vol. 141, p. 3114, 2016, doi: 10.1039/c6an00071a.

[9] C. H. Goey, S. Alhuthali, and C. Kontoravdi, “Host cell protein removal from biopharmaceutical preparations: Towards the implementation of quality by design,” *Biotechnology Advances*, vol. 36, no. 4, pp. 1223–1237, Jul. 2018, doi: 10.1016/j.biotechadv.2018.03.021.

[10] A. L. Tscheliessnig, J. Konrath, R. Bates, and A. Jungbauer, “Host cell protein analysis in therapeutic protein bioprocessing – methods and applications,” *Biotechnology Journal*, vol. 8, no. 6, pp. 655–670, Jun. 2013, doi: 10.1002/BIOT.201200018.

[11] X. Wang, A. K. Hunter, and N. M. Mozier, “Host cell proteins in biologics development: Identification, quantitation and risk assessment,” *Biotechnol Bioeng*, vol. 103, no. 3, pp. 446–458, Jun. 2009, doi: 10.1002/BIT.22304.

[12] J. Briggs and P. R. Panfili, “Quantitation of DNA and protein impurities in biopharmaceuticals,” *Anal Chem*, vol. 63, no. 9, pp. 850–859, May 1991, doi: 10.1021/AC00009A003.

[13] S. Clavier *et al.*, “Improving the analytical toolbox to investigate copurifying host cell proteins presence: N-(4)-(β-acetylglucosaminyl)-L-asparaginase case study,” *Biotechnology and Bioengineering*, vol. 117, pp. 3368–3378, 2020, doi: 10.1002/bit.27514.

[14] C. Ludwig, L. Gillet, G. Rosenberger, S. Amon, B. C. Collins, and R. Aebersold, “Data-independent acquisition-based SWATH - MS for quantitative proteomics: a tutorial,” *Molecular Systems Biology*, vol. 14, no. 8, Aug. 2018, doi: 10.15252/MSB.20178126/FORMAT/PDF.

[15] E. Engvall and P. Perlmann, “Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G,” *Immunochemistry*, vol. 8, no. 9, pp. 871–874, 1971, doi: 10.1016/0019-2791(71)90454-X.

[16] “The principle and method of ELISA | MBL Life Science -JAPAN-.” https://ruo.mbl.co.jp/bio/e/support/method/elisa.html (accessed Nov. 11, 2021).

[17] “Overview of ELISA | Thermo Fisher Scientific - PT.” https://www.thermofisher.com/pt/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-elisa.html (accessed Nov. 11, 2021).

[18] “CHO HCP ELISA Kit, 3G.” https://www.cygnustechnologies.com/cho-hcp-elisa-kit-3g-resupply.html (accessed Nov. 11, 2021).

[19] S. Gilgunn and J. Bones, “Challenges to industrial mAb bioprocessing—removal of host cell proteins in CHO cell bioprocesses,” *Current Opinion in Chemical Engineering*, vol. 22, pp. 98–106, Dec. 2018, doi: 10.1016/j.coche.2018.08.001.

[20] S. Gilgunn *et al.*, “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*, vol. 1595, pp. 28–38, Jun. 2019, doi: 10.1016/j.chroma.2019.02.056.

[21] S. Flatman, I. Alam, J. Gerard, and N. Mussa, “Process analytics for purification of monoclonal antibodies,” *J Chromatogr B Analyt Technol Biomed Life Sci*, vol. 848, no. 1, pp. 79–87, Mar. 2007, doi: 10.1016/J.JCHROMB.2006.11.018.

[22] C. Ludwig, L. Gillet, G. Rosenberger, S. Amon, B. C. Collins, and R. Aebersold, “Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial,” *Molecular Systems Biology*, vol. 14, no. 8, p. e8126, Aug. 2018, doi: 10.15252/MSB.20178126.

[23] R. Aebersold and M. Mann, “Mass spectrometry-based proteomics,” *Nature 2003 422:6928*, vol. 422, no. 6928, pp. 198–207, Mar. 2003, doi: 10.1038/nature01511.

[24] F. Xie, T. Liu, W. J. Qian, V. A. Petyuk, and R. D. Smith, “Liquid Chromatography-Mass Spectrometry-based Quantitative Proteomics *,” *Journal of Biological Chemistry*, vol. 286, no. 29, pp. 25443–25449, Jul. 2011, doi: 10.1074/JBC.R110.199703.

[25] T. Shi *et al.*, “Advances in targeted proteomics and applications to biomedical research,” *PROTEOMICS*, vol. 16, no. 15–16, pp. 2160–2182, Aug. 2016, doi: 10.1002/PMIC.201500449.

[26] L. C. Gillet *et al.*, “Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis *,” *Molecular & Cellular Proteomics*, vol. 11, no. 6, Jun. 2012, doi: 10.1074/MCP.O111.016717.

[27] AB SCIEX, “TripleTOF ® 6600+ System System User Guide,” 2019.

[28] AB SCIEX, “6500 and 6500+ Series of Instruments System User Guide,” 2015.

[29] “Multiple Reaction Monitoring Mass Spectrometry (MRM-MS). A schematic of... | Download Scientific Diagram.” https://www.researchgate.net/figure/Multiple-Reaction-Monitoring-Mass-Spectrometry-MRM-MS-A-schematic-of-a-triple_fig3_51034246 (accessed Nov. 11, 2021).

[30] C. Lindemann *et al.*, “Strategies in relative and absolute quantitative mass spectrometry based proteomics,” *Biological Chemistry*, vol. 398, no. 5–6, pp. 687–699, May 2017, doi: 10.1515/HSZ-2017-0104/MACHINEREADABLECITATION/RIS.

[31] M. Bantscheff, S. Lemeer, M. M. Savitski, and B. Kuster, “Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present”, doi: 10.1007/s00216-012-6203-4.

[32] M. C. Wiener, J. R. Sachs, E. G. Deyanova, and N. A. Yates, “Differential Mass Spectrometry: A Label-Free LC-MS Method for Finding Significant Differences in Complex Peptide and Protein Mixtures,” 2004, doi: 10.1021/ac0493875.

[33] P. Mallick *et al.*, “Computational prediction of proteotypic peptides for quantitative proteomics,” *Nature Biotechnology*, vol. 25, no. 1, pp. 125–131, Jan. 2007, doi: 10.1038/nbt1275.

[34] M. Nikolov, C. Schmidt, and H. Urlaub, “Quantitative Mass Spectrometry-Based Proteomics: An Overview,” *Methods in Molecular Biology*, vol. 893, pp. 85–100, 2012, doi: 10.1007/978-1-61779-885-6_7.

[35] J. Vowinckel, F. Capuano, K. Campbell, M. J. Deery, K. S. Lilley, and M. Ralser, “The beauty of being (label)-free: sample preparation methods for SWATH-MS and next-generation targeted proteomics,” *F1000Res*, vol. 2, Apr. 2013, doi: 10.12688/F1000RESEARCH.2-272.V2.

[36] L. Tuli and H. W. Ressom, “LC–MS Based Detection of Differential Protein Expression,” *J Proteomics Bioinform*, vol. 2, no. 10, p. 416, Oct. 2009, doi: 10.4172/JPB.1000102.

[37] S. Cappadona, P. R. Baker, P. R. Cutillas, A. J. R. Heck, and B. van Breukelen, “Current challenges in software solutions for mass spectrometry-based quantitative proteomics,” *Amino Acids 2012 43:3*, vol. 43, no. 3, pp. 1087–1108, Jul. 2012, doi: 10.1007/S00726-012-1289-8.

[38] D. E. Walker, F. Yang, J. Carver, K. Joe, D. A. Michels, and X. C. Yu, “A modular and adaptive mass spectrometry-based platform for support of bioprocess development

toward optimal host cell protein clearance," *MAbs*, vol. 9, no. 4, pp. 654–663, May 2017, doi: 10.1080/19420862.2017.1303023.

[39] K. H. Sim *et al.*, "A comprehensive CHO SWATH-MS spectral library for robust quantitative profiling of 10,000 proteins," *Scientific Data*, vol. 7, no. 1, p. 263, Dec. 2020, doi: 10.1038/s41597-020-00594-z.

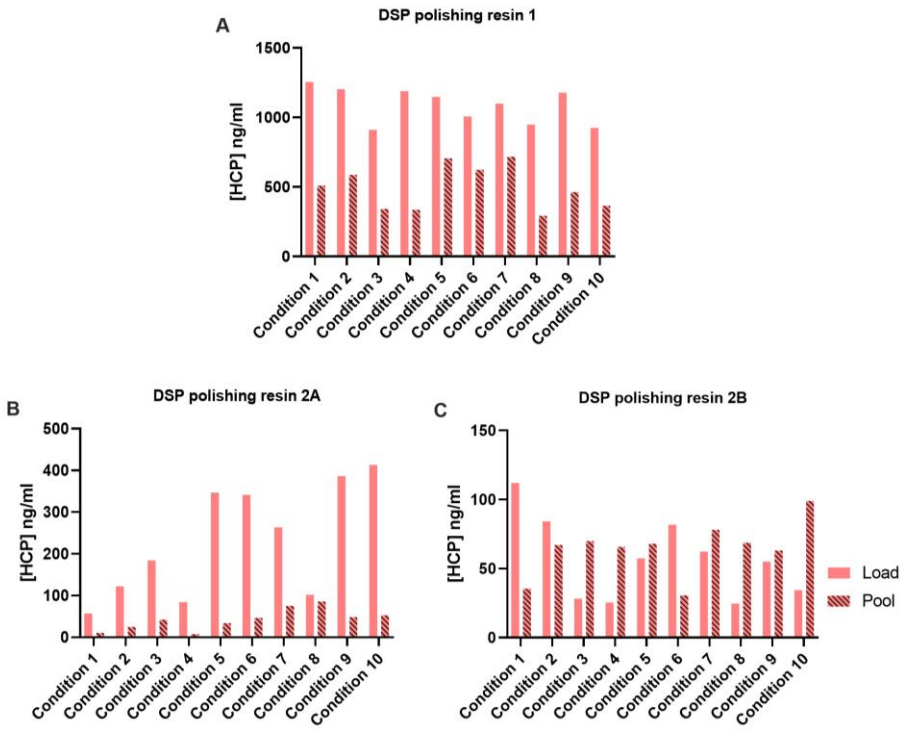
[40] Y. Huang, R. Molden, M. Hu, H. Qiu, and N. Li, "Toward unbiased identification and comparative quantification of host cell protein impurities by automated iterative LC–MS/MS (HCP-AIMS) for therapeutic protein development," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 200, p. 114069, Jun. 2021, doi: 10.1016/j.jpba.2021.114069.

[41] K. Pilely *et al.*, "Monitoring process-related impurities in biologics–host cell protein analysis," *Analytical and Bioanalytical Chemistry*, vol. 414, no. 2, pp. 747–758, Jan. 2022, doi: 10.1007/s00216-021-03648-2.

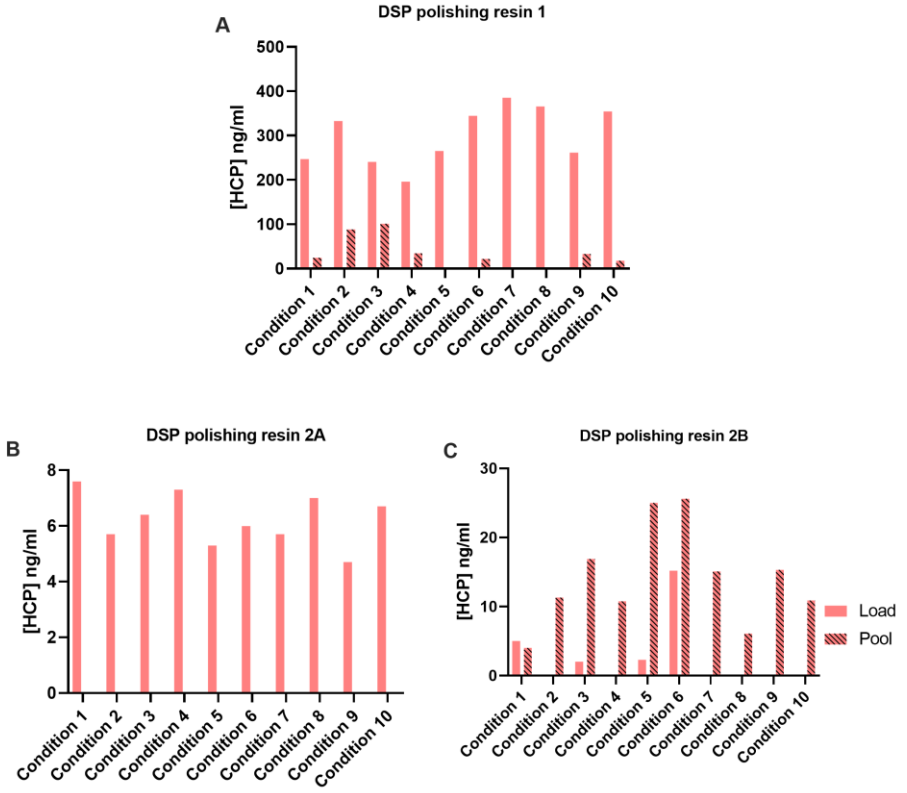
7. Appendices

Appendix 7.1 – Conversion between fmol (obtained directly from the calibration curve) and the ppm level for each HCP (correlates with their molecular mass)

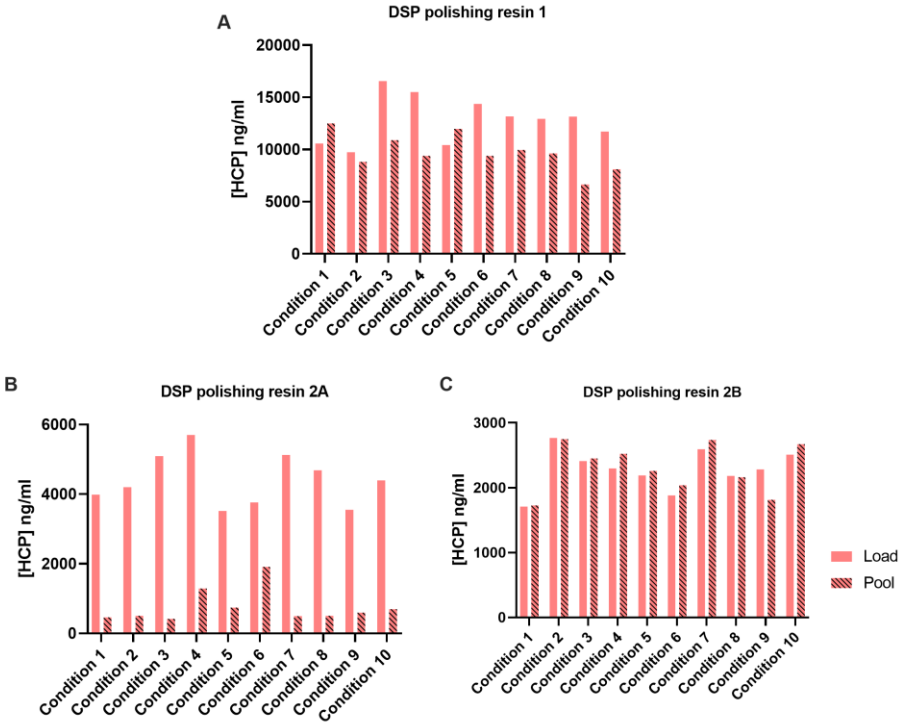
HCP	fmol	ppm
HCP_A	1	65.5
HCP_B	1	37.5
HCP_C	1	44.1
HCP_D	1	50.5
HCP_E	1	22.2
HCP_F	1	54.4
HCP_G	1	28.7
HCP_H	1	48.7
HCP_I	1	14.8



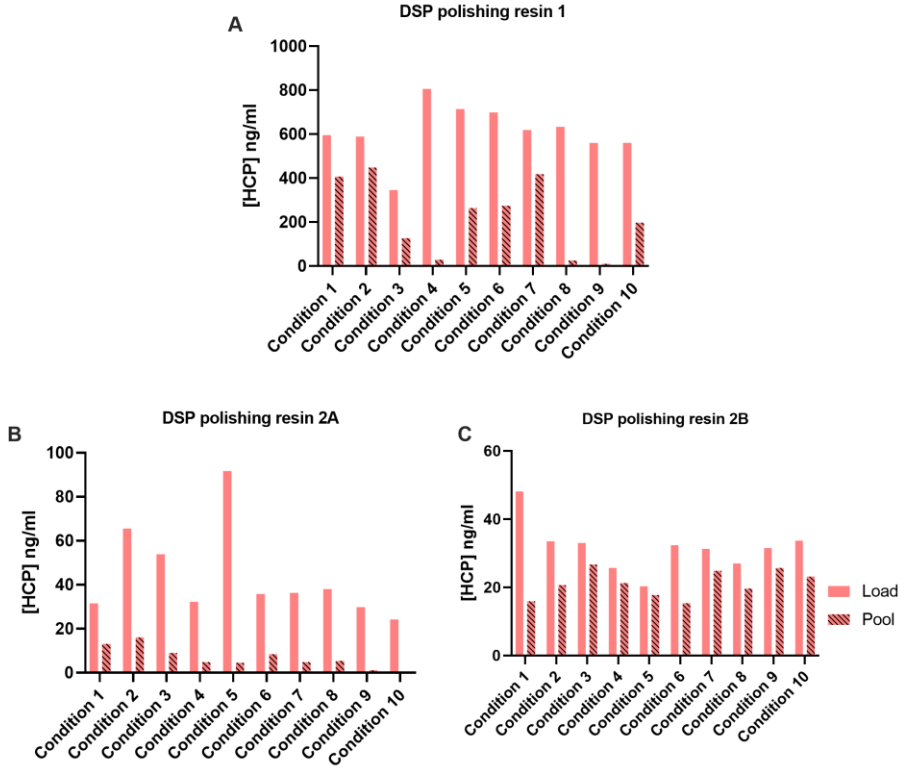
Appendix 7.2 – Measurement of CHO HCPs by ELISA in mAb1 samples for the 10 conditions tested in each of the three DSP polishing resins under evaluation (yy axes were zoomed to the most adequate scale for visualization).



Appendix 7.3 **–** Measurement of CHO HCPs by ELISA in mAb2 samples for the 10 conditions tested in each of the three DSP polishing resins under evaluation (yy axes were zoomed to the most adequate scale for visualization).



Appendix 7.4 – Measurement of CHO HCPs by ELISA in mAb3 samples for the 10 conditions tested in each of the three DSP polishing resins under evaluation (yy axes were zoomed to the most adequate scale for visualization).



Appendix 7.5 **–** Measurement of CHO HCPs by ELISA in mAb4 samples for the 10 conditions tested in each of the three DSP polishing resins under evaluation (yy axes were zoomed to the most adequate scale for visualization).