Development and Application of Charge Variant Profiling Platforms for Molecular Triage of Candidate Monoclonal Antibodies

A thesis submitted to Dublin City University for the degree of Ph.D.

By

Anne Trappe BSc.

School of Biotechnology, Faculty of Science and Health

The research described in this thesis was performed under the supervision of

Dr. Paula Meleady, National Institute for Cellular Biotechnology (NICB), DCU and Dr. Jonathan Bones, National Institute for Bioprocessing Research & Training (NIBRT), Co. Dublin.

September 2018

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____

Anne Trappe

Student ID: 59334831

Date: _____

Acknowledgements

I have been extremely fortunate in my life, to have been given the opportunities to pursue my dreams, for lack of a better word.

In no small way, has Dr. Bones been instrumental in helping me realise this achievement. I am forever grateful to him for taking a risk and sticking with me all these years. Through the ups and downs, he has always had an open door and uncanny ability to reassure and support. He has been a great mentor, advisor and friend and it has been a great privilege to work with him. To Prof. Clynes and Dr. Paula Meleady for supervising my thesis and giving me great help and advice whenever I needed it.

To Mr. Flo, who swooped in like an analytical knight in shining armour, I would like to thank you from the bottom of my black heart. For all of his help and advice in this work and for being a great friend, always having a pirate joke on hand, in times of stress.

To all of the NIBRT pals, for all of your encouragement and especially laughs. It has been a really amazing four years. I have been so lucky to have met you all in a professional capacity but has been an even greater privilege to have you as friends. Particular thanks to Paddy, Jo, Catriona, Roisin and Sarah for being great friends and advisors. To Amy and Josh, my science sister and brother, thank you for writing such amazing theses and for all your guidance on how to get this done. To the CCL members, past and present, who I have learned so much from. Particular thanks to Stef and Silvia who, without their guidance I would have been lost in a sea of glycans. Without the amazing work of Izabela, whit her expertise in all things cells, none of this could have been possible. Thank you so much Iza!

Thank you to Science Foundation Ireland for enabling this research to take place.

Thank you to my friends from other departments, contract research, training and glycolab who have allowed me root around chem stores and hijack equipment.

To the facilities team, Anthony, Stephen, Robbie, Carlos and Barry who have been so generous with their time and help. Sorry for breaking the autoclave so much and for setting so many alarms off at weekends. Thank you for your care.

Thank you to my special friends the salad fingers for so so many laughs over the years and undying support for the smallest achievements. I am so lucky to have had you all in my life, you really have helped make me who I am today.

A special thanks to my Mr. Quinn. I can't really express the gratitude I feel, for having you as a friend, for all of the listening and understanding and great times you've given me I thank you very, very much!

Thanks to my family, Pat, Tessie, Rebecca, Stephen, Donal and the best boys Conor, Jack and Ben. Your encouragement and support has been amazing all these years, I couldn't have done it without knowing you were there every step of the way. To all my aunties, uncles and cousins; SCIENCE, SICENCE, SICENCE! You are the people who know me best, thanks so much for putting up with me all this time! I love you all.

Last, but definitely not least, thanks to my one and tobleroney, Mr Matthew Waters. Thanks for believing in me so thoroughly, that no matter what I could believe in myself. For putting me up there with Valentina and Marie. For your patience and love and letting me have my way. You are the absolute best.

P.s Thanks Nanuq, for being the best dog and for your soulful wisdom.

Publications and Conference Presentations

Publications

- Trappe, A.; Füssl, F.; Carillo, S.; Zaborowska, I.; Meleady, P.; Bones, J., Rapid charge variant analysis of monoclonal antibodies to support lead candidate biopharmaceutical development. *Journal of Chromatography B* **2018,** *1095*, 166-17
- Füssl, F., Trappe, A., Cook, K., Scheffler, K., Bones, J., *Comprehensive Characterisation of the Heterogeneity of Adalimumab via Charge Variant Analysis Hyphenated On-Line to Native High Resolution Orbitrap Mass Spectrometry.* mAbs. Submitted manuscript.

Conference Presentations

- Trappe, A., Mittermayr, S., Millan-Martin, S., Lambert, M., McDonnell, B., Bones, J., Characterisation of commercial drug products and comparison with in-house produced biosimilars.

  Presented as both an oral and poster presentation at Cambridge Healthtech Institute, Peptalk – Protein Science Week, San Diego, January 2017.

- Trappe, A., Mittermayr, S., Millan-Martin, S., Lambert, M., McDonnell, B., Bones, J., Characterisation of commercial drug products and comparison with in-house produced biosimilars.

  Presented as a poster at CASSS - Analytical Technologies Europe, Brussels, March 2017.

- Trappe, A., Mittermayr, S., Millan-Martin, S., Lambert, M., McDonnell, B., Bones, J., Characterisation of commercial drug products and comparison with in-house produced biosimilars.

Presented as a poster at the 16th Human Proteome Organisation World Congress, Dublin, September 2017.

- Trappe, A., Füssl, F., Carillo, S., Zaborowska, I., Meleady, P., Bones, J., Rapid Charge Variant Analysis of Monoclonal Antibodies to Support Lead Candidate Biopharmaceutical Development.

  Presented orally at School of Biotechnology 10th Annual Research Day, Dublin, January 2018.

- Trappe, A., Füssl, F., Carillo, S., Zaborowska, I., Meleady, P., Bones, J., Rapid Charge Variant Analysis of Monoclonal Antibodies to Support Lead Candidate Biopharmaceutical Development.

  Presented as a poster at the 47th International Symposium on High Performance Liquid Phase Separations and Related Techniques (HPLC 2018), Washington, DC, USA, 29th July to 2nd August.

Awards

- Cambridge Healthtech Institute Student Fellowship for attendance and presentation at PepTalk – Protein Science Week, San Diego, CA, USA, January 2018.

- CASSS Travel Reimbursement Grant to facilitate research presentation at CASSS - Analytical Technologies Europe, Brussels, March 2017.

- CASSS Travel Reimbursement Grant to facilitate research presentation at HPLC 2018, Washington, DC, USA, July 29th – August 2nd 2018.

Table of Contents

Abbreviations

2-AA	Anthranilic acid
AAS	Accelerated Aging Study
ABS	Sialidase A
ADCC	Antibody Dependent Cell Mediated Cytotoxicity
AGC	Automatic Gain Control
AMF	α1-3,4 Fucosidase
APR	Aggregation Prone Region
AU	Arbitrary Units
B	Magnetic Sector
BKF	α1-2,4,6 Fucosidase O
Blas	Blasticidin
BPC	Base Peak Chromatogram
BsAbs	Bispecific Antibodies
CDC	Complement Dependent Cytotoxicity
CDR	Complimentary Determining Region
CHO	Chinese Hamster Ovary
CID	Collisional induced Dissociation
cIEF	Capillary Isoelectric Focusing
CpB	Carboxy Peptidase B
CPP	Critical Process Parameters
CQA	Critical Quality Attribute
CV	Charge Variant
CVA	Charge Variant Analysis
Da	Dalton
DC	Direct Current
dd/DDA	Data Dependent
DI	Deionised
DIA	Data independent acquisition
DMSO	Dimethyl sulfoxide
DO	Dissolved Oxygen
DSP	Downstream Processing
DTT	DL-dithiothreitol
EGFR	Epidermal Growth Factor Receptor
ER	Endoplasmic Reticulum
ESA	Electric Sector
ESI	Electrospray Ionisation
eV	Electron volts
Fab	Fragment Antigen Binding
Fc	Fragment Crystallisable
FTICR	Fourier Transform Ion Cyclotron Resonance
FT-OT	Fourier Transform Orbitrap
GFC	Gel Filtration Chromatography
GMP	Good Manufacturing Process
GP	Glycan Profile
GPC	Gel Permeation Chromatography
GUH	β-N-Acetylhexosaminidase

HACA	Human Anti Chimeric Antibodies
HC	Heavy Chain
HCD	High Energy Collisional Dissociation
HCP	Host Cell Proteins
HEK	Human Embryonic Kidney
HIC	Hydrophobic Interaction Chromatography
HILIC	Hydrophilic Interaction Chromatography
HMWH	High Molecular Weight Fragment
IAA	Iodoacetamide
ICH	International Council for Harmonisation
IEC	Ion Exchange Chromatogrpahy
IgG	Immunoglobulin G
IND	Investigational New Drug
IPA	Isopropanol
IT	Injection Time
IT	Ion Trap
K_{app}	Apparent Retention Factor
kV	Kilovolts
LB	Lysogeny Broth
LC	Light chain
LMWF	Low molecular Weight Fragment
LOD	Limit of Detection
LOQ	Limit of Quantitation
mAb	Monoclonal Antibody
MALDI	Matrix Assisted Laser Desorption Ionisation
MAM	Multiple Attribute Monitoring
MCI	Multiply Charge Ions
MES	(N-Morpholino)ethanesulfonic acid
NK	Natural Killer
PAT	Process Analytic Technologies
PBS	Phosphate Buffered Saline
Pc	Peak Capacity
PCR	Polymerase Chain Reaction
ppm	Parts per million
PTM	Post Translational Modification
Q	Quadrupole
QbD	Quality by Design
QE	QExactive
QToF	Quadrupole Time of flight
QTPP	Quality Target Product Profile
RDP	Reference Drug Product
RF	Radio Frequency
RP	Reverse Phase
Rs	Resolution
RSD	Relative Standard Deviation
SCX	Strong Cation Exchange
SEC	Size Exclusion Chromatography

SOB	Super Optimal Broth Medium
SPG	β 1-4 Galactosidase S
SRM	Single Reaction Monitoring
TB	Terrific Broth
TBE	Tris-borate EDTA
TCA	Trichloroacetic acid
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TE	Transfection Enhancer
TIC	Total Ion Chromatogram
UF/DF	Ultrafiltration/Diafiltration
UPLC	Ultra-Performance Liquid Chromatography
USP	Upstream Processing
VH	Variable Heavy
VL	Variable Light
WCX	Weak Cation Exchange
Zeo	Zeocin

List of Figures

2.3.11 Ultrafast 3 minute separations of the trastuzumab sample set consisting of the reference drug product and in-house produced candidate biosimilar material using either transient transfection of Expi293 and ExpiCHO cells and a stably transfected CHO cell line.

3.3.1 (A) Gradient from 0 – 100% buffer B (pH 10.9) in 10 minutes followed by 4 minutes of column flushing and 12 minutes of column equilibration. pH of the eluate was monitored in real time and the trace is shown in red. (B) Two-step re-equilibration with 100 % buffer A (pH 5.3) for 4 minutes followed by 50% buffer B for 12 minutes.

3.3.2 (A) Charge variant separation of seven monoclonal antibodies using a gradient from 0-100% buffer B in 10 min. (B) Charge variant separation after gradient optimisation.

3.3.3 UV-chromatogram highlighting the charge variant separation of adalimumab. Data were acquired using the optimized gradient for adalimumab (Table 3.2.4.1).

3.3.4 (A) BPC's of adalimumab at resolution settings 17,500, 35,000 and 70,000. (B) Mass spectra of the main variant peak. (C) Magnification of charge state +26, with labels of the three most abundant isoforms.

3.3.5 (A) BPC of adalimumab acquired at a resolution setting of 70,000. The main peak in blue corresponds to adalimumab without both C-terminal lysine residues, orange represents the variant containing one C-terminal lysine and green the variant carrying both C-terminal lysine residues. (B) BPC of CpB digested adalimumab. The arrows mark the positions where the lysine variants were present. (C) Relative abundance of the 3 most intense glycoforms of all 3 lysine variants identified.

3.3.6 (A) Magnification of the BPC of an undigested and unstressed adalimumab sample. The peaks labelled from 1 – 3 are the main lysine variants, peaks 4, 5 &6 correspond to the succinimide Asp containing forms. (B) BPC of a CpB digested sample. Peak positions of CpB digested variants are indicated by a red "X". The co-eluting substance, suspected of being a basic variant derived by N-terminal Asp loss is indicated in yellow. (C.) BPC's of samples which were subject to accelerated aging. (D) Bar chart showing the increase of succinimide formation of Asn284 over time.

3.3.7 MS2 spectrum of N-terminal peptide, from the antibody light chain, after aspartic acid loss.

3.3.8 (A) BPC of adalimumab. The presence and retention times of glycated antibody forms is indicated by blue, orange and green labels, in peak fronts of all three major lysine variants. Antibody variants that were found to carry deamidation are indicated in red. (B) Table showing the mass differences between the main glycoform of the three acidic species and the major lysine variant (no C-terminal lysine residues and G0F/G0F glycoform). The average masses were experimentally observed. C.) Ratios of the three most abundant glycoforms, of all three major lysine variants, of the peak centres compared to peak fronts. D.) Asn329 deamidation of the main peak compared to acidic peaks 1 & 2. Mass difference tolerance was set at 5 Da in the deconvolution algorithm

List of Tables

Abstract

Title: Development and application of charge variant profiling platforms for molecular triage of candidate monoclonal antibodies

Candidate: Anne Trappe

Monoclonal antibodies (mAbs) are highly valued biopharmaceuticals. To patients, these molecules are medicines for conditions that have previously been untreatable, improving quality of life and saving lives. To pharmaceutical companies, a successful product can be worth billions. For companies, however, the development of these therapies is a long and expensive process. The patent protected, marketable “life”, of these molecules is relatively short and so companies have only a few years to recoup the cost of development. The cost of these treatments to health services and patients alike can, therefore, be prohibitive. If the development process could be de-risked to identify molecules that will be easily manufactured early in the process, this could lead to significantly reduced costs.

MAbs are complex molecules produced by living organisms. Their production process leads to inherent differences. The charge variant (CV) profile of a mAb is arguably one of the most important critical quality attributes (CQAs) monitored during manufacturing. The CV profile is constituted by mAb isoforms with heterogeneous net charge caused by enzymatic and non-enzymatic processes. Strong cation exchange (SCX) represents the predominant analytical procedure used in this project. The molecular composition of each isoform is usually determined using a multifaceted analytical strategy consisting of mass spectrometry (MS) analysis on many levels, peptide mapping and glycosylation profiling in addition to other methods to evaluate structure and function. By hyphenation of SCX to MS the main isoforms of the mAb can be identified in one analysis rather than a longer and less reliable multi-step process.

To fully understand SCX in depth investigations into the fundamental absorption mechanisms of salt and pH elution modes were performed. From this investigation the optimum parameters were utilised to develop a rapid charge variant method. Finally, to investigate the equivalence of SCX-MS to traditional characterisation studies a comprehensive study of in-house produced Cetuximab was performed. These studies combined to develop a platform for the rapid and in-depth characterisation of candidate molecules by SCX-MS.

Chapter 1.

1. Advances in the characterization of biopharmaceuticals using cation exchange chromatography and high-resolution chromatography.

1.1. Introduction

The Biopharmaceutical industry is rapidly growing, overtaking traditional small molecule pharmaceuticals. Monoclonal antibodies have been, for a number years, the candidate therapy of choice for investigational new drugs. *In silico* modelling and peptide engineering are enabling the development of bispecific antibodies (BsAbs), Fc fusion proteins and antibody-drug conjugates (ADC). BsAbs are antibodies which have two antigen binding sites and therefore have a "two-target" action[1]. Fc fusion proteins consist of the Fc portion of an immunoglobulin fused with a peptide which can have ligand binding functionalities[2]. ADC's have been a particular interest in biopharmaceutical pipelines, as they combine the specificity of an antibody with the potency of chemotherapeutics[3]. These biotherapeutics have revolutionised therapeutic areas such as oncology and rheumatology with expansion into other medical fields such as optometry underway[4]. The process of developing and commercialising a new therapy, however, is still extremely costly.

The expansion of emerging markets has also led to an increasing demand for these products. Equally, there is a growing demand for biosimilars in these markets and this has led to many companies increasing research activities in biosimilar development, as many mAb market leaders come off patent in the next few years[5]. There is lower risk associated with their toxicity, compared to small molecule pharmaceuticals so the approval process can often be shorter and less costly[5]. MAbs are highly specific to the targets they are developed for and their production is easily adaptable with scalable technologies[5].

Chapter 1.

In 1976 Köhler and Milstein first described the fusion of myeloma cells and antibody producing cells[6, 7]. Figure 1.1.1 shows the timeline of antibody advancements from this ground breaking paper.

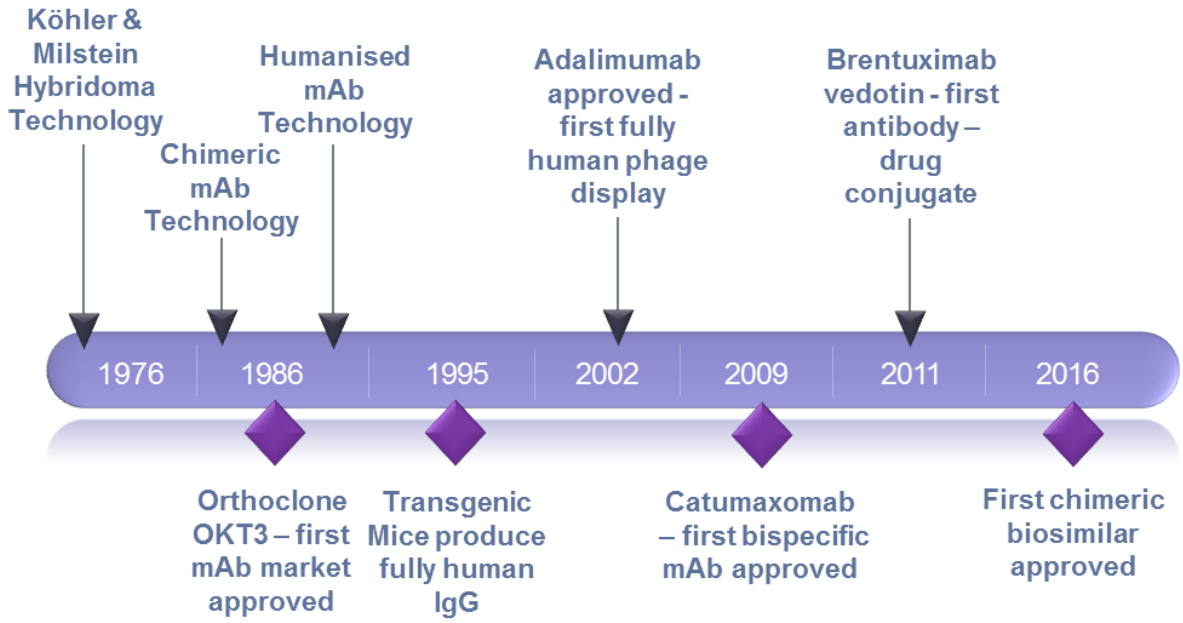


Figure 1.1.1: Timeline of mAb evolution from initial hybridoma technology first biosimilar to market[6, 7]

The development of a biopharmaceutical is a long process, often taking decades to get to market. However, with rapidly advancing knowledge and technology development, time is decreasing.

Emicizumab, owned by Roche, is a humanized bispecific IgG4 mAb for the treatment of haemophilia[8]. It was approved in the US in November 2017 and received recommendation for market approval by the EMA in January 2018[9, 10]. The original patent was given a priority date in 2003[11]. This mAb took 15 years from initial molecule development to market approval. This does not include lead candidate selection which could add 4.5 years[12]. Typically, patent protection covers 20 years beyond the initial approval. Taking the above as an example, Roche have 5 years of market monopoly to re-coup the development costs of this drug. Inevitably, the cost of this drug will be high. The cost per patient per year in the U.S

will be $482,000, however the use of this drug by haemophilic patients is expected to reduce the overall budget by $1.8 billion per patient per year, in the U.S[13]. This is an example of the significant impact mAb therapies can have on the health and comfort of patients and ultimately how their adaptation in therapeutic plans can ultimately reduce costs.

1.2. The Development process – Novel therapeutics

A novel biotherapeutic begins with the selection of an appropriate "druggable" target. This is an antigen to which an antibody can bind and affect a desired response. At this point in the development process the candidate pool is approximately 10^3 molecules. Two methods of generating the initial monoclonal antibody candidates are typically used; transgenic animals and phage display[14]. Rigorous screening eliminates molecules which have poor functional profiles. Target binding assays, cell-based assays and in vivo disease models are some of the screening processes which take place at this stage[15]. Screening reduces the candidate pool to approximately 10-20 molecules. Humanization and re-engineering can then increase that number to roughly 100. It is at this point where more in-depth biophysical characterisation takes place and a molecule's manufacturability is assessed[15].

In recent years there has been an increase in academic involvement in target selection[16]. This has led to the wider availability and transparency of biomedical data. This in turn has led to an increase in the ability to "data mine" using bioinformatics approaches, enabling a high throughput approach to be used in selecting potential target candidates[17].

The phage display method for antibody production allows fully humanised antibodies to be produced. This method negates the need for immunization and harvest of animals. The process begins with a library of genes encoding a selected

antibody's fragment-antigen binding (Fab) region, the part of the antibody which interacts with the target. These genes are inserted into the plasmid DNA of a bacteriophage. The bacteriophage transcribes the gene and displays it on a coating protein on the exterior of the phage[18]. The use of phage display in mAb production also enables greater potential for genetic manipulation and the addition of "manufacturability" characteristics. Built in early, these characteristics can be of great benefit during production scale up and potentially prevent project and product attrition[19].

Focused screening is a knowledge driven method which selectively filters the mAb library for molecules which are statistically more likely to elicit an effect on the target. This type of hit screening is dependent on prior knowledge of the target protein and the nature of therapies which are likely to have an effect on it[17]. Structural based drug design methods such as in-depth target binding assays, pharmacodynamic/pharmacokinetic modelling and preclinical toxicity assessments are used to further filter the hit compounds[17]. Solubility and permeability assessments are important in this stage, ruling out molecules which cannot be readily formulated is key.

Lead optimization is the next stage and the final one before submission to regulatory authorities. The purpose of this stage is to optimise the target molecule profile and ensure that the lead candidate fits this profile[17]. In addition to a focused characterisation of the lead compound, researchers continue to investigate alternative compounds related to the lead. This is continued in case of pre-clinical/clinical failure of the lead compound but also as an investment for future lead compounds[17]. Collectively the data acquired during lead optimization is used to form the submission of an Investigational New Drug (IND) application[17].

Chapter 1.

1.3. The Development Process – Biosimilars

The development of a biosimilar is in some regards a more difficult process. The burden of approval for a novel biotherapeutic rests on proving efficacy and tolerability, but is unrestricted in terms of adherence to a molecular format. For a biosimilar, however, the development process begins with in-depth characterisation studies and retro-engineering a product to suit these constraints[20]. The first step in biosimilar development is obtaining multiple batches of the reference compound, fully characterising and defining the quality target product profile (QTPP) and critical quality attributes (CQAs) for the reference product. From there, the manufacturing process is designed, which results in a suitably similar product. CQA's can be defined, measured and controlled[20]. In developing a novel biotherapeutic the emphasis is on clinical efficacy, in biosimilar development the emphasis is, arguably, on quality.

1.4. Large Scale Manufacture

Production of the mAb candidate at the initial stages of drug development is through transient transfection in a mammalian cell line, once the IND application is successful a stably transfected cell line is produced. The most commonly used cell line for bench scale production is Human Embryonic Kidney (HEK 293). These cells are suited for this scale as they readily accept the plasmid DNA and can produce large quantities of protein over a short period of time[21].

CHO is the mammalian cell line of choice for production of commercialised mAbs as it is readily adaptable to large scale cultures, is more malleable to stable transfection than HEK and there are decades of research into the optimisation of CHO culture processes. CHO cells also have the ability to produce human-like glycosylation patterns, all these factors contribute to the confidence regulatory agencies have in the safety of therapies produced in CHO, which can lead to faster

approval times[14]. It is well understood that there are inherent differences in the glycosylation patterns of mammalian cell lines[21, 22]. However, the differences between cell line expression of other PTM's are not frequently addressed. There is potential for, positive or negative, modifications occurring when a well characterised molecule is produced in HEK and then transferred to CHO.

Once the stable cell line has been generated, process development begins. Figure 1.4.1 illustrates the manufacturing process, from upstream processing through downstream processing to fill/finish. The principles of Good Manufacturing Practice (GMP) must be built into the process at this stage. Information gained from bench scale production, literature and experience are used to develop the optimum process parameters for the mAb. At this point the CQA of the product are established, concurrent with this the critical process parameters (CPP) and a risk evaluation takes place. A benefit of using CHO in large scale production is the numerous studies which have taken place on characterising the process of producing recombinant proteins in CHO. One such study by Rouiller *et al.* (2012) describes a quality by design (QbD) approach to the production of an Fc fusion protein in CHO. They identified the main process parameters which affected the quality of their product with a particular focus on upstream processing (USP)[23]. They identified; culture pH, dissolved oxygen (DO) and culture duration as being the cell culture process parameters which had the largest effect on product quality[23]. Fortunately, these are process parameters which can be carefully monitored and controlled. In Downstream Processing (DSP) however some of these parameters are inherent parts of the process; mechanical stresses, pH fluctuations and high concentrations are all unavoidable critical process parameters.

From harvest to fill/finish the mAb protein is exposed to mechanical stresses through centrifugation, filtering, pumping and stirring. Before complete clearance of

process impurities, it is exposed to host cell proteins (HCP), which include proteases, and dissolved oxygen in the culture media which can induce oxidation of amino acids, such as methionine and tryptophan. After these gross impurities are removed the mAb is exposed to pH fluctuations as part of Protein A purification and high ionic strength buffers in the polishing chromatography steps. The mAb also experiences temperature fluctuations as part of the process including freezing before final fill. The protein interacts with steel, glass and plastic through DSP and the effect which extractables and leachables, from these materials, have on the protein structure is only beginning to be understood[24]. In effect, the entire process exposes the mAb to stresses which may alter its native state.

MAbs have a very specific binding affinity to Protein A; hence it is used in almost all industrial purification processes of mAbs and Fc fusion proteins. However, due to this high affinity, a low pH buffer is required to elute the protein. At this acidic pH, aggregation readily occurs. Aggregation levels are one of the most important CQAs in mAb processing. Aggregation can cause viscosity problems during formulation and more importantly adverse immunogenic responses in patients following administration. Unfortunately, aggregation is an unavoidable phenomenon in production; interestingly low pH may not be the sole cause of aggregation during Protein A purification. Ejima *et al.* (2005) demonstrated that; using the same pH but different buffer formulations had an observable effect on aggregate formation. They found that, protein A elution with a 2 M arginine solution resulted in higher recovery titres and lower aggregate formation, compared to other buffers[25]. Hence aggregate formation could be lessened by buffer choice informed by quality target product profiling (QTPP)[26].

Ultrafiltration/diafiltration (UF/DF) is the high pressure filtration of sample through molecular weight cut off membranes that allow only molecules of the correct size

through[27]. It is used in mAb processing for concentration and final formulation buffer exchange[28]. During UF/DF the mAb is being pumped under high pressure through the membrane which causes mechanical stress on the protein and ultimately can lead to aggregates. At this stage of DSP the final formulation solution is employed to protect the product. A formulation which can stabilize the native structure of the mAb while it is under these mechanical stresses is essential[28]. The surface of the membrane in UF/DF is a location of high concentration of the product. A localised high concentration can result in a decrease of membrane flux which in turn leads to the exposure of the product to air-water interfaces which causes aggregation and ultimately further membrane fouling[26]. A formulation designed to reduce solution viscosity can help reduce the occurrence of membrane fouling. Excipients such as the addition of arginine have been shown to reduce viscosity as well as increasing the ionic strength of the buffer[28].

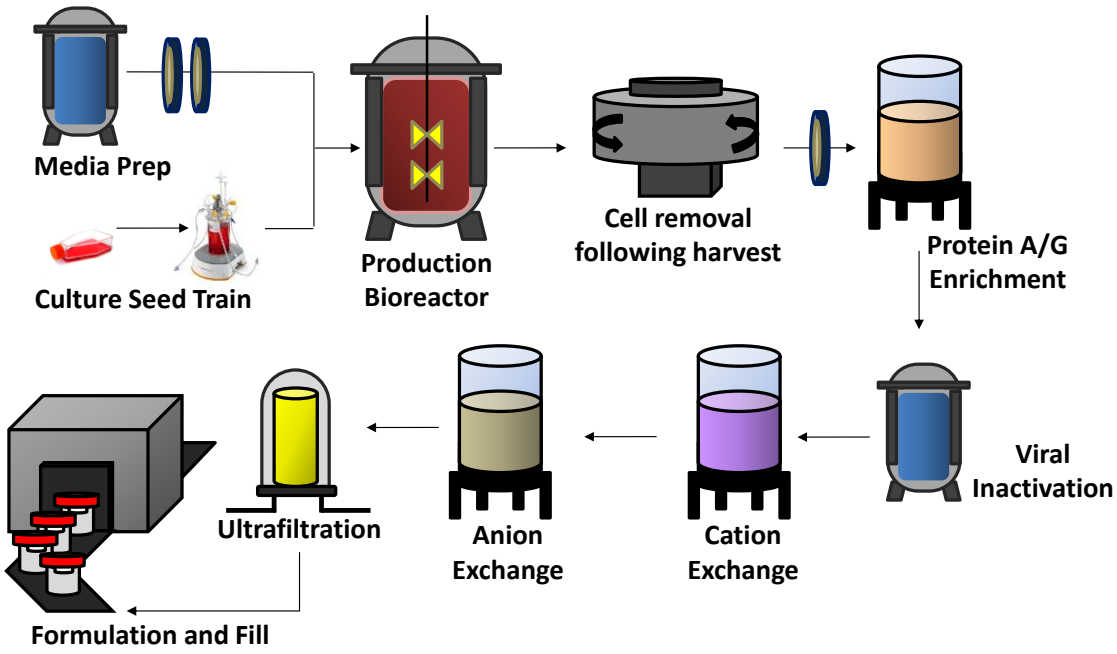


Figure 1.4.1: Schematic of upstream and downstream processing of a mAb

The design of formulation buffer requires detailed knowledge of the products CQA's. The protein's conformational stability and colloidal stability are related to

the proteins propensity for aggregation and are used to establish the required excipients for formulation[29]. Conformational stability can be thermodynamically described as the difference in free energy between an antibody's native and denatured state given by the formula (ΔG_{ND})[29]. The larger the value for ΔG_{ND}, the more stable the protein is. A mAb's colloidal stability is dependent on the protein's interaction either with other proteins or with the solution. The osmotic second virial coefficient incorporates these physical characteristics and is used as a score of stability in solution, $B_{22} = \frac{2\pi}{M^2}\int_0^\infty r^2\left(1-e^{-u(r)/kT}\right)dr$,[30] where; *M* is the molecular weight of the protein, *r* intermolecular separation distance, *k* is the Boltzmann constant, *T* is absolute temperature and *u(r)* interaction potential. The value *u(r)* describes all interaction forces between two proteins including hard-sphere, electrostatic van der Waals and short range interactions. The B_{22} score encompasses all potential protein-protein interactions and protein-solution interactions, a thorough break-down of the equation is discussed by Chi *et al* (2003)[30].

In short, a large and positive B_{22} value implies high colloidal stability whereas a negative value implies low colloidal stability and a propensity for aggregation. These two scores are used together to choose excipients tailored for the protein. For example; in a situation where a protein has a large ΔG_{ND} and a negative B_{22} value, increasing ΔG_{ND} leads to a more stable protein. This increase has been shown to be achieved by the addition of a sugar, such as sucrose to the formulation[30]. The sugar, changes the protein structure to a less solvent exposed configuration. In the case where a protein has a large and positive B_{22} value, the protein-protein interaction in solution is repulsive. Thus, if a protein becomes denatured in this solution it will have a lesser propensity for aggregation due to

these repulsive forces, formulation characteristics for this condition would favour conservation of the native structure[30]. Chi *et al.* (2003) found, in this instance, that the best formulation for the protein being studied was a low pH and low ionic strength solution. .However, it is necessary to have a complete knowledge of a mAbs QTPP, which should include ΔG_{ND} and B_{22}, to configure a formulation which is best suited[31]. In instances where there is no optimum formulation, targeted mutation of amino acids is then performed. Kuhn and colleagues (2017) described the molecular dynamics simulation driven, mutation of certain amino acids in the Fv regions of two mAbs to charged amino acids which contributed to a more evenly distributed charge profile[32]. The resulting mAbs showed increased value for ΔG_{ND} and a decrease in aggregate quantity, overall[32].

After fill/finish mAbs are stored for long durations at low temperatures. Under these stresses, aggregation is common and often unpredictable. A key excipient consideration is the addition of cryopreservants which can protect the mAb during long-term freezing conditions, which could be as long as 5 years at -20°C[28]. These excipients, such as sugars, must be resistant to crystallization during freezing, sugars are such an excipient. Shire *et al*.(2009) stated that disaccharides such as sucrose are effective cryoprotectants, however sucrose, at low pH, can hydrolyse to form glucose which has potential to produce non-enzymatic glycation on amino acids and thus contribute to protein instability[28]. Polysorbates are plant derived emulsifiers which are used in antibody formulation to protect the protein stability by preventing the exposure of the antibody's hydrophobic regions to the air-surface interface[29]. Polysorbates, however, can be degraded by near-UV radiation and host cell proteins to produce free fatty acids, which can have negative effects in regards to the alteration of primary amino acids[29]. Putative phospholipase B-like 2 (PLBL2) is a host cell protein, expressed by CHO cells under extended culturing. PLBL2 has

been reported to degrade polysorbates in formulation buffers due to their structural similarity to triglycerides[33]. If the mAb is preserved by freeze drying, lyophilisation or spray-drying, excipients for these methods must also be considered. The addition of disaccharides to the formulation buffer helps replace hydrogen lost during drying and contributes to maintenance of stability in this form[34].

Chapter 1.

1.5. Post Translational Modifications

Post translation modifications (PTMs) are enzymatic or non-enzymatic alterations to a protein, which occur within the cell, following excretion into media or during processing. For the purposes of this review they are discussed under two main categories; amino acid modifications and glycosylation.

1.5.1. Amino Acid modification

Understanding the post translational modifications, which affect protein function, efficacy and potency, adds value to formulation development and stability assessment of a mAb. The following section explores the different modifications used to establish a product's CQAs, as reported in literature. Table 1.5.1 summarises these modifications.

Methionine (Met or M) is a hydrophobic amino acid. The thioether functional group (R-S-R') has a low oxidation potential, therefore it is easily oxidized when the Met residue is surface exposed. Methionine sulphoxide (Met-SO) is typically formed through the two-electron transfer of oxygen to the thioether[35]. Met-SO is a stable compound and usually will not be further oxidised, however, Met-SO_2 may also occur. Met oxidation has been shown to disrupt the binding of an IgG to the neonatal Fc receptor, FcRN. FcRN is responsible for the uptake of IgGs into the bloodstream, and also for the IgGs characteristic long half-life in serum[36].

Chapter 1.

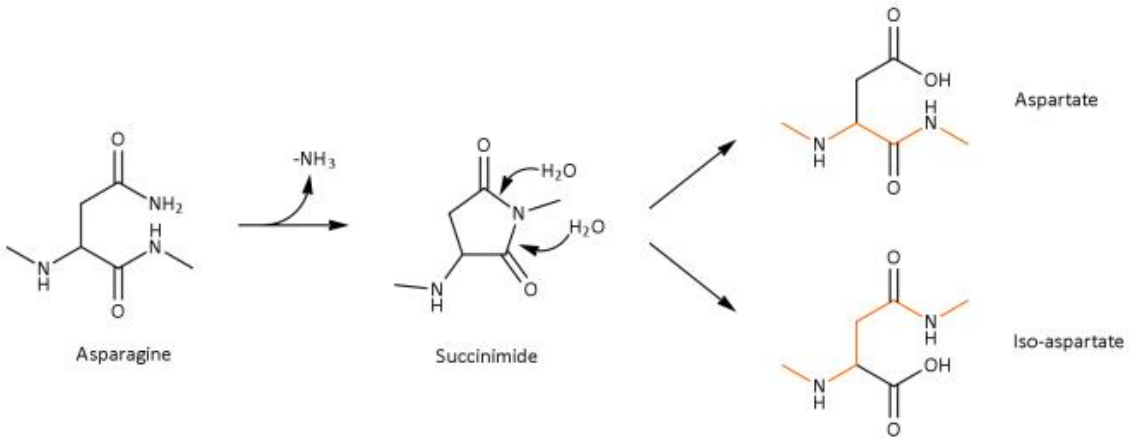


Figure 1.5.1: Deamidation mechanism of Asn to Asp/IsoAsp

The deamidation of asparagine to aspartic acid is initiated in a neutral/basic environment when the peptide bond NH becomes deprotonated, fig 1.5.1. This induces a nucleophilic attack on the NH_3^+ by the carbonyl group, with the loss of ammonia and the formation of the succinimide intermediate. This unstable group then undergoes hydrolysis to form a mixture of Asp and IsoAsp. The ratio of these is usually in favour of Asp, as Asp is better accomodated within the protein structure[37]. When Asn is followed by a Glycine (Gly) amino acid on the C-terminal side, it is more likely to become deaminated. Serine has also been implicated in Asn deamidation[38].

One of the most common methods for analysing deamidation in a protein is the use of isotopic labeling in conjunction with HPLC and MS. Samples are prepared in $H_2{}^{18}O$. The heavy oxygen is taken up in the hydrolysis of succinimide. The molecular weight increase from Asn to Asp is 3 Da with H_2O^{18} incorporated, which allows the Asp or IsoAsp to be easily distinguished with MS[39, 40]. However, ^{18}O labeling can result in complicated spectra due to the uptake of O^{18} onto the C-terminal carbonyl[40].

A glutamic acid at the N-terminal is common to half of all reported anitbody's heavy chains, while 6% of κ light chains have an N-terminal Glu[41]. The formation of

PyroGlu from Glu involves the loss of a water molecule. While the formation of PyroGlu from Gln involves a deamidation reaction, figure 1.5.1.3 shows these. These reactions can occur enzymatically and non-enzymatically. PyroGlu formation from Glu can be identified with MS by the loss of 18 Da[41]. While the formation pyroGlu from Gln is characterised by a -17 Da shift, the lack of a free amino group in Pyro-Glu prevents its own removal by the enzyme and the removal of subsequent amino acids from the N-terminus.

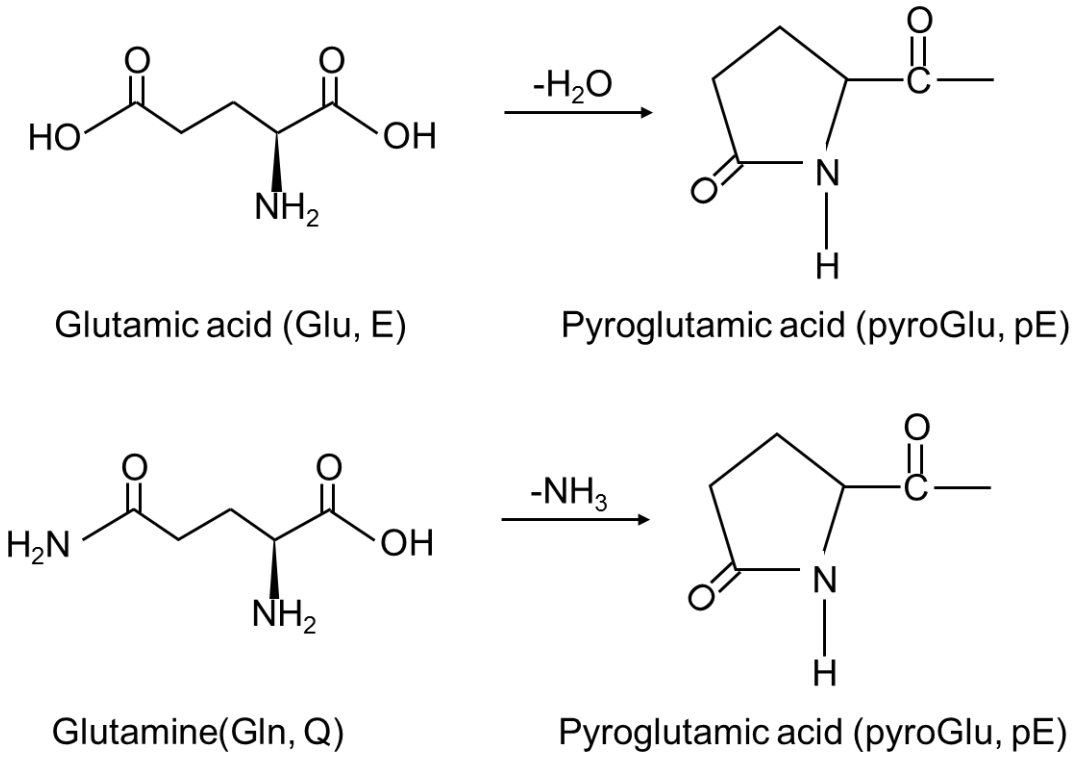


Figure 1.5.2: Dehydration of Glu to pyroGlu (Top). Deamidation of Gln to pyroGlu (bottom).

C-terminal lysine clipping is reported to occur *in vivo* within 1 hour, post injection[42]. Efficacy remains unchanged as the heavy chain is not involved with antigen binding. The cleavage of Lysine has, however been implicated in the disruption of the charge within an antibody and contributes to lot-to-lot variation of antibodies. Sundaram *et al.* (2011) investigated the cause of an observed variation in a certain mAb's pI from batch to batch. It was found that in one lot, there were seven isotypes of one mAb with a pI range of 7.9-8.9. After treatment with carboxypeptidase and

a sialydase there were only 3 isotypes remaining. This further suggests that the adherence of lysine to the C-terminal is not necessarily a natural and beneficial phenomenon, especially when considering product conformity[43].

Table 1.5.1: Summary of amino acid modifications, effect they have on LC elution and the structural and functional impact they have[41, 44, 45].

Modification	Formation	Elution in LC modes	Structural	Function
Met/Trp Oxidation	Reaction of AA with oxidative species	Elutes prior to the main peak in RP and HIC, but can be distributed across variants in SCX	Contributes to protein unfolding	FcRN binding is disrupted which reduces the serum half-life of the IgG
Asp deamidation	Loss of NH_3 formation of succinimide intermediate	Elutes prior to and post of the main peak in RP and HIC, elutes in basic variants in SCX	Contributes to chemical degradation	No significant functional effect
N-terminal PyroGlu formation	Loss of H_2O	Elutes after the main peak in RP, HIC and SCX	Infers protection from peptidases	No reported effect on function or efficacy
C-Terminal Lys clipping	Carboxypeptidase cleaves C-terminal Lys	Elutes prior to the main peak in RP and HIC, main contributor to basic variants in SCX	Contributes to charge heterogeneity	No effect

1.5.2. Glycosylation

On a whole organism level, glycosylation has many roles in the correct function of biological systems including blood grouping and the phenotype of genetic disorders[46]. Glycosylation is a complex process and one of the most important CQA's of a biopharmaceutical. Glycosylation can have a positive impact; ensuring correct folding of the protein, increasing serum half-life and eliciting ADCC activity.

However, when it has a negative impact, such as non-human glycoforms on therapies, it can have dangerous consequences resulting in adverse immune responses[47].

The synthesis of a glycopeptide involves three major steps; (1) the dolichol cycle and oligosaccharide transfer to the nascent protein, (2) trimming of the oligosaccharide and (3) branching[48]. Steps (1) & (2) take place in endoplasmic reticulum (ER) and step (3) in the Golgi apparatus. Steps (2) and (3) are illustrated in figure 1.5.3, below. The formative structure of an oligosaccharide is a dolichyl phosphate, which sugars are added to until a conformation of $GlcNAc_2Man_5$ is attained. At this point the molecule is recognised by flippase co-factors and the oligosaccharide portion of the molecule is transferred to the lumen side of the ER[48]. In the outer part of the ER, four additional mannose and three glucose residues are added to the molecule. Oligosaccharyltransferase transfers the oligosaccharide portion of the molecule to an accepting asparagine residue of the nascent polypeptide chain. Prior to addition to the polypeptide, the asparagine must be in the consensus sequence Asn-X-Thr or Asn-X-Ser, where X can be any amino acid except Pro. It must also be located on the solvent accessible portion of the 3D structure and must be orientated toward the luminal part of the ER[46, 48].

Once the oligosaccharide is attached to the polypeptide, steps (2) & (3) proceed. Trimming of the oligosaccharide begins with removal of three glucose residues and proceeds with removal of the terminal mannose residues. Once the folded protein has been transferred to the Golgi apparatus further trimming and branching occurs. Branching involves the addition of more complex sugar moieties than occurred in the initial oligosaccharide formation.

Chapter 1.

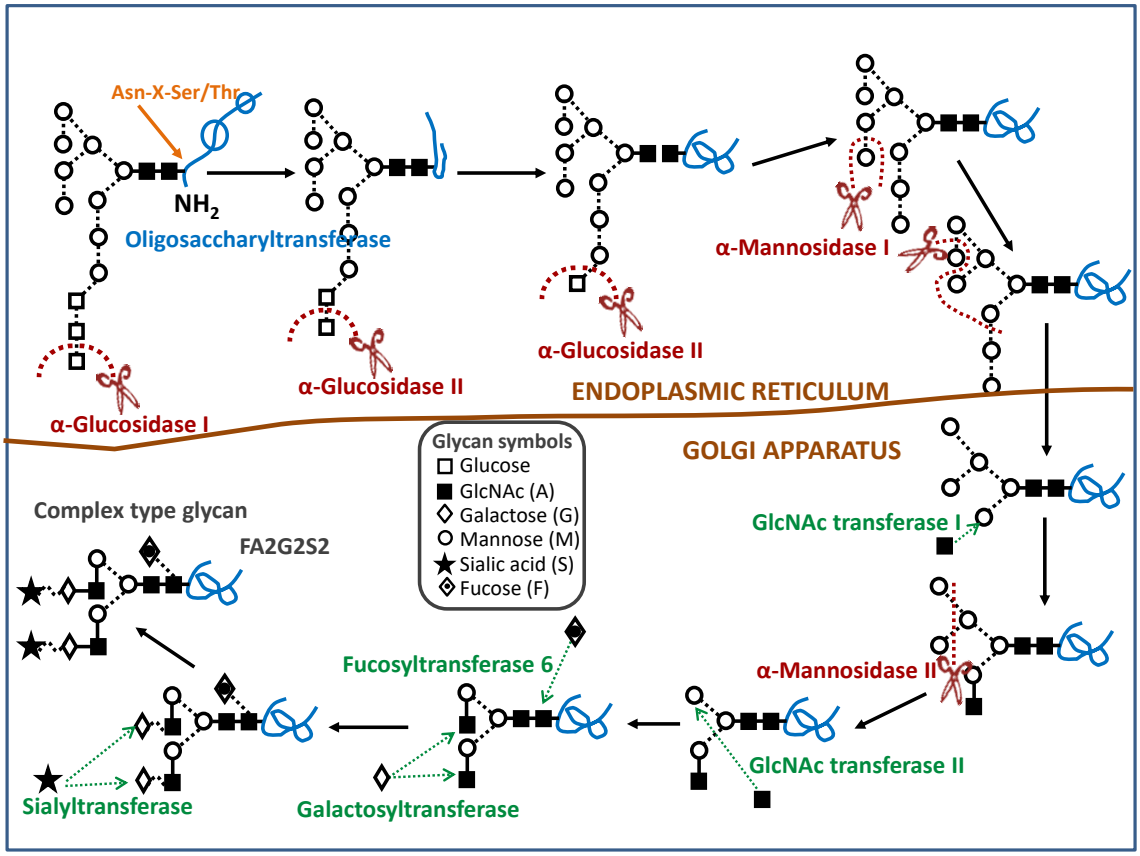


Figure 1.5.3: The synthesis of a glycopeptide from oligosaccharide trimming in the ER to complex glycoform finishing in the outer Golgi.

There are three main classifications of glycan moieties. Figure 1.5.4 illustrates these.

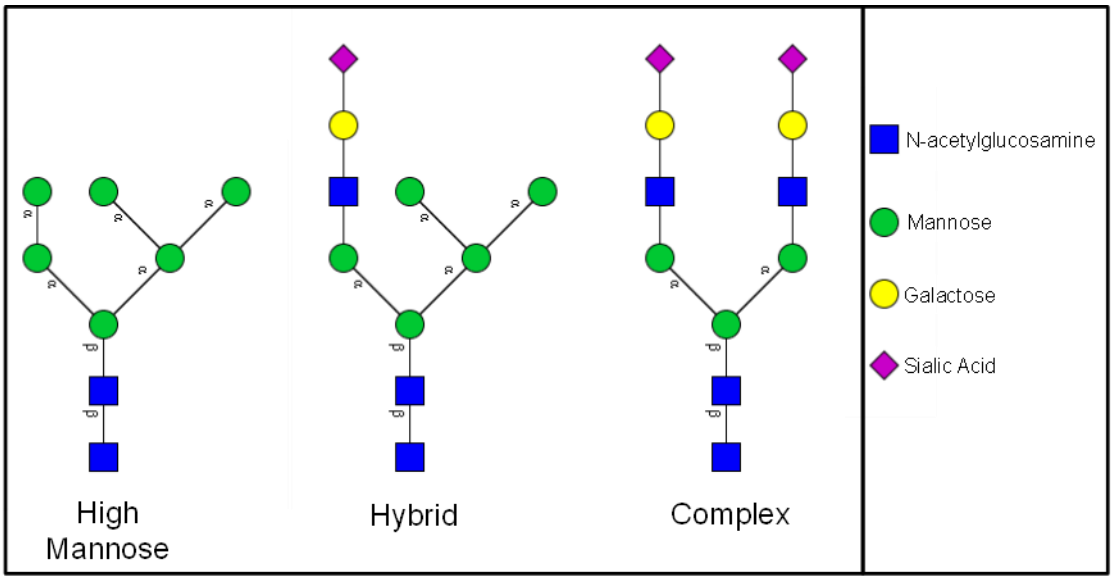


Figure 1.5.4: Three main glycan archetypes; high mannose, hybrid and complex. Right of the image is a legend key.

Chapter 1.

Glycosylation has been shown to have a significant effect on antibody function and also for the efficacy and safety of protein therapeutics. Glycosylation has influenced the shift from simple cell factories such as yeast and bacteria to mammalian cells such as CHO and HEK in an effort to replicate the complex mammalian glycosylation activity that occurs *in vivo*[48]. The most important site of glycosylation is on the fragment crystallisable (Fc) of IgG antibodies, on asparagine residue 297 (Asn297). There is a large amount of heterogeneity beyond the core pentasaccharide; however, tight process controls can ensure a relative amount of predictability of a products glycan profile. The Fc region contains effector ligands for various receptors which enable the Ab to trigger responses such as antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC). The type of glycan addition can both positively and negatively affect the antibody's elicitation of these immune responses, Table 1.5.2 summarises these. The absence of core fucose has been shown to improve the binding affinity of Human IgG to the FcγRIII receptor and thus, increases ADCC activity, as FcγRIIIa is one of the main receptors to elicit the ADCC response[49, 50]. Galactosylation has an effect on the binding of the C1q complex of an IgG to initiate the complement system. Hodoniczky *et al.* showed that, while terminal Galactosylation is necessary for binding of C1q, excess gal can inhibit binding[51]. Sunadaram *et al*. (2011) illustrated that, along with C-terminal lysine clipping, increased sialylation also contributed to charge distribution variation of different batches of the same mAb[43].

Table 1.5.2: Summary of N-glycan sugars[43, 49-56]

PTM	Location on core heptasaccharide	Structural effects	Functional effects
Fucose	Commonly on the core GlcNac residue, also found on outer residues	Absence has an effect on the antibodies binding ability to a receptor	Absence increases binding affinity to FcγRIIIa, increases ADCC activity
Galactose	Either α (1-6) or α (1-3) of the terminal GlcNAc residues. Can also be bisected.	Excess Gal can disrupt binding to receptor	α (1-3) is not a human configuration and can cause immunogenic response to the antibody
Mannose	$GlcNAc_2Man_5$ is most common isoform where terminal GlcNAc are replaced with Mannose	Receptor binding disruption	Interferes with mannose binding receptor, reduces serum half life
Sialic Acid	Often found attached to Gal residues	Negatively charged Sialic Acid increases electrostatic repulsion on the mAb	Interferes with FcγRIIIa binding, reduces ADCC activity
N-Acetylglucosamine (GlcNac)	Key constituent to the core heptasaccharide	Main building block for further glycosylation	Bisection GlcNac induces a stronger ADCC activity, particularly in combination with the absence of core fucosylation

Chapter 1.

1.6. Analytical Techniques for Biopharmaceutical characterisation

Regulatory agencies and the International Council for Harmonisation (ICH) encourage in-depth product characterisation as part of the development and manufacturing process. The more information which can be gathered about a biopharmaceuticals CQA's and the effects processing and other conceivable stresses can have on the quality of product, aids in regulatory approval and also ensures consistent product quality. Table 1.6.1 summarises the attributes of biopharmaceutical which are used as an indicator of manufacturability.

Table 1.6.1: Structural and sequence characteristics commonly used as indicators of manufacturability in the selection/optimisation phase[15, 57, 58]

Attribute	Role in developability	Method
Isoelectric Point (pI)	Designing downstream processing and informing formulation studies	Sequence based calculation, SCX
Charge Distribution	Even distribution of charge is a desirable trait. Decreases risk of aggregation	SCX
Aggregation prone regions (APR)	Highly hydrophobic / hydrophilic molecules have tendency to aggregate	HIC/HILIC
Amino Acid positions for PTMs in the CDR	Unpaired cysteines can lead to disulphide scrambling and potentially encourage aggregation. Consensus sequences for N-glycosylation to asparagine should be identified.	Peptide Mapping by UHPLC-MS
Protease recognition sites	A molecule with a high number of protease recognition sites has a higher risk of degradation during culturing	In vitro binding assays

1.6.1. Liquid Chromatography

Liquid chromatography (LC) is an analytical technique which separates a complex mixture into its constituent parts by the analytes interaction, either chemical or physical, with a mobile phase and a stationary phase. The analyte partitions

between a moving solvent, the mobile phase, and a solid particulate, the stationary phase. The stationary phase is typically supported by a column casing made of plastic or steel in which the mobile phase and sample flow through. Ultra-performance liquid chromatography (UPLC) is characterised by having high selectivity and resolution and is afforded its name due a combination of technological advances in column technology, high pressure pumps and increasingly sensitive detectors[59]. Smaller particle sizes, narrower internal diameters, more stable phases and increased knowledge of column chemistries have revolutionised the separation of biomolecules[60]. These factors, and more, influence the theoretical plate number of a separation and have led to increasing resolution and peak capacity[61]. The theoretical plate number is an abstract representation of the number of separations which can occur on column.

LC can be used for many applications. In the biopharmaceutical arena, it used as a sample purification tool, for process control and for hyphenation to complex detectors to yield in-depth information. In this body of work, LC is used to two ends; as a robust screening tool for quality and preliminary information and then for in-depth analysis of molecules both on-line and off line to mass spectrometry. Figure 1.6.1, shows a basic schematic of an LC system. Solvent is pumped from the reservoir through the injector, where sample is collected and flowed through the column, the sample interacts/elutes and travels through a detector where the signal is transmitted to a data processor. This set-up is universal for LC systems. The specificity of LC is gained from the column type and detector[59].

Chapter 1.

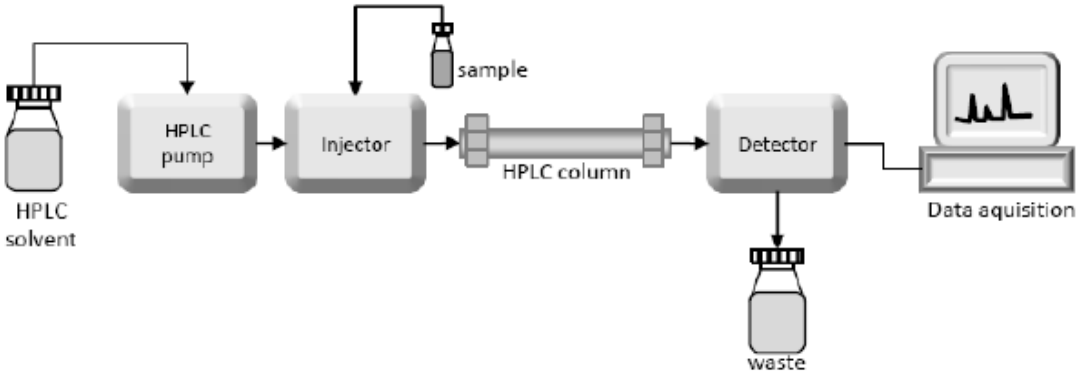


Figure 1.6.1: Basic schematic of an LC

In this work, UHPLC for biomolecules was employed to characterise large monoclonal antibodies. The LC modes utilised in this body of work are discussed in detail below.

1.6.1.1. Ion Exchange Chromatography (IEC)

IEC is one of the most versatile UHPLC methods used in the biopharmaceutical industry. It is used as a purification step as well as a technique for monitoring consistent product quality. As it is a non-denaturing method it enables the analyst to study the complex form of the protein in its native state[62].

Ion exchange chromatography works on the principal that a molecule with a charge will be attracted to a stationary phase with an opposing charge and will be displaced from the column when the mobile phase has a pKa value equal to the net charge on the molecule, it is then that the molecule exhibits no net charge and elutes when using pH gradient elution, or may be displaced by competitive interaction with a counter ion when using salt gradient elution[63]. Isoforms of the protein having differing pIs can be separated using IEC. Anion exchange and cation exchange are the ion exchange phases used in IEC. Anion exchange chromatography is best suited for samples which have a predominantly negative charge, whereas the opposite is applicable for cation exchange. Figure 1.6.2 shows a break-through graph which represents the shift across pI where proteins bind to IEX resins.

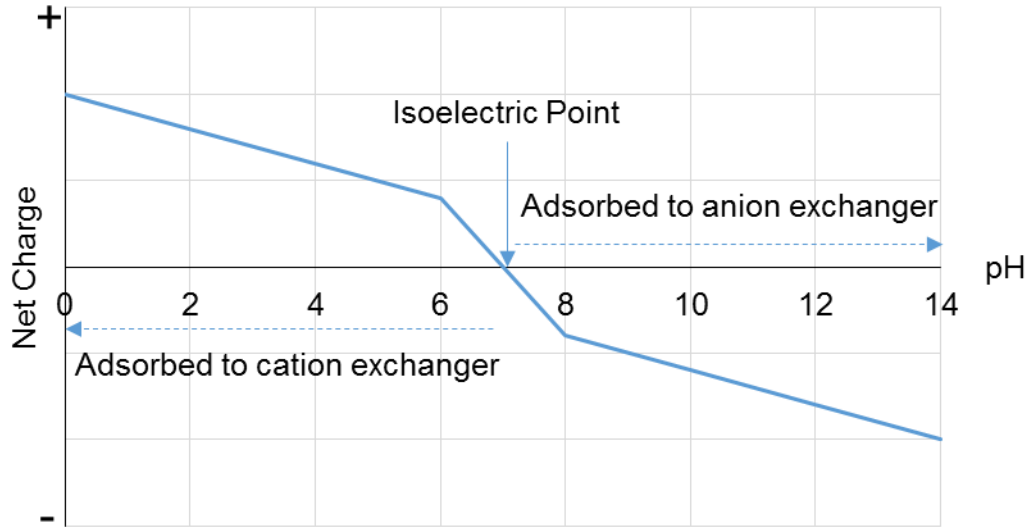


Figure 1.6.2: Graphical representation of adsorption characteristics depending on pH and overall charge

Table 1.6.2 below shows the most commonly used ion exchangers in IEC stationary phases. Zwitterionic exchangers are useful for complex sample mixtures as they enable the analyst to separate a variety of analytes with differing pIs. Weak to medium exchangers have the ability to be "switched on" by the ionic strength/pH of the mobile phase and so require longer equilibration times. Strong exchangers on the other hand buffer the changing on-column conditions and so can be operated at extreme ends of the pH/ionic spectrum.

Table 1.6.2: Ion exchange resin functional groups

Ionic mode	Functional Groups
Weak Cation	COO^-, C_6H_4-O^-, AsO_3H^-
Medium Cation	PO_3H^-
Strong Cation	PO_4H^-, SO_3^-
Weak Anion	NH_3^+, $[NH_2(CH_3)]^+$
Medium Anion	$[N(CH_3)_2(CH_2CH_2OH)]^+$
Strong Anion	$[N(CH_3)_3]^+$, $[N(C_2H_5)(CH_3)_2]^+$
Zwitterionic	$N(CH_3)_2^+$-$(CH_2)_n$-SO_3^-, $CH(SO_3^-)$-$(CH_2)_n$-$N(CH_3)_3^+$

1.6.1.1.1. Strong Cation Exchange

Strong cation exchange chromatography (SCX) is considered a gold standard in protein characterisation[64]. Due to the overall basic nature of mAbs they exhibit an overall anionic charge with pI's typically ranging from 6.0 – 9.5 and so are best suited to cation exchange[64]. The main contributors to charge heterogeneity on mAbs are; C-terminal lysine clipping, glycosylation, deamidation, N-terminal pyroglutamate formation, and methionine oxidation[65].

There are two elution modes for SCX; salt gradient elution and pH gradient elution, both of which are discussed, in detail, in Chapter two. pH gradient elution is the elution method employed throughout this body of work. pH gradient has been found to be superior to salt gradient in terms of separation efficiency, robustness and universality[62, 66]. In pH gradient, the ionic strength of the mobile phase is constant while the pH of the phase changes over gradient time. There are two methods of applying a pH gradient; internally and externally. In an external pH gradient two buffers are used which are mixed *via* the LC pump to create a gradient, before the column[67]. In an internal pH gradient, often referred to column chromatofocusing, the column is equilibrated to a high pH, prior to injection, and step wise progresses to a low pH, creating a "focusing" of analytes along the column in each pH band[68]. Hyphenation to mass spectrometry has, until recently, been difficult to achieve. Due to the relatively high ionic strength of elution buffers. Determination of charge variant identities typically required a 2D LC method, IEC followed by RP for buffer exchange.

However, Fussl *et al.* (2018) developed a buffer system which is suitably volatile and enabled identification of the major charge isoforms, in their native state, in one step[69]. This was achieved by using buffer constituents such as ammonium hydroxide and ammonium bicarbonate which were able to achieve a suitable pH gradient while also allowing hyphenation to MS. This buffer system required a

convex curved gradient be applied to achieve adequate separation of mAbs which had pIs outside of the pH range of the buffer (5.3 – 10.2). This method is employed in this body of work to analyse the complex charge variant patterns of in-house produced mAb variants. The pH range of the buffer system was broadened to counteract the difficulties encountered previously. The formation of carbon dioxide adducts was also a problem which was addressed in Chapter 3, of this work.

1.6.1.2. Size Exclusion Chromatography (SEC)

SEC is the separation of molecules based on size, where, higher molecular weight substances are excluded by the small particulates of the stationary phase and so move to the void and elute. Smaller molecules enter the tightly packed stationary phase and move through the column at a slower rate[70]. The column packing material is typically inert and so this is a physical separation method, utilising isocratic elution.

SEC began as gel filtration chromatography (GFC) or gel permeation chromatography (GPC) depending on the polarity of the mobile phase. Commercialised by Waters Corporation in the 1960's, it enabled polymer chemists to observe the molecular weight of large polymers[59]. However, modern SEC columns are not strictly gel based but are often porous silica or monolithic and can be used with both aqueous and organic solvents[59]. The chemistry of the column is less significant in SEC than in other LC modes, however intra particle pore diameter and particle diameter are arguably more important[63]. Liu *et al.* (2009) demonstrated this, in a study on optimisation of SEC for MS analysis. They showed the effects particle size had on resolution by using the same separation method on five columns, each with the same length and pore size. The column with the smallest particle size had the highest resolution value[71]. However, it can be said that this is true for all HPLC columns, however SEC separation is a physical separation and it is interesting to note, in this case, that columns with equal dimensions and

chemistries have significantly different resolution scores based on particle size alone.

SEC can be applied to many aspects of biopharmaceutical production, similarly to IEC; it is used as a purification method prior to polishing, as a tool for rapid determination of physical characteristics and hyphenated to MS for in-depth analysis of size variants and the potential identification of aggregation prone regions. One of the primary uses is for the separation and subsequent analysis of aggregation in mAbs [72, 73]. Mono-molecular size, aggregation or fragmentation is determined by the generation of a standard curve using protein size standards, prediction error of molar-mass accuracy of roughly 12% has been reported[63]. During formulation development, SEC is a particularly useful tool, incubating the mAb in various candidate buffers and analysing both the native conformation while simultaneously characterising aggregated forms, enables developers to confidently select the optimum formulation[74]. SEC is also used for the confirmation of correct conjugate binding during antibody-drug conjugate engineering[75].

SEC-MS can often be a difficult process, non-denaturing phosphate buffers are widely used in SEC and can cause ion suppression when used with MS. Use of other, more MS-friendly organic based denaturing buffers, can result in the loss of some structural features which are indicative of aggregation propensity[76]. Indirect coupling can be performed by first dimensional separation and fraction collection by SEC, buffer exchange and subsequent analysis with ESI-MS[77].

In this study, SEC was used to determine the levels of aggregation of in-house, produced variants and subsequently for the fractionation of Immunoglobulin-degrading Enzyme –Staphylococcus (IdeS) and reduced mAb fragments.

1.6.1.3. Reverse Phase Chromatography (RP)

Reverse phase chromatography employs a non-polar stationary phase with a polar mobile phase, the “reverse” of normal phase chromatography. It’s mode of action depends on an analytes affinity for the hydrophobic groups of the stationary phase. Elution with a polar mobile weakens the hydrophobic bonds and releases the analyte[78]. Amino acids have a hydrophobic core, with hydrophobic/hydrophilic functional groups, which in turn infers an overall hydrophobic nature to proteins[79]. RP is a staple for protein analysis, as it can be easily coupled to MS and features a high separation efficiency for proteins[80]. Figure 1.6.3 shows the structure of a silica bed column with a hydrocarbon tail, which are the most commonly, used RP columns. For protein analysis, C_4 or C_8 carbon lengths are usually employed, as proteins exhibit high retention on longer, more hydrophobic, chains. End capping of RP stationary phases is typically employed in the production of other chromatographic modes. It is also a necessity when secondary interactions occur as a result of un-bonded silanol groups on the support bed[81]. Figure 1.6.3 illustrates the common structure of an RP-LC stationary phase.

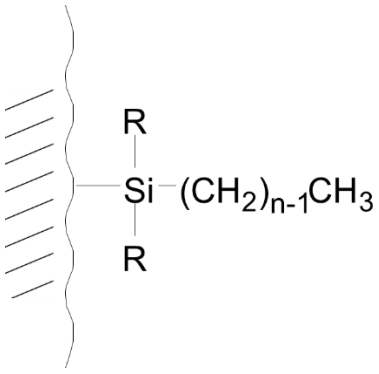


Figure 1.6.3: Illustration of silica bonded Cn stationary phases. Where n is the total carbon in the polymer chain.

For the identification and quantitation of modifications on mAbs, RP-MS is widely employed. The expansive application of RP biopharmaceutical characterisation should be acknowledged. In terms of a singular universal mode for sequence,

structure and modification determination, RP-LC can achieve each of these feats while also enabling MS detection[45]. It is also often employed as the second dimension in 2D separations due to universality. RP-LC was primarily used in the studies for this thesis for peptide mapping analysis.

1.6.1.4. Hydrophilic Interaction Chromatography (HILIC)

HILIC is a form of normal phase chromatography, a polar stationary phase with an aqueous-organic mobile phase. The constituents of HILIC mobile phases are similarly MS compatible, as with NP, but are found in higher, non-polar concentrations. This is a benefit over NP, as many polar and hydrophilic molecules cannot be dissolved in the highly organic solvents. Many of the solvents used in NP are also highly dangerous and not environmentally friendly[82].

HILIC enables MS analysis of compounds which would typically be excluded from RP analysis[59]. Though buffer constituents are the same as RP, the strong, eluting solvent of HILIC is aqueous, typically water, while the organic component is the weaker eluting component. As the stationary phase is polar, aqueous molecules concentrate there and form an aqueous rich layer. The retention mechanism of HILIC is due to the polar analytes partitioning between the organic-aqueous mobile phase and the aqueous rich layer[83]. Figure 1.6.4 illustrates the inter-connectivity between HILIC, RP, IE and NP chromatography. Where HILIC and IE have similar separation mechanism for analytes, a separation related to charge. RP and HILIC have similar mobile phase constituents (eluent), which enable MS hyphenation, and NP and HILIC have the same polarity stationary phases(adsorbent)[84].

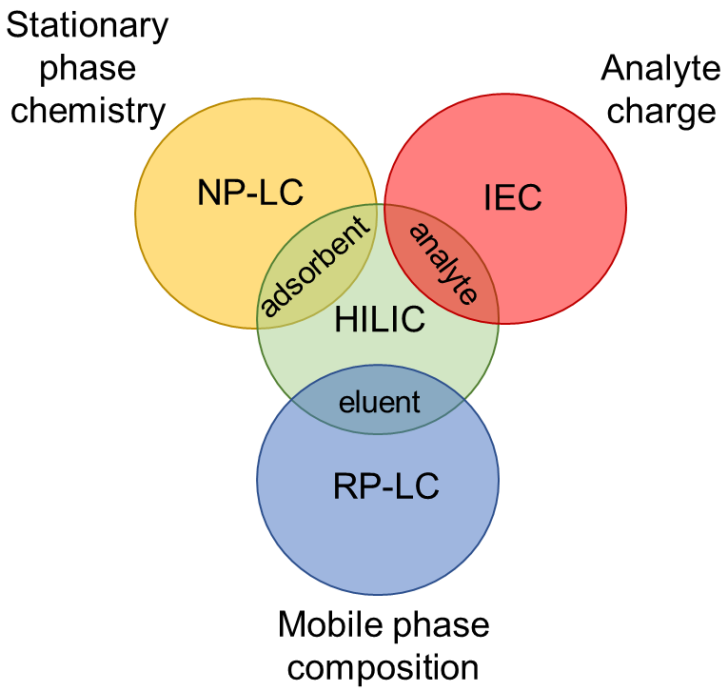


Figure 1.6.4: Illustration of the inter-relation between HILIC and other chromatography modes[84]

Due to the mobile phase compatibility and with selection of an appropriately charged HILIC stationary phase, a 2D separation with HILIC and RP has been shown to be a good alternative to IEC for MS determination of mAb variants[85]. Though it can be readily used for the analysis of proteins, organic acids and nucleotides, the arguable niche of HILIC is in carbohydrate analysis[59, 86]. HILIC was used extensively in this thesis for the purification and analysis of N-linked glycans. The detection mode for HILIC is fluorescence, when applied to carbohydrate analysis. This requires that glycans be fluorescently labelled prior to analysis, Chapter 4, section 2.9 discusses the reaction involved. Larger, more hydrophilic carbohydrate moieties have increased affinity for the stationary phase and so are retained longer and so a separation-by-size aspect is also at play[85, 86].

1.6.1.5. Hydrophobic Interaction Chromatography (HIC)

As HILIC is a form of NPLC so the reverse is true for hydrophobic interaction chromatography (HIC), it is a form of RPLC. The stationary phase is non-polar with polar eluents. However, HIC is used with non-denaturing mobile phases and so is

particularly useful for protein characterisation as it maintains the molecules native state[63]. In section 1.4, ΔG_{ND} for protein conformational stability was discussed, an important contributor to a proteins higher ΔG_{ND} value, and therefore increased conformation stability, is the number of hydrophobic "pockets" on the molecule[87]. Using HIC, the level of hydrophobic regions on the protein can be determined. HIC was not used during this thesis work; however the usefulness of the method is worth stating, in relation to biopharmaceutical characterisation.

1.6.2. Mass Spectrometry

A mass spectrometric analysis determines the mass to charge ratio (m/z) of an analyte. Therefore, characterisation of any analyte requires that the sample be in a charged state. The charged analyte is separated in the mass analyser regions of the instrument. Traditionally, the mass analyser consisted of a magnetic field, through which, the ions travel in a vacuum[88]. However, advancements in mass analyser technology have seen the emergence of new families of instrument, such as Time of Flight and Orbitrap which have revolutionised biomolecule mass spectrometry in terms of resolution, mass accuracy and sensitivity. Figure 1.6.5 illustrates the basic schematic of MS instrumentation.

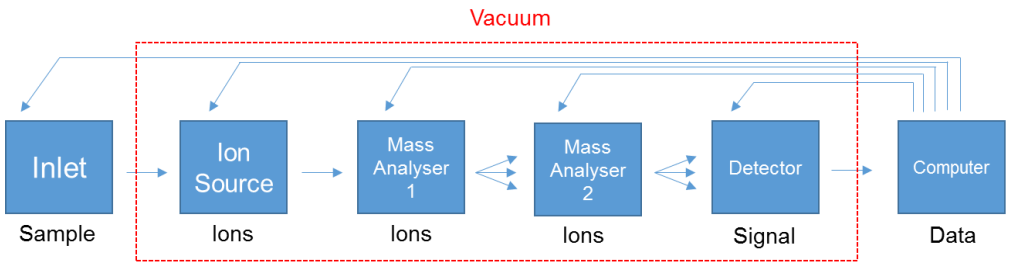


Figure 1.6.5: Basic schematic of key components of a mass spectrometer

1.6.2.1. Ion Sources for biopharmaceutical analysis

Prior to the introduction of "soft" ionisation techniques, a biomolecule's mass was determined by gel electrophoresis or chromatographic methods. However, other

physical characteristics of the molecule interfered with accurate determination[88]. Mass spectrometers, at the time, required the analyte to be in the gas phase upon ionisation and so analyses of biomolecules required derivatization.

Howard Morris and co-workers pioneered the analysis of large biomolecules by mass spectrometry. In his 1979 paper he describes approaches for “mass spectrometric attacks” on peptide sequences[89]. Derivatization required three overall processes, though also required prior testing of the sample: (1) Acetylation by acetic anhydride in methanol; (2) permethylation, which could be unsuccessful on peptides containing Cys, His or Met; (3) finally peptides containing Arg required converting the residue to ornithine[89]. Morris recognised the potential of analysing biomolecules by MS. At that time, chemical ionisation and electron ionisation were the primary ionisation strategies. Morris and co-workers used a magnetic sector instrument hyphenated to a gas chromatography system to undertake their initial studies[90]. Field desorption and fast atom bombardment ionisation techniques rapidly overtook derivatization as a means to analyse large, polar and thermally labile compounds[91].

The time of flight (TOF) mass analyser had been available commercially since the 1960’s and had been used primarily for isotopic analysis of organic molecules. With the advent of matrix assisted laser desorption ionisation (MALDI) and electrospray ionisation (ESI) in 1988 and 1989 respectively[92, 93], and pairing of ionisation and mass analyser enabled the analysis of large biomolecules. The breadth of knowledge and research on ionisation techniques for mass spectrometry is expansive and expanding. Therefore, for the purposes of this review, ESI is the focus as it was the sole ionisation method employed experimentally.

1.6.2.1.1. Electrospray Ionisation (ESI)

The principle of electrospray ionisation is; a sample passes through a capillary needle and disperses in a fine spray. As it passes out of the needle it is introduced to an electric field. The field is produced by creating a potential between the capillary and a counter electrode.

The exact mechanism of droplet formation is well researched, as it provides a fascinating insight in to fluid dynamics in an electric field[94, 95]. However, it is also highly debated. There are two theories on ionisation mechanism in ESI; the charge residue model and the ion evaporation model[96]. Spray formation is the main source of ions in ESI, independent of the ionisation mechanism.

Figure 1.6.6, below, illustrates the spray formation in ESI. On the left hand side of the image, a low voltage is applied, the drop is spherical, upon increasing voltage, charge accumulates on the molecules present in the liquid and the surface tension comes under pressure, elongating the drop. When the onset voltage is applied, the surface tension breaks and a Taylor cone forms. At this voltage, the charge field at the apex of the drop is counter balanced by the surface tension energy of the liquid, this phenomenon is referred to as the Rayleigh limit.

Figure 1.6.6: The formation of droplets in an electrospray potential with increasing voltage. Left; low voltage, centre; 1-2 kV, right; at onset voltage. Image reproduced with permission from Hoffmann (2007)[88]

The Rayleigh limit is a mathematical equation which describes the formation of droplets from a larger body of liquid[96]. The onset voltage is specific for each fluid, for example water has an onset voltage of 4 kV while acetonitrile has an onset

voltage of 2.5 kV[88]. This process repeats in increasingly smaller droplets each with locally formed Taylor cones. It is from this point that ionisation models play a more crucial role.

Figure 1.6.7 illustrates the proposed model of ion evaporation, in this instance ΔG represents the energy required to increase the solvated surface area versus the columbic pull of the ion in the field. As shown, the energy involved in these opposing forces increases the overall energy of the system, until the ion is expelled, where the overall energy of the system is decreased and so the energy required to expel another ion is also decreased and so this process continues until all ions have been expelled[96].

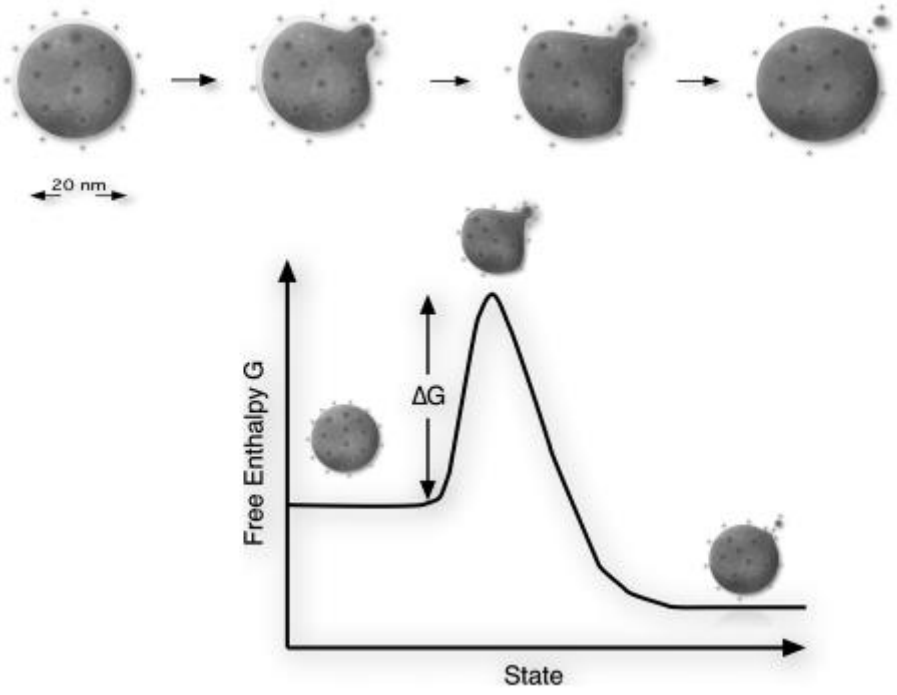


Figure 1.6.7: The ion evaporation model. Reproduced with permission from Wilm *et al* (2011)[96]

The assumption of the charge residue model, shown in Figure 1.6.8, is that the spray process creates a one-molecule-per-droplet dynamic. This is the fundamental difference between the two models. Where ion evaporation disregards molecule concentration within the droplet and bases ionisation on the evaporation

of the droplet and the attractive forces of the field. The charge residue models assumes that the charge on the final gas phase model comes from the evaporated solvent rather than the molecule itself[97].

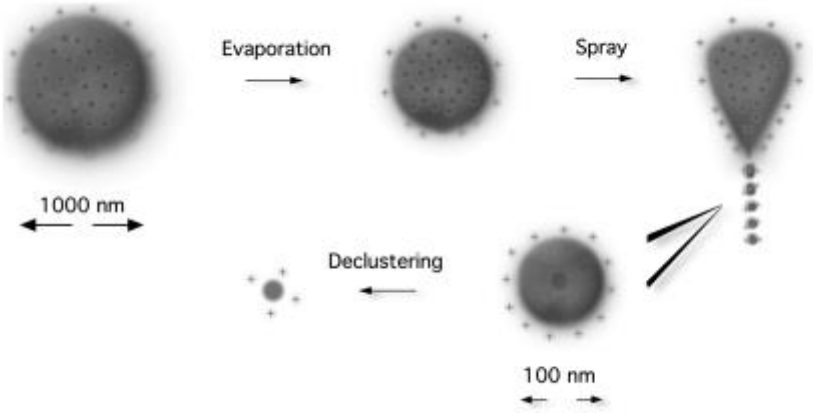


Figure 1.6.8: The charge residue model. Reproduced with permission from Wilm *et al* (2011)[96]

It is the accumulation of charge during the desorption process which gives rise to one of the main advantages of ESI over other ionisation sources; the formation of multiply charged ions (MCI). The formation of MCI on high molecular weight compounds translates to increased sensitivity at the detector, as MS detects m/z ratios rather than mass alone. A mass analyser with relatively low mass range will still be capable of detecting large molecules, if they are multiply charged[96]. Sensitivity in ESI is also linked to the concentration, per unit of flow i.e.; a high concentration of sample delivered in a small volume to the ESI source increases sensitivity at the detector[88].

Due to its relative ease of use, robustness and increased sensitivity, ESI is a staple for biomolecular analysis. The advent of nanoelectrospray ESI (nESI) has furthered the distance between ESI and MALDI. The benefit of nESI over wider diameter ESI

is that the analytes reach the final droplet stage, just prior to gas phase, in less time[97]. This benefits biochemical analysis, especially with precious sample, by an increased efficiency in ionisation and by reducing the effect of ion suppressing contaminants[96]. The production of multiply charged ions allows use with modest mass analysers and so it is arguably more accessible than MALDI.

In this body of work, one of the main advantages of ESI was utilised, the ability to analyse mAbs in the native state. ESI has been capable of analysing intact biopharmaceuticals since its inception, the caveat being that they were denatured by the acidified aqueous/organic buffers used. This results in the loss of the protein's 3D structure but enabled mass accuracies of ≤1 Da[63]. With the arrival of larger mass range analysers, volatile buffers which retained the physiological conditions of the mAb could be used, at the expense of highly charged analytes, to analyse the protein in its truly native conformation[98]. Buffers such as aqueous ammonium acetate preserve the 3D conformation, while inferring fewer charges on the molecule. This results in fewer native charge states which occur at a higher m/z, with the added benefit of greater space between each state allowing for greater resolving power and ease of assignment[63]. Figure 1.6.9 illustrates the observable differences between a protein analyses in the native and denatured intact states.

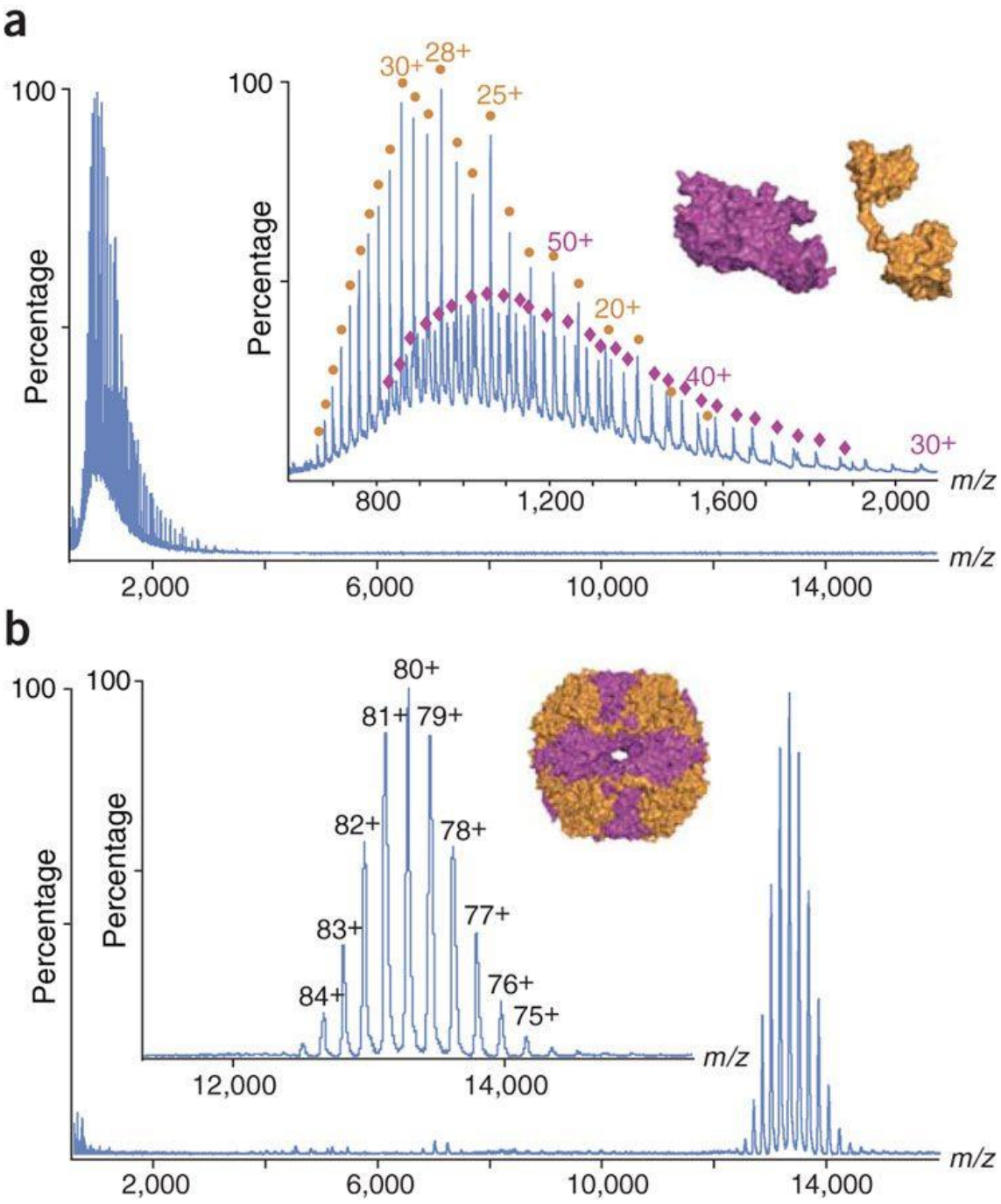


Figure 1.6.9: The analysis of the enzyme H. pylori urease by native and denaturing MS. Reproduced with permission from Heck *et al*. (2008)[99].

1.6.2.2. Mass Analysers

Mass analysers can be broadly defined by the separation field they induce, the two types are by electric or magnetic fields, and within these a static or dynamic field. The two are not exclusive however, and a combination of magnetic and electric fields can be employed[88]. Most modern mass spectrometers have a number of mass analysers coupled together, which allow for further analysis of selected ion;

this will be further discussed in section 1.6.2.2.3. Table 1.6.3 summarises the individual mass analyser technologies, their separation technologies and the fields employed to achieve separation.

Table 1.6.3: Summary of mass analyser technology and their separation principle [88]

Type of Analyser	Separation Principle	Applied field
Electric Sector (E/ESA)	Kinetic energy	Electric
Magnetic Sector (B)	Momentum	Magnetic
Quadrupole (Q)	m/z (trajectory stability)	Electric
Ion Trap (IT)	m/z (resonance frequency)	Electric
Time-of-flight (TOF)	Velocity (flight time)	No field. Reflectron technology induces an electric field to increase sensitivity
Fourier transform ion cyclotron resonance (FTICR)	m/z (resonance frequency)	Magnetic/electric
Fourier transform Orbitrap (FT-OT)	m/z (resonance frequency)	Electric

During the course of this work, a hybrid mass spectrometer was used which featured a quadrupole-Orbitrap mass analyser. Each of these mass analysers are discussed in further detail below. The other features of hybrid instruments and the application of tandem mass spectrometry will be discussed in section 1.6.2.3, "*state of the art instrumentation."*

1.6.2.2.1. The Quadrupole

Quadrupole analysers consist of four parallel rods. Each adjacent rod has an opposing charge; while opposite rods have the same charge. This opposing charge creates an electric field which oscillates at a given frequency (RF). When an ion, of any charge, travels into the quadrupole, its path is disrupted from a straight line to an oscillation due to the attractive and repulsive forces on the rod charges[100].

The RF and the direct current (DC) applied to the rods can be selected by the MS operator to exclude certain ions which have an m/z that would be unstable in the

oscillating field and so would spin out toward the rod. This is illustrated in figure 1.6.10 above, where, the green line represents an ion with the selected m/z and the red line is the path of the ion which is filtered out[101].

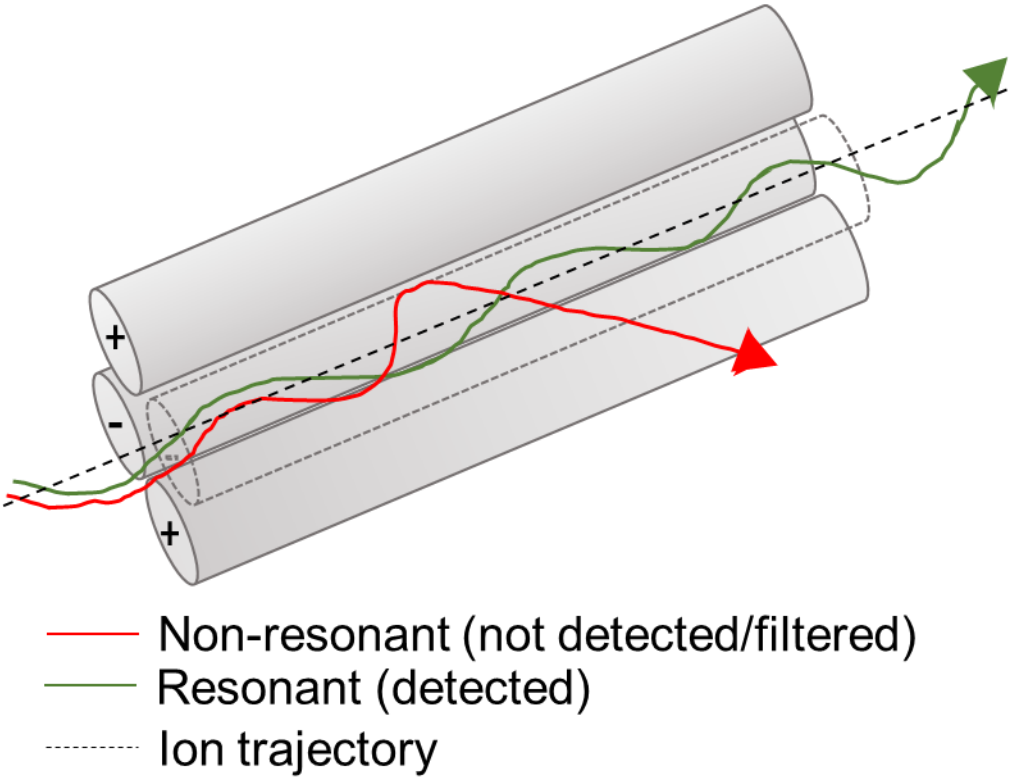


Figure 1.6.10: A basic representation of an ion being filtered in a quadrupole mass analyser.

1.6.2.2.2. Ion traps

The Orbitrap mass analyser was developed by Alexander Makarov in the 1990's. Kingdon (1923) had first put forth the idea of trapping ions in an electrostatic field created by an outer and inner electrode[102]. Ions injected into the Orbitrap are electrostatically trapped there, oscillating and rotating, around and along, the central electrode. The oscillating ions generate a charge image on the outer electrode which is Fourier transformed to a frequency corresponding to their m/z[103]. Figure 1.6.11 shows the schematic of the Orbitrap mass analyser.

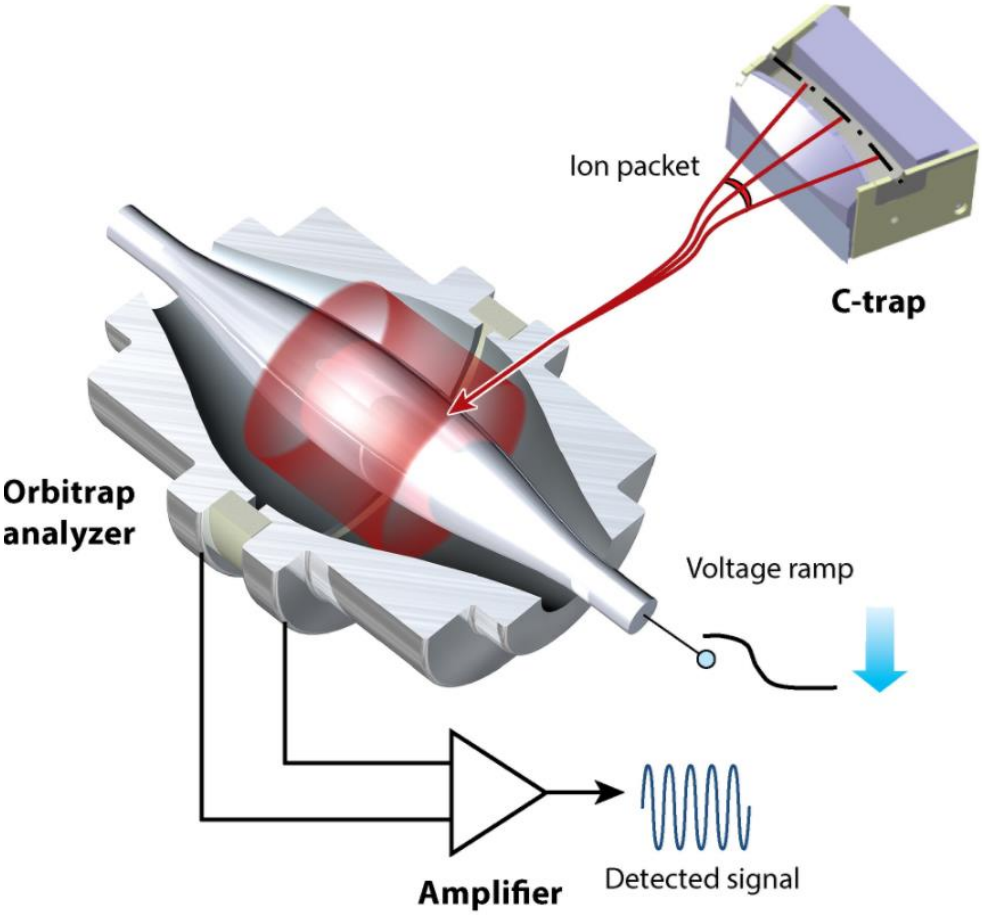


Figure 1.6.11: Schematic of Orbitrap mass analyser with C-trap ion accumulation device[104]

Comisarow and Marshall first reported on the development of the ion cyclotron resonance (ICR) mass analyser in 1974[105]. ICR technology works in an applied magnetic field, an rf oscillating field is applied which focuses the ions and creates a cyclotron frequency. The cyclotron frequency is then converted by fourier transformation into a digitized mass spectrum[106]. It is this cyclotron frequency which enables the highest mass resolution of any mass analyser and provides a mass accuracy 10x times higher than the Orbitrap[106]

The ion trap has revolutionised mass spectral analysis of proteins. Higher resolution, sensitivity and vast mass ranges are creating new avenues for researchers outside of the analytical field.

The extraordinary capabilities and the potential this technology offers are only beginning to be understood, and even now classical biological techniques are being foregone for the wealth of information that can be obtained from MS[103].

1.6.2.2.3. Tandem Mass Spectrometry

Tandem mass spectrometry is used for the accurate identification and quantitation of molecules. A tandem MS (MS/MS) instrument consists of two mass analysers connected either in space or in time. MS/MS in space is the physical linkage of two mass analysers in sequence, where selection, MS^1, analysis occurs in the first analyser and fragmentation occurs in the second. MS/MS in time requires ion trap mass analysers which can simultaneously hold and analyse ions. In time MS/MS, ion selection analysis is usually performed by a quadrupole.

There are two modes of MS/MS data acquisition; data dependent acquisition (DDA) and data independent acquisition (DIA). DDA selects the most abundant ions in a spectrum for further fragmentation and analysis[107]. Selected reaction monitoring (SRM) in DDA allows the analyst to select a single ion of a particular mass for further fragmentation, thus enabling focused analysis. DIA fragments and analyses ions from a user defined m/z range. The benefit of DIA over DDA is the ability to fragment ions that would have not have been selected without the need to perform a single ion SRM[108]. Sequential window acquisition of all theoretical fragment ion spectra (SWATH) is a DIA strategy which selects windows of m/z ranges for MS/MS analysis. It has been a significantly important advantage for large proteome experiments[109]. A disadvantage of DIA however is the noisy spectra which are produced and the extensive data processing which is required[110].

1.6.2.3. State of the art instrumentation

The work horses of a mass spectrometer are arguably the mass analyser and the ion source. However, these are common features to all MS instrumentation. In modern instruments, which easily deliver accurate, high resolution data, there is a

lot more going on behind the cover. In this section, the specific instrument features of the ThermoScientific Q Exactive™ Plus BioPharma Mass Spectrometer (QE) will be discussed. Figure 1.6.12, below, shows a schematic representation of the QE.

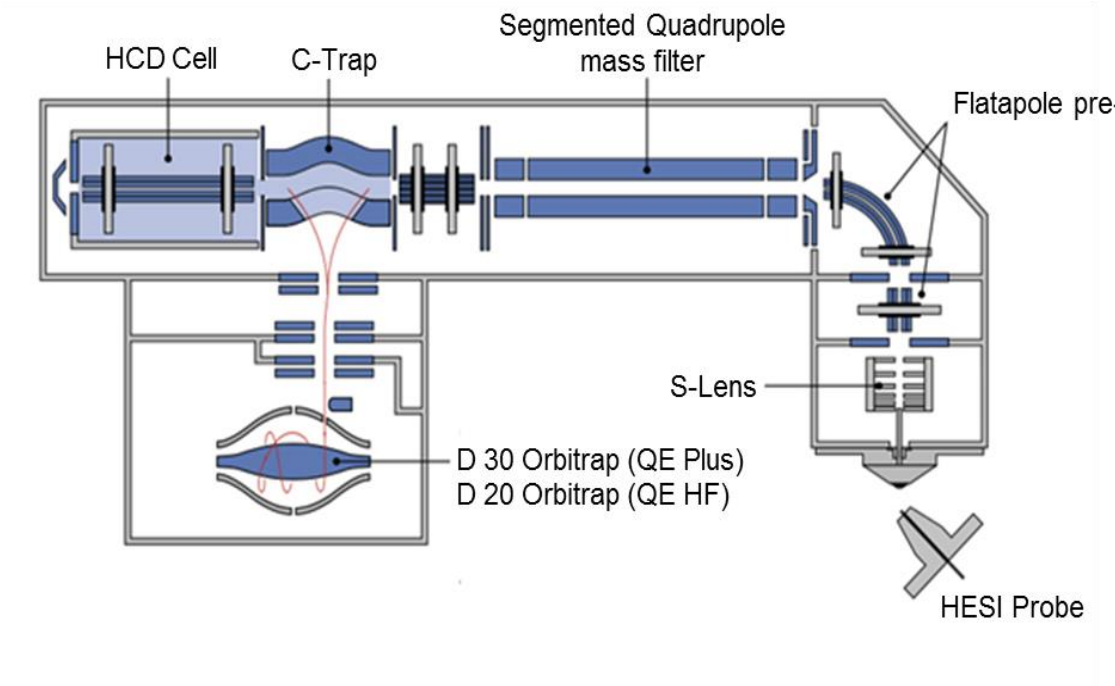


Figure 1.6.12: Schematic of a Q Exactive BioPharma Plus Mass Spectrometer

A heated ESI source (HESI) differs from a standard ESI by heating the auxiliary gas, to temperatures ranging from 200°C to 600°C. An ion beam generated under ambient conditions, often has poor ion transfer efficiency. The use of a HESI source has been shown to positively affect the velocity, density and mobility of ions[111]. In a comparison of MS instrumentation, with heated or unheated ESI, analysing *E.coli* cell protein and metabolites, Lu *et al*. (2008) showed that the heated source had five-fold increase in absolute ion counts[112].

In previous versions of this instrumentation, it was found that, excluded peptide ions were building up and contaminating the quadrupole[113]. In more recent versions, the flatapole pre-filter has been introduced. This device is a low resolution quadrupole in function but physically consists of four rounded rods with flat exteriors; Figure 1.6.13 shows an image of the flatapole filter. For intact protein

analysis, this results in lower fragmented peptides and other contaminants being excluded from the main quadrupole filter. This in turn enables more highly resolved peaks.

Figure 1.6.13: The "flatapole"

The inclusion of a pre-filter, was shown to benefit selectivity and result in a 90% decrease in signal detected, from excluded ions[113].

After pre-filtering, the bulk of the ion beam then travels through the second filtering quadrupole, which is more specific than the flatapole. The segmented quadrupole acts as a mass filter, which can isolate very narrow mass windows. This is an advantage in data independent acquisition, and in particular SWATH approaches.

The advantage of having the quadrupole mass filter was highlighted by Michalski *et al*. (2011) where a HeLa cell lysate was analysed in a SRM experiment. The MS instrument was an Exactive series which, comprises of an Orbitrap analyser and a multipole ion focus. The instrument was adapted to have a quadrupole filter prior to the Orbitrap[114]. They chose three peptide mass ranges, which were known to be low abundant, following analysis ,they showed a marked increase in sensitivity for these peptides and additional peptides which had previously been undetected[114]. For complex proteomic experiments, such as intact, native mAb analysis, the ability to filter out overly abundant proteins is key.[111, 113]

The C-trap is a curved RF flatapole which provides both, collisional cooling of the ion beam and accumulation of selected ions. Ions enter the Orbitrap from the C-trap; ions also enter the HCD cell from the C-trap. Once the selected amounts of ions have cooled in the C-trap, the RF is ramped down and in the absence of the field, the ions are no longer held in the C-trap. The ions are then accelerated into the Orbitrap where mass analysis takes place.

In MS/MS mode, the precursor ions are fragmented in the higher energy collisional dissociation (HCD) cell. For accurate peptide sequencing, the quality of fragmented ions is very important. Diedrich *et al.* (2013) used a HCD cell to determine the effect of stepped collisional energy on the analysis of HEK293 cell lysates[115]. The stepped method involves fragmenting ions at low, medium and high energies producing peptide fragments of varying sizes. This results in a larger pool of overlapping ions and aids in database searching and accurate identification of peptide sequences. Figure 1.6.14 below, is an excerpt from the Diedrich study where they found that, using a stepped energy HCD strategy resulted in more precise determination of phosphorylation sites on HEK cell lysate peptides. In a comparison with standard collision induced dissociation (CID) as a fragment ion source, they found that HCD provided more “usable” fragments[115].

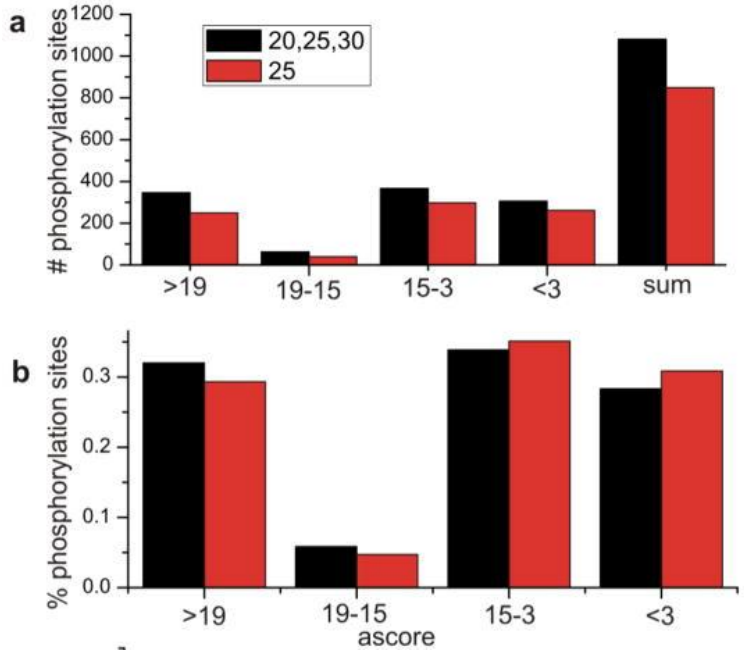


Figure 1.6.14: (A) & (B) comparing the number and percentage of accurately identified phosphorylation sites (Ascore) in a single energy and stepped energy experiments. Reproduced with permission from Diedrich *et al.* (2013)[115]

Orbitrap mass analysers provide higher mass accuracy than the majority of mass analysers. However, FT-ICR has been shown to provide mass accuracies well above those achieved by the Orbitrap.

In this body of work a D30 Orbitrap was used. The rapid advancement in mass analyser technology is illustrated by the increasingly frequent updates to the Orbitrap design. Figure 1.6.15 shows the differences in scale between the two types of Orbitrap analyser available.

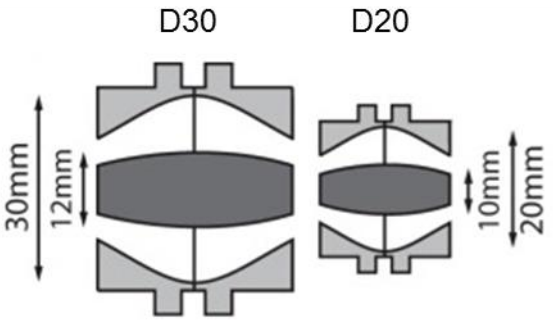


Figure 1.6.15: A basic schematic illustrating the differences in the D30 and D20 Orbitrap mass analysers

The difference between these two is the size. The decrease in size results in an increase in the electric field and therefore, an increase in the detected frequency[113]. In turn, this means a higher mass resolution, in a reduced amount of time, which enables a higher MS/MS scan rate to be performed providing enhanced mass detail.

1.6.2.4. Informatics

Data analysis for peptide and protein profiling experiments require powerful algorithms to enable the determination of sequence and PTM's.

1.6.2.4.1. Peptide mapping analysis

There are three approaches to peptide mapping data interpretation; *de novo* sequencing, spectral database matching and sequence data base matching. *De novo* sequencing is typically used for a sample which there is no sequence available. It identifies peptides by peak pattern recognition in MS/MS spectra. Figure 1.6.16 shows the fragmentation of amino acids in HCD.

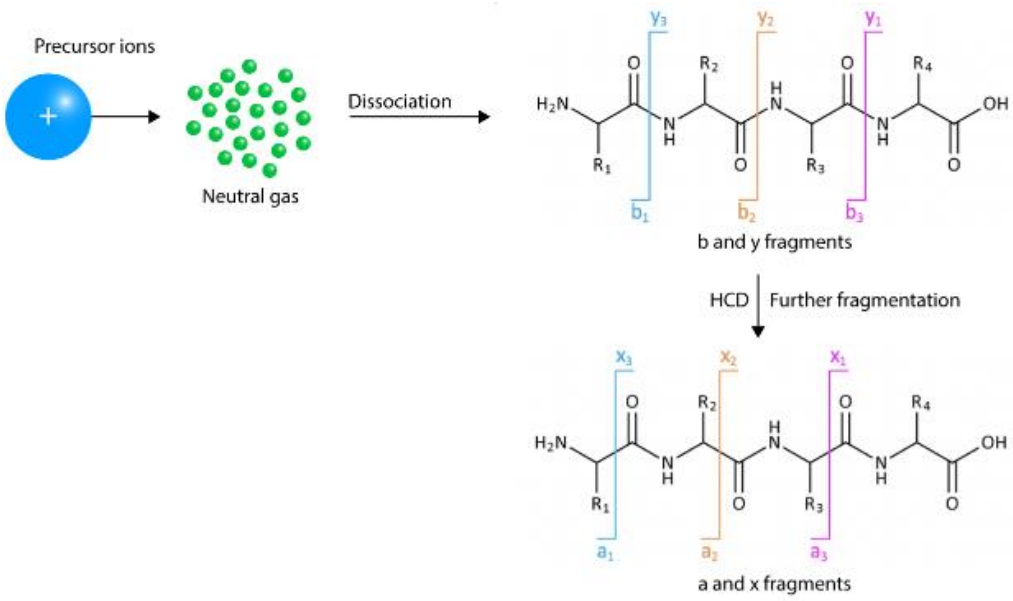


Figure 1.6.16: Fragment ion formation in HCD

The PEAKS platform computes the possible combinations of all amino acids for a detected precursor ion. The PEAKS *de novo* sequencing algorithm considers the

a, b, x, y and b/y−17/18 ions for the prediction of amino acid identity and sequence[116].

Spectral database identification searches a library of spectra with reference to the sample MS/MS and assigns identifications based upon the statistical similarities.

In this body of work, the sequence of all mAbs were known and so sequence matching was performed using both ThermoScientific's BioPharma Finder 3.0 and Protein Metric's Byonic and Bylogic. Software platforms such as these produce *in silico* digestions of the sequence submitted and compare the theoretical masses of each fragmentation to the real MS/MS spectrum. The output typically consists of a sequence coverage map and a table of identified peptides[117].

1.6.2.4.2. Intact and middle up protein identification

Spectral library searching is also used to identify protein domains in top-down and middle up proteomic experiments. Multiply charged protein signals are de-convoluted to produce theoretical, singly charged ions, which are then compared to the reference sequence and the theoretical spectra to accurately identify protein domains. Frequently used deconvolution algorithms are those such as ReSpect, Xtract and Sliding Windows in the BioPharma Finder software. The Xtract algorithm detects isotopically resolved peaks and uses the peak spacing of the cluster to estimate an initial mass. It then uses a distribution model to analyse the peak distribution and fits a best monoisotopic identity. Finally it combines all observed charge states and combines them to produce a single mass value[118]. The ReSpect algorithm is used to deconvolute isotopically unresolved peaks to the neutral average mass. To aid deconvolution, it first subtracts the baseline, then performs a peak deconvolution, followed by a charge deconvolution to convert from m/z to mass. Figure 1.6.17 below shows the de-convoluted mass spectra of an anti-IL8 IgG1 using the ReSpect algorithm[118].

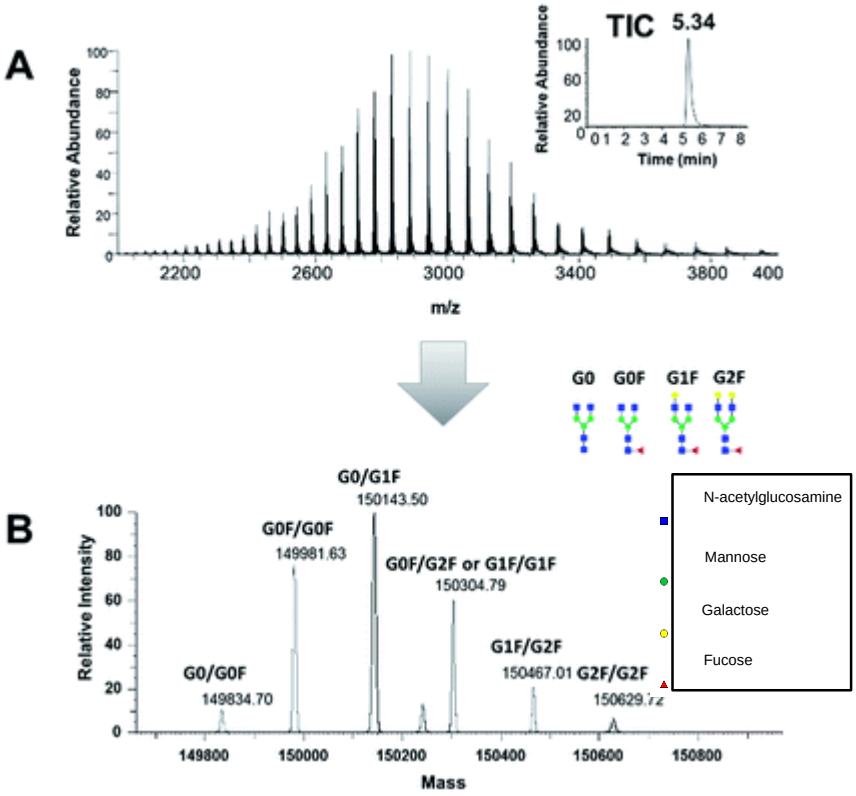


Figure 1.6.17: (A) Full MS spectra of a mAb with Total Ion Chromatogram (TIC) trace in outset. (B) The ReSpect de-convoluted spectra showing the accurate identification of various glycoforms[118]. Reproduced with permission from Farrell *et al.* (2018)

This IgG was analysed in the native state and so de-convolution was simplified due to the space between charge states. As (B) above shows analysing in the native form allows identification to a very precise level, up to the absence/presence of a fucose sugar molecule on a glycan on one amino acid site not always occurring at 100% abundance[118].

In this work, the main deconvolution approach used was the sliding window algorithm. This process utilises either Xtract, for isotopically resolved peaks, or ReSpect, for unresolved peaks. Initially it scans peaks by retention time windows and creates an average of each successive window. These windows are then deconvoluted and the component peaks are identified. In the case of this body of work these peaks are used to analyse mAb subunits. The next step merges the masses of these peaks and then applies a user defined filter, either by a mass

window of constant width, used for isotopically resolved peaks or by a defined mass difference, for unresolved.

1.6.2.4.3. N-glycan identification

Glycan structure determination depends on the separation and detection mode employed. The most widely used method for N-glycan annotation is the use of an exoglycosidase panel followed by MS analysis. An exoglycosidase panel is a series of enzymes which have cleavage specificity for the sugar linkages on N-glycans. By having prior knowledge of the enzyme specificity, the structural order can be elucidated by sequentially digesting the N-glycan and analysing by HILIC with either fluorescent or MS detection.

In this work, three methods were used; release of the glycan from the protein by PNGase F followed by an exoglycosidase panel, secondly MS analysis of intact N-glycans with reference to the exoglycosidase panel and finally analysis of the glycopeptide by native MS. The third method is to perform MS/MS analysis on released glycans.

The informatics approach to MS spectra of glycan structures involves manual searching base peak chromatograms and subsequent searching of the m/z in GlycoWorkbench, an open source program developed by Ceroni and co-workers (2008)[119]. GlycoWorkbench searches the user m/z input against databases of known glycan structures and outputs potential structures based on the user defined mass accuracy.

Similarly to proteomic experiments, there are informatics tools which can aid in glycan identification. One such program is Premier BioSoft's SIMglycan. It uses a scoring algorithm similar to *de novo* sequencing which matches MS/MS fragments to theoretical digests based on the sample precursor ions[120].

Chapter 1.

Proteomics platforms are incorporating glycan structures into their PTM searches. In this work, major glycoforms were accurately identified by proteomic-centric software.

Chapter 1.

1.7. Conclusion

Monoclonal antibodies are rapidly becoming the vital model for future therapies. They are specific, versatile and a powerful means of targeting diseases. Their production and manufacturing process is being constantly reviewed and updated. In this thesis, two of the most commonly used cell lines are transfected to produce fully human variants of commercially available products. The use of HEK293 and CHO provides interesting insight into the effect the production process has on the 3D conformation of a mAb. PTM's are assessed by rigorous state of the art instrumentation and methodology. The analytical method which provides the most information about a protein is arguably LC-MS.

In this work an SCX-MS was developed to determine the PTM's of various mAbs. The method was then compared to traditional liquid phase separations of data generation using state of the art instrumentation and modern bioinformatics approaches. This method is a step towards developing a manufacturability platform for the triage of biotherapeutic candidates.

During the course of this review of literature pertaining to these methods and technologies, valuable insight has been gained and applied to the investigations for this thesis. No review, however, can ever be a definitive representation of the breadth of knowledge available. This is evident from the rapidly advancing MS technology discussed here, particularly in the area of mass analysers. As well as from a technological perspective, the arena of regulatory requirements is also changing as more biosimilars come to market, the question is asked how similar is similar and now, with the level of information that can be gained, the question is asked even louder.

Chapter 1.

1.8. Author Contributions

Literature review topic devised by Jonathan Bones and Anne Trappe; Literature review written by Anne Trappe and reviewed by Jonathan Bones.

Chapter 1.

1.9. Specific Aims

The overall aim of this thesis was to develop a manufacturability platform for the molecular triage of biotherapeutics.

In the pursuit of this overall goal the specific aims were:

- To construct and transform monoclonal antibody VH and VL genes into *E. coli* cells and subsequently transfect in HEK293 and CHO cells
- To investigate the optimum elution mode of SCX chromatography and develop a method for the high throughput analysis for the determination of the CV profiles of the in-house produced mAbs.
- To develop upon a CVA-MS method for the inclusion of extreme pI antibodies and to reduce the formation of adducts.
- To fully characterise the in-house molecule cetuximab by middle up, intact and peptide mapping approaches and to fully characterise the glycan structures present.
- Finally, to compare CVA-MS approach of analysing cetuximab to the data generated by traditional means.

Chapter 1.

1.10. References

1.Kontermann, R. E.; Brinkmann, U., Bispecific antibodies. *Drug Discovery Today* **2015,** *20* (7), 838-847.

2.Czajkowsky, D. M.; Hu, J.; Shao, Z.; Pleass, R. J., Fc-fusion proteins: new developments and future perspectives. *EMBO Molecular Medicine* **2012,** *4* (10), 1015-1028.

3.Beck, A.; Goetsch, L.; Dumontet, C.; Corvaïa, N., Strategies and challenges for the next generation of antibody-drug conjugates. *Nature Reviews Drug Discovery* **2017,** *16*, 315.

4.Fekete, S.; Guillarme, D.; Sandra, P.; Sandra, K., Chromatographic, Electrophoretic, and Mass Spectrometric Methods for the Analytical Characterization of Protein Biopharmaceuticals. *Analytical Chemistry* **2016,** *88* (1), 480-507.

5.Ecker, D. M.; Jones, S. D.; Levine, H. L., The therapeutic monoclonal antibody market. *mAbs* **2015,** *7* (1), 9-14.

6.Ribatti, D., From the discovery of monoclonal antibodies to their therapeutic application: An historical reappraisal. *Immunology Letters* **2014,** *161* (1), 96-99.

7.Rodgers, K. R.; Chou, R. C., Therapeutic monoclonal antibodies and derivatives: Historical perspectives and future directions. *Biotechnology Advances* **2016,** *34* (6), 1149-1158.

8.Lenting, P. J.; Denis, C. V.; Christophe, O. D., Emicizumab, a bispecific antibody recognizing coagulation factors IX and X: how does it actually compare to factor VIII? *Blood* **2017**.

9.Kaplon, H.; Reichert, J. M., Antibodies to watch in 2018. *mAbs* **2018**, 1-21.

10.European Medicines Agency, Meeting highlights from the Committee for Medicinal Products for Human Use (CHMP) 22-25 January 2018. 2018.

11.Hattori, K.; Kojima, T.; Miyazaki, T.; Soeda, T., Bispecific antibody substituting for functional proteins. Google Patents: 2007.

12.Lonza, Drug development cycle and the 'valley of death'. Schematics of the drug development lifecycle, duration of each one of the stages, capitalized costs and probability of failure. 2015.

13.*Emicizumab for Hemophilia A: Effectiveness and Value*; Institute for Clinical and Economic Review: 2018.

14.Gaughan, C. L., The present state of the art in expression, production and characterization of monoclonal antibodies. *Molecular Diversity* **2016,** *20* (1), 255-270.

15.Jarasch, A.; Koll, H.; Regula, J. T.; Bader, M.; Papadimitriou, A.; Kettenberger, H., Developability Assessment During the Selection of Novel Therapeutic Antibodies. *J Pharm Sci-Us* **2015,** *104* (6), 1885-1898.

16.Schuurman, J.; Graus, Y. F.; Labrijn, A. F.; Ruuls, S. R.; Parren, P. W. H. I., Opening the door to innovation. *mAbs* **2014,** *6* (4), 812-819.

17.Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L., Principles of early drug discovery. *British Journal of Pharmacology* **2010,** *162* (6), 1239-1249.

18.Bazan, J.; Całkosiński, I.; Gamian, A., Phage display—A powerful technique for immunotherapy: 1. Introduction and potential of therapeutic applications. *Human Vaccines & Immunotherapeutics* **2012,** *8* (12), 1817-1828.

19.Fennell, B. J.; McDonnell, B.; Tam, A. S. P.; Chang, L.; Steven, J.; Broadbent, I. D.; Gao, H.; Kieras, E.; Alley, J.; Luxenberg, D.; Edmonds, J.; Fitz, L. J.; Miao, W.; Whitters, M. J.; Medley, Q. G.; Guo, Y. J.; Darmanin-Sheehan, A.; Autin, B.; Shúilleabháin, D. N.; Cummins, E.; King, A.; Krebs, M. R. H.; Grace, C.; Hickling, T. P.; Boisvert, A.; Zhong, X.; McKenna, M.; Francis, C.; Olland, S.; Bloom, L.; Paulsen, J.; Somers, W.; Jensen, A.; Lin, L.; Finlay, W. J. J.; Cunningham, O., CDR-restricted engineering of native human scFvs creates highly stable and

soluble bifunctional antibodies for subcutaneous delivery. *MAbs* **2013,** *5* (6), 882-895.

20.Declerck, P.; Farouk Rezk, M., The road from development to approval: evaluating the body of evidence to confirm biosimilarity. *Rheumatology* **2017,** *56* (suppl_4), iv4-iv13.

21.Croset, A.; Delafosse, L.; Gaudry, J.-P.; Arod, C.; Glez, L.; Losberger, C.; Begue, D.; Krstanovic, A.; Robert, F.; Vilbois, F.; Chevalet, L.; Antonsson, B., Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells. *Journal of Biotechnology* **2012,** *161* (3), 336-348.

22.Böhm, E.; Seyfried, B. K.; Dockal, M.; Graninger, M.; Hasslacher, M.; Neurath, M.; Konetschny, C.; Matthiessen, P.; Mitterer, A.; Scheiflinger, F., Differences in N-glycosylation of recombinant human coagulation factor VII derived from BHK, CHO, and HEK293 cells. *BMC Biotechnology* **2015,** *15* (1), 87.

23.Rouiller, Y.; Solacroup, T.; Deparis, V.; Barbafieri, M.; Gleixner, R.; Broly, H.; Eon-Duval, A., Application of Quality by Design to the characterization of the cell culture process of an Fc-Fusion protein. *European Journal of Pharmaceutics and Biopharmaceutics* **2012,** *81* (2), 426-437.

24.Bee, J. S.; Randolph, T. W.; Carpenter, J. F.; Bishop, S. M.; Dimitrova, M. N., Effects of Surfaces and Leachables on the Stability of Biopharmaceuticals. *J Pharm Sci-Us* **2011,** *100* (10), 4158-4170.

25.Ejima, D.; Yumioka, R.; Tsumoto, K.; Arakawa, T., Effective elution of antibodies by arginine and arginine derivatives in affinity column chromatography. *Anal.Biochem.* **2005,** *345* (2), 250-257.

26.Cromwell, M. E. M.; Hilario, E.; Jacobson, F., Protein aggregation and bioprocessing. *The AAPS Journal* **2006,** *8* (3), E572-E579.

27.Freshney, I. R., *Culture of Animal Cells - A manual of Basic Technique*. 5th ed.; John Wiley & Sons: 2005.

28.Shire, S. J., Formulation and manufacturability of biologics. *Current opinion in biotechnology* **2009,** *20* (6), 708.

29.Uchiyama, S., Liquid formulation for antibody drugs. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2014,** *1844* (11), 2041-2052.

30.Chi, E. Y.; Krishnan, S.; Kendrick, B. S.; Chang, B. S.; Carpenter, J. F.; Randolph, T. W., Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor. *Protein Science* **2003,** *12* (5), 903-913.

31.Galm, L.; Amrhein, S.; Hubbuch, J., Predictive approach for protein aggregation: Correlation of protein surface characteristics and conformational flexibility to protein aggregation propensity. *Biotechnology and Bioengineering* **2017,** *114* (6), 1170-1183.

32.Kuhn, A. B.; Kube, S.; Karow-Zwick, A. R.; Seeliger, D.; Garidel, P.; Blech, M.; Schäfer, L. V., Improved Solution-State Properties of Monoclonal Antibodies by Targeted Mutations. *The Journal of Physical Chemistry B* **2017,** *121* (48), 10818-10827.

33.Thakur, A.; Huang, M.; Lum, L. G., Bispecific antibody based therapeutics: Strengths and challenges. *Blood Reviews* **2018**.

34.Cicerone, M. T.; Pikal, M. J.; Qian, K. K., Stabilization of proteins in solid form. *Advanced Drug Delivery Reviews* **2015,** *93*, 14-24.

35.Drazic, A.; Winter, J., The physiological role of reversible methionine oxidation. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2014,** *1844* (8), 1367-1382.

36.Wang, W.; Vlasak, J.; Li, Y.; Pristatsky, P.; Fang, Y.; Pittman, T.; Roman, J.; Wang, Y.; Prueksaritanont, T.; Ionescu, R., Impact of methionine oxidation in human IgG1 Fc on serum half-life of monoclonal antibodies. *Molecular Immunology* **2011,** *48* (6-7), 860-866.

37.Vlasak, J.; Bussat, M. C.; Wang, S.; Wagner-Rousset, E.; Schaefer, M.; Klinguer-Hamour, C.; Kirchmeier, M.; Corvaïa, N.; Ionescu, R.; Beck, A., Identification and characterization of asparagine deamidation in the light chain CDR1 of a humanized IgG1 antibody. *Anal.Biochem.* **2009,** *392* (2), 145-154.

38.Geiger, T.; Clarke, S., Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *Journal of Biological Chemistry* **1987,** *262* (2), 785-794.

39.Terashima, I.; Koga, A.; Nagai, H., Identification of deamidation and isomerization sites on pharmaceutical recombinant antibody using H218O. *Anal.Biochem.* **2007,** *368* (1), 49-60.

40.Liu, H.; Wang, F.; Xu, W.; May, K.; Richardson, D., Quantitation of asparagine deamidation by isotope labeling and liquid chromatography coupled with mass spectrometry analysis. *Anal.Biochem.* **2013,** *432* (1), 16-22.

41.Chelius, D.; Jing, K.; Lueras, A.; Rehder, D. S.; Dillon, T. M.; Vizel, A.; Rajan, R. S.; Li, T.; Treuheit, M. J.; Bondarenko, P. V., Formation of Pyroglutamic Acid from N-Terminal Glutamic Acid in Immunoglobulin Gamma Antibodies. *Analytical Chemistry* **2006,** *78* (7), 2370-2376.

42.Brorson, K.; Jia, A. Y., Therapeutic monoclonal antibodies and consistent ends: terminal heterogeneity, detection, and impact on quality. *Current Opinion in Biotechnology* **2014,** *30* (Supplement C), 140-146.

43.Sundaram, S.; Matathia, A.; Qian, J.; Zhang, J.; Hsieh, M.-C.; Liu, T.; Crowley, R.; Parekh, B.; Zhou, Q., An innovative approach for the characterization of the isoforms of a monoclonal antibody product. *mAbs* **2011,** *3* (6), 505-512.

44.Chelius, D.; Rehder, D. S.; Bondarenko, P. V., Identification and Characterization of Deamidation Sites in the Conserved Regions of Human Immunoglobulin Gamma Antibodies. *Analytical Chemistry* **2005,** *77* (18), 6004-6011.

45.Sandra, K.; Vandenheede, I.; Sandra, P., Modern chromatographic and mass spectrometric techniques for protein biopharmaceutical characterization. *Journal of Chromatography A* **2014,** *1335*, 81-103.

46.Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J., *Essentials of Glycobiology*. First ed.; Cold Spring Harbour Laboratory Press: 1999.

47.Chung, C. H.; Mirakhur, B.; Chan, E.; Le, Q.-T.; Berlin, J.; Morse, M.; Murphy, B. A.; Satinover, S. M.; Hosen, J.; Mauro, D.; Slebos, R. J.; Zhou, Q.; Gold, D.; Hatley, T.; Hicklin, D. J.; Platts-Mills, T. A. E., Cetuximab-Induced Anaphylaxis and IgE Specific for Galactose-α-1,3-Galactose. *New England Journal of Medicine* **2008,** *358* (11), 1109-1117.

48.Ozturk, S.; Hu, W.-S.; Ozturk, S. S.; Hu, W.-S., *Cell Culture Technology for Pharmaceutical and Cell-Based Therapies*. Chapman and Hall/CRC: Baton Rouge, UNITED STATES, 2005.

49.Raju, T. S., Terminal sugars of Fc glycans influence antibody effector functions of IgGs. *Current Opinion in Immunology* **2008,** *20* (4), 471-478.

50.Shields, R. L.; Lai, J.; Keck, R.; O'Connell, L. Y.; Hong, K.; Meng, Y. G.; Weikert, S. H. A.; Presta, L. G., Lack of Fucose on Human IgG1 N-Linked Oligosaccharide Improves Binding to Human FcγRIII and Antibody-dependent Cellular Toxicity. *Journal of Biological Chemistry* **2002,** *277* (30), 26733-26740.

51.Hodoniczky, J.; Zheng, Y. Z.; James, D. C., Control of Recombinant Monoclonal Antibody Effector Functions by Fc N-Glycan Remodeling in Vitro. *Biotechnology Progress* **2005,** *21* (6), 1644-1652.

52.Jefferis, R., Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action. *Trends in Pharmacological Sciences* **2009,** *30* (7), 356-362.

53.Zhou, Q.; Shankara, S.; Roy, A.; Qiu, H.; Estes, S.; McVie-Wylie, A.; Culm-Merdek, K.; Park, A.; Pan, C.; Edmunds, T., Development of a simple and rapid method for producing non-fucosylated oligomannose containing antibodies with

increased effector function. *Biotechnology and Bioengineering* **2008,** *99* (3), 652-665.

54.Dong, X.; Storkus, W. J.; Salter, R. D., Binding and Uptake of Agalactosyl IgG by Mannose Receptor on Macrophages and Dendritic Cells. *The Journal of Immunology* **1999,** *163* (10), 5427-5434.

55.Scallon, B. J.; Tam, S. H.; McCarthy, S. G.; Cai, A. N.; Raju, T. S., Higher levels of sialylated Fc glycans in immunoglobulin G molecules can adversely impact functionality. *Molecular Immunology* **2007,** *44* (7), 1524-1534.

56.Abès, R.; Teillaud, J.-L., Impact of Glycosylation on Effector Functions of Therapeutic IgG. *Pharmaceuticals* **2009,** *3* (1), 146-157.

57.Potocnakova, L.; Bhide, M.; Pulzova, L. B., An Introduction to B-Cell Epitope Mapping and In Silico Epitope Prediction. *Journal of Immunology Research* **2016,** *2016*, 6760830.

58.Schmidt, S., Strategies to Predict the Developability of Biopharmaceuticals. *American Pharmaceutical Review* **2017**, 122-125.

59.Fanali, S.; Haddad, P.; Lloyd, D. K.; Poole, C.; Schoenmakers, P., *Liquid Chromatography : Fundamentals and Instrumentation*. Elsevier: Saint Louis, UNITED STATES, 2013.

60.Naijun, W.; M., C. A., Fundamental and practical aspects of ultrahigh pressure liquid chromatography for fast separations. *Journal of Separation Science* **2007,** *30* (8), 1167-1182.

61.Neue, U. D., Theory of peak capacity in gradient elution. *Journal of Chromatography A* **2005,** *1079* (1), 153-161.

62.Fekete, S.; Beck, A.; Fekete, J.; Guillarme, D., Method development for the separation of monoclonal antibody charge variants in cation exchange chromatography, Part I: Salt gradient approach. *J.Pharm.Biomed.Anal.* **2015,** *102*, 33-44.

63.Tassi, M.; De Vos, J.; Chatterjee, S.; Sobott, F.; Bones, J.; Eeltink, S., Advances in native high-performance liquid chromatography and intact mass spectrometry for the characterization of biopharmaceutical products. *J Sep Sci* **2017,** *41* (1), 125-144.

64.Fekete, S.; Guillarme, D., Ultra-high-performance liquid chromatography for the characterization of therapeutic proteins. *TrAC Trends in Analytical Chemistry* **2014,** *63* (Supplement C), 76-84.

65.Fekete, S.; Beck, A.; Veuthey, J. L.; Guillarme, D., Ion-exchange chromatography for the characterization of biopharmaceuticals. *J Pharm Biomed Anal* **2015,** *113*, 43-55.

66.Fekete, S.; Beck, A.; Fekete, J.; Guillarme, D., Method development for the separation of monoclonal antibody charge variants in cation exchange chromatography, Part II: pH gradient approach. *J.Pharm.Biomed.Anal.* **2015,** *102*, 282-289.

67.Farnan, D.; Moreno, G. T., Multiproduct High-Resolution Monoclonal Antibody Charge Variant Separations by pH Gradient Ion-Exchange Chromatography. *Analytical Chemistry* **2009,** *81* (21), 8846-8857.

68.Yansheng, L.; Anderson, D. J., Gradient chromatofocusing high-performance liquid chromatography II. Theoretical aspects. *Journal of Chromatography A* **1997,** *762* (1), 47-54.

69.Fussl, F.; Cook, K.; Scheffler, K.; Farrell, A.; Mittermayr, S.; Bones, J., Charge Variant Analysis of Monoclonal Antibodies using Direct Coupled pH Gradient Cation Exchange Chromatography to High Resolution Native Mass Spectrometry. *Anal Chem* **2018**.

70.Paul-Dauphin, S.; Karaca, F.; Morgan, T. J.; Millan-Agorio, M.; Herod, A. A.; Kandiyoti, R., Probing Size Exclusion Mechanisms of Complex Hydrocarbon Mixtures: The Effect of Altering Eluent Compositions. *Energy & Fuels* **2007,** *21* (6), 3484-3489.

71.Liu, H.; Gaza-Bulseco, G.; Chumsae, C., Analysis of Reduced Monoclonal Antibodies Using Size Exclusion Chromatography Coupled with Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2009,** *20* (12), 2258-2264.

72.van der Kant, R.; Karow-Zwick, A. R.; Van Durme, J.; Blech, M.; Gallardo, R.; Seeliger, D.; Aßfalg, K.; Baatsen, P.; Compernolle, G.; Gils, A.; Studts, J. M.; Schulz, P.; Garidel, P.; Schymkowitz, J.; Rousseau, F., Prediction and Reduction of the Aggregation of Monoclonal Antibodies. *Journal of Molecular Biology* **2017,** *429* (8), 1244-1261.

73.Paul, A. J.; Schwab, K.; Hesse, F., Direct analysis of mAb aggregates in mammalian cell culture supernatant. *BMC Biotechnology* **2014,** *14*, 99.

74.Hong, P.; Koza, S.; Bouvier, E. S. P., A Review Size Exclusion Chromatography for the Analysis of Protein Biotherapeutics and Their Aggregates. *Journal of Liquid Chromatography & Related Technologies* **2012,** *35* (20), 2923-2950.

75.Wagner-Rousset, E.; Janin-Bussat, M.-C.; Colas, O.; Excoffier, M.; Ayoub, D.; Haeuw, J.-F.; Rilatt, I.; Perez, M.; Corvaïa, N.; Beck, A., Antibody-drug conjugate model fast characterization by LC-MS following IdeS proteolytic digestion. *mAbs* **2014,** *6* (1), 173-184.

76.Fekete, S.; Beck, A.; Veuthey, J.-L.; Guillarme, D., Theory and practice of size exclusion chromatography for the analysis of protein aggregates. *J.Pharm.Biomed.Anal.* **2014,** *101*, 161-173.

77.Kükrer, B.; Filipe, V.; van Duijn, E.; Kasper, P. T.; Vreeken, R. J.; Heck, A. J. R.; Jiskoot, W., Mass Spectrometric Analysis of Intact Human Monoclonal Antibody Aggregates Fractionated by Size-Exclusion Chromatography. *Pharmaceutical research* **2010,** *27* (10), 2197-2204.

78.Shire, S. J., 2 - Analytical tools used in the formulation and assessment of stability of monoclonal antibodies (mAbs). In *Monoclonal Antibodies*, Woodhead Publishing: 2015; pp 17-44.

79.Astefanei, A.; Dapic, I.; Camenzuli, M., Different Stationary Phase Selectivities and Morphologies for Intact Protein Separations. *Chromatographia* **2017,** *80* (5), 665-687.

80.Fekete, S.; Rudaz, S.; Fekete, J.; Guillarme, D., Analysis of recombinant monoclonal antibodies by RPLC: Toward a generic method development approach. *J.Pharm.Biomed.Anal.* **2012,** *70*, 158-168.

81.. Reverse Phase Chromatography *Theory of HPLC* [Online], Accessed 2018.

82.Zhang, Q.; Yang, F. Q.; Ge, L.; Hu, Y. J.; Xia, Z. N., Recent applications of hydrophilic interaction liquid chromatography in pharmaceutical analysis. *Journal of Separation Science* **2017,** *40* (1), 49-80.

83.Grumbach, E. S.; Wagrowski-Diehl, D. M.; Mazzeo, J. R.; Iraneta, P. C. Hydrophilic Interaction Chromatography Using Silica Columns for the Retention of Polar Analytes and Enhanced ESI-MS Sensitivity *LCGC* [Online], 2004.

84.Buszewski, B.; Noga, S., Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. *Analytical and Bioanalytical Chemistry* **2012,** *402* (1), 231-247.

85.Lam, M. P. Y.; Siu, S. O.; Lau, E.; Mao, X.; Sun, H. Z.; Chiu, P. C. N.; Yeung, W. S. B.; Cox, D. M.; Chu, I. K., Online coupling of reverse-phase and hydrophilic interaction liquid chromatography for protein and glycoprotein characterization. *Analytical and Bioanalytical Chemistry* **2010,** *398* (2), 791-804.

86.Wuhrer, M.; Boer, A. R. d.; Deelder, A. M., Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry. *Mass Spectrometry Reviews* **2009,** *28* (2), 192-206.

87.Pace, C. N.; Fu, H.; Fryar, K. L.; Landua, J.; Trevino, S. R.; Shirley, B. A.; Hendricks, M. M.; Iimura, S.; Gajiwala, K.; Scholtz, J. M.; Grimsley, G. R., Contribution of Hydrophobic Interactions to Protein Stability. *Journal of Molecular Biology* **2011,** *408* (3), 514-528.

88.Hoffmann, E. d.; Stroobant, V., *Mass Spectrometry: Principles and Applications*. Third ed.; John Wiley & Sons: 2007.

89.Morris, H. R., Research on Peptides and Glycopeptides. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **1979,** *293* (1400), 39-51.

90.Morris, H. R.; Geddes, A. J.; Graham, G. N., Some problems associated with the amino acid-sequence analysis of proteins by mass spectrometry. *Biochemical Journal* **1969,** *111* (5), 38P.

91.Morris, H. R.; Dell, A.; Etienne, A. T.; Judkins, M.; McDowell, R. A.; Panico, M.; Taylor, G. W., New methods for the characterisation of biopolymers. In *Pure and Applied Chemistry*, 1982; Vol. 54, p 267.

92.Karas, M.; Hillenkamp, F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry* **1988,** *60* (20), 2299-2301.

93.Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C., Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989,** *246* (4926), 64-71.

94.Rosell-Llompart, J.; Grifoll, J.; Loscertales, I. G., Electrosprays in the cone-jet mode: from Taylor cone formation to spray development. *Journal of Aerosol Science* **2018**.

95.Gomez, A.; Tang, K., Charge and fission of droplets in electrostatic sprays. *Physics of Fluids* **1994,** *6* (1), 404-414.

96.Wilm, M., Principles of Electrospray Ionization. *Molecular & Cellular Proteomics : MCP* **2011,** *10* (7), M111.009407.

97.Kebarle, P.; Verkerk, U. H., A Brief Overview of the Mechanisms Involved in Electrospray Mass Spectrometry. In *Reactive Intermediates*, Santos, L. S., Ed. 2010.

98.Heck, A. J. R., Native mass spectrometry: A bridge between interactomics and structural biology. *Nature Methods* **2008,** *5* (11), 927-933.

99.Heck, A. J. R., Native mass spectrometry: a bridge between interactomics and structural biology. *Nature Methods* **2008,** *5*, 927.

100.Downard, K., The mass spectrometer. In *Mass Spectrometry: A Foundation Course*, Downard, K., Ed. The Royal Society of Chemistry: 2004; pp 22-66.

101.Bier, M. E., Coupling ESI and MALDI Sources to the Quadrupole Mass Filter, Quadrupole Ion Trap, Linear Quadrupole Ion Trap, and Orbitrap Mass Analyzers. In *Electrospray and MALDI Mass Spectrometry*, Cole, R. E., Ed. 2012.

102.Kingdon, K. H., A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures. *Physical Review* **1923,** *21* (4), 408-418.

103.Kalli, A.; Smith, G. T.; Sweredoski, M. J.; Hess, S., Evaluation and Optimization of Mass Spectrometric Settings during Data-Dependent Acquisition Mode: Focus on LTQ-Orbitrap Mass Analyzers. *Journal of proteome research* **2013,** *12* (7), 3071-3086.

104.Eliuk, S.; Makarov, A., Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annual Review of Analytical Chemistry* **2015,** *8* (1), 61-80.

105.Comisarow, M. B.; Marshall, A. G., Fourier transform ion cyclotron resonance spectroscopy. *Chemical Physics Letters* **1974,** *25* (2), 282-283.

106.Marshall, A. G.; Chen, T., 40 years of Fourier transform ion cyclotron resonance mass spectrometry. *International Journal of Mass Spectrometry* **2015,** *377*, 410-420.

107.Doerr, A., DIA mass spectrometry. *Nature Methods* **2015,** *12*, 35.

108.Edmond, d. H., Tandem mass spectrometry: A primer. *Journal of Mass Spectrometry* **1996,** *31* (2), 129-137.

109.Krasny, L.; Bland, P.; Kogata, N.; Wai, P.; Howard, B. A.; Natrajan, R. C.; Huang, P. H., SWATH mass spectrometry as a tool for quantitative profiling of the matrisome. *Journal of Proteomics* **2018**.

110.Egertson, J. D.; Kuehn, A.; Merrihew, G. E.; Bateman, N. W.; MacLean, B. X.; Ting, Y. S.; Canterbury, J. D.; Marsh, D. M.; Kellmann, M.; Zabrouskov, V.; Wu, C. C.; MacCoss, M. J., Multiplexed MS/MS for improved data-independent acquisition. *Nature Methods* **2013,** *10*, 744.

111.Bernier, L.; Pinfold, H.; Pauly, M.; Rauschenbach, S.; Reiss, J., Gas Flow and Ion Transfer in Heated ESI Capillary Interfaces. *Journal of The American Society for Mass Spectrometry* **2018,** *29* (4), 761-773.

112.Lu, W.; Bennett, B. D.; Rabinowitz, J. D., Analytical strategies for LC-MS-based targeted metabolomics. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2008,** *871* (2), 236-242.

113.Scheltema, R. A.; Hauschild, J.-P.; Lange, O.; Hornburg, D.; Denisov, E.; Damoc, E.; Kuehn, A.; Makarov, A.; Mann, M., The Q Exactive HF, a Benchtop Mass Spectrometer with a Pre-filter, High-performance Quadrupole and an Ultra-high-field Orbitrap Analyzer. *Molecular & Cellular Proteomics : MCP* **2014,** *13* (12), 3698-3708.

114.Michalski, A.; Damoc, E.; Hauschild, J.-P.; Lange, O.; Wieghaus, A.; Makarov, A.; Nagaraj, N.; Cox, J.; Mann, M.; Horning, S., Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics* **2011,** *10* (9).

115.Diedrich, J. K.; Pinto, A. F. M.; Yates, J. R., Energy dependence of HCD on peptide fragmentation: Stepped collisional energy finds the sweet spot. *Journal of the American Society for Mass Spectrometry* **2013,** *24* (11), 10.1007/s13361-013-0709-7.

116.Bin, M.; Kaizhong, Z.; Christopher, H.; Chengzhi, L.; Ming, L.; Amanda, D. K.; Gilles, L., PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Communications in Mass Spectrometry* **2003,** *17* (20), 2337-2342.

117.Bern, M.; Kil, Y. J.; Becker, C., Byonic: Advanced Peptide and Protein Identification Software. *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]* **2012,** *CHAPTER*, Unit13.20-Unit13.20.

118.Farrell, A.; Carillo, S.; Scheffler, K.; Cook, K.; Bones, J., Monoclonal antibody sequence assessment using a hybrid quadrupole-Orbitrap mass spectrometer. *Analytical Methods* **2018**.

119.Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A.; Haslam, S. M., GlycoWorkbench: A Tool for the Computer-Assisted Annotation of Mass Spectra of Glycans. *Journal of Proteome Research* **2008,** *7* (4), 1650-1659.

120.Meitei, N. S.; Apte, A.; Snovida, S. I.; Rogers, J. C.; Saba, J., Automating mass spectrometry-based quantitative glycomics using aminoxy tandem mass tag reagents with SimGlycan. *Journal of Proteomics* **2015,** *127*, 211-222.

Chapter 2.

2.Rapid Charge Variant Analysis of Monoclonal Antibodies to Support Lead Candidate Biopharmaceutical Development

2.1.Introduction

Over the next five years, five of the top 10 best-selling pharmaceuticals[1], of both small molecule and biopharmaceutical markets, will see their patent protection expire[2,3,4]. Some of these molecules are already off patent in the EU and biosimilars are beginning to emerge[4]. In 2016, many of these molecules including; adalimumab, bevacizumab, trastuzumab, rituximab and infliximab, had collective sales revenue of almost $46 billion[1]. The cost of these life changing treatments to international health services is a widely discussed topic with increasing pressure for cheaper alternatives, emerging as a result of population demand for these medicines. This high demand and lucrative potential for manufacturers means more companies are developing their own versions of blockbuster therapies. The requirements to achieve 'biosimilarity' have been outlined by the International Council for Harmonization (ICH) and regulatory agencies have adapted these guidelines to their own approval processes[5,6]. The process of proving similarity involves extensive comparability tests. In order for a candidate molecule to be considered a biosimilar it must show comparable "physicochemical properties, biological activity, immunochemical properties, purity and impurities" to the originator product [7,8].

Certain charge variants (CV) of a monoclonal antibody (mAb) are important critical quality attributes (CQA's) which have to be monitored during comparability studies[9,10]. They are significantly related to its physicochemical, immunochemical properties and its biological activity[11]. This heterogeneity results from mAb isoforms with varying net charge caused by enzymatic and non-enzymatic processes, within the cell and following secretion into the culture medium[10]. One of the most common

modifications contributing to charge heterogeneity is C-terminal lysine clipping, which results in the loss of one or two positively charged lysine residues leading to the formation of basic variants[12]. Deamidation, glycation and the presence of negatively charged sialic acids on N- and O-linked glycans all lead to an increased negative charge and the appearance of acidic species[12,9,11].

Cation exchange chromatography (CEX) and capillary isoelectric focusing (cIEF) are routinely used for the determination of the CV profile[13,14,15,16,17,18]. In CEX, mAb variants are separated based on differences in the net charge leading to varying interactions with the opposing weak (WCX) or strong (SCX) ion exchange functionalities on the stationary phase[19]. Within SCX, two modes of elution, salt and pH gradient elution, are widely used[20]. A salt gradient progresses from a starting buffer of low ionic strength which increases over time, while the pH of the buffer is kept constant. A pH gradient starts with a buffer of low pH which increases over time with the ionic strength kept constant[21,22]. Salt gradient elution is considered as the classic elution mode, whereas, pH gradient elution has been described as the chromatographic analogy to IEF, using changes in the mobile phase pH to create a gradient along the column[23,21]. Relative to salt elution mode, pH elution has shown to provide better resolution and is adaptable for rapid analytical methods due to the separation mechanisms, apparent independence from column length and chemistry [24,25,26,27].

The CV profile of a mAb, as an indicator for the presence and abundance of CQA's, is used to monitor the stability of the production process[10]. Minor alterations in processing can cause an evident change in isoform distribution[28]. Therefore, the determination, optimisation and monitoring of a mAbs CV profile is important at all stages of the products lifecycle, from development through to manufacturing of the therapeutic drug substance and stability assessment of the formulated drug product. CEX was recently incorporated as a process analytical technology (PAT)

monitoring tool in a continuous biological production process using salt based elution with buffers of different pH[28]. An important requirement to facilitate the incorporation into PAT workflows and indeed, to support lead candidate development and clonal selection workflows, is the need for rapid CEX methods that offer high throughput capabilities however, without compromising resolution. Rapid CEX methods for the separation of mAb charge variants have been investigated, predominantly using linear and non-linear salt gradient based elution[29,30].

In this chapter, salt and pH elution SCX chromatography were compared and the optimised separation conditions, for each, were used to develop a rapid CV screening protocol for mAbs. The ultrafast (< 5 minutes) pH gradient method was chosen for validation and further application in the analysis of mAbs with a range of pI. The suitability of the method for in-process testing was investigated by repeat injections (n=45) of an in-house produced adalimumab variant.

The validated method was then applied as a molecular triage method for the rapid charge variant profiling of in-house trastuzumab variants to assess differences in the charge variant patterns between mAbs expressed transiently in either Chinese hamster ovary (CHO) or human embryonic kidney (HEK 293) cells.

Chapter 2.

2.2.Materials & Methods

2.2.1. Reagents

Acetone HPLC grade, acetonitrile LC-MS grade, agar plates, agarose, ammonium hydroxide, bottle top filters 0.22 µm & 0.45 µm, CX-1 pH gradient buffers, dimethyl sulfoxide (DMSO), erlenmeyer flasks, ethanol, ethidium bromide, formic acid, formic acid (0.1% v/v) in water LC-MS grade, glacial acetic acid, isopropanol (IPA), loops for microbiology, Oakridge tubes, sodium chloride, L-shaped spreaders, Millipore Ultrafree-MC sterile centrifugal filters, polyethylenimine, valproic acid, lithium acetate, Super Optimal Broth Medium (SOB), tris-borate-EDTA (TBE) water LC-MS grade were all purchased from ThermoFisher Scientific (Dublin, Ireland).

Citric acid monohydrate, (N-Morpholino)ethanesulfonic acid (MES), Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), caffeine and trisodium citrate dihydrate were purchased from Sigma Aldrich (Wicklow, Ireland).

Promega sequencing grade modified Trypsin was purchased from MyBio (Kilkenny, Ireland). Origene PowerPrep HP Plasmid Purification System purchased from Insight Biotechnology Ltd. (Middlesex, United Kingdom). pFUSE plasmids, Blasticidin and Zeocin, liquid and agar, were purchased from InvivoGen (Toulouse, France). Life Technologies Expi293 and ExpiCHO expression system kits, Trypan blue, Countess™ slides, GeneAmp™ dNTP Blend, MAX EfficiencyDH5α chemically competent cells and Platinum Taq DNA Polymerase High Fidelity kit were purchased from Bio-Sciences Ltd. (Dublin, Ireland). QIAquick PCR Purification Kit and QIAquick Gel Extraction Kit purchased from Qiagen (Manchester, United Kingdom). MycoTool mycoplasma real-time PCR chain reaction kit purchased from Roche Diagnostics (Dublin, Ireland) Fabricator (IdeS) enzyme was purchased from Genovis (Lund, Sweden).

Chapter 2.

The NIST Monoclonal Antibody Reference material was obtained from the National Institute of Standards & Technology (Gaithersburg, MD, USA). Trastuzumab drug product was kindly provided by the Hospital Pharmacy Unit of the University Hospital of San Cecilio, Granada, Spain.

2.2.2. V-gene amplification, restriction, ligation and transformation

The trastuzumab and adalimumab protein produced in this project contained the CDR sequence of the reference product on a human IgG1 frame.

2.2.2.1.Polymerase Chain Reaction (PCR)

Table 2.2.1: Primer sequence, Tm, GC content and annealing temperature for adalimumab encoding V-genes

Antibody	Primer For-Rev (5'-3')	Tm (°C)	GC%	Ann. Temp
Adalimumab Heavy Chain V-gene	For-GCACTTGTCACGAATTCGGAAGTTCAGCTGGTT	71.1	48.48	66.1
	Rev-TGGGCCCTTGGTGCTAGCGGAGGACACGGTAAC	79.2	63.64	
Adalimumab Light Chain V-gene	For-GCACTTGTCACGAATTCGGAAGTTCAGCTGGTT	60	48.48	55
	Rev-TGGTGCAGCCACCGTACGCTTGATTTCAACCTTGGTGCCC	78.9	57.5	

Table 2.2.2: Primer sequence, Tm, GC content and annealing temperature for trastuzumab encoding V-genes

Antibody	Primer For-Rev (5'-3')	Tm (°C)	GC%	Ann. Temp
Trastuzumab Heavy Chain V-gene	For-GCACTTGTCACGAATTCGGAGGTGCAGCTGGTG	75.2	57.58	66.1
	Rev-TGGGCCCTTGGTGCTAGCGCTAGACACGGTGAC	79.0	63.64	
Trastuzumab Light Chain V-gene	For-GCACTTGTCACGAATTCGGACATCCAGATGACCCAGTC	73.7	52.63	72
	Rev-TGGTGCAGCCACCGTACGCTTGATTTCCACCTTGGTGCC	79.3	58.97	

Platinum Taq DNA Polymerase High Fidelity was used to perform DNA amplification. Table 2.2.3 shows the constituents and appropriate concentrations of each component of the PCR master mix, volumes shown are per 50 µL PCR reaction. Primer sequences were not added to the reaction tube until immediately before placing in the PCR cycler.

Table 2.2.3: PCR reagents and relative concentrations

Reagent	Concentration
High Purity Water	To 50 µL
PCR Buffer 10X	5 µL
10 mM dNTP	0.2 mM
50 mM $MgSO_4$	2 mM
***Taq* Polymerase (5 U/ µL)**	1 U (0.2 µL)
Forward Primer	1 µM
Reverse Primer	1 µM
Template DNA	<500 ng
Total Volume	50 µL

All PCR amplifications were performed in an Eppendorf MasterCycler EP S. The temperatures and duration of each step is shown in Table 2.2.4 below. Tm for the VH and VL chains were determined using the Thermo Scientific Tm calculator.

Table 2.2.4: Thermocycler program

Step		Temperature (°C)	Time
Initial Denaturation		94	30 seconds
25 - 35 PCR Cycles	Denaturation	94	15 seconds
	Anneal	See table 2.2.1 & 2.2.2	30 seconds
	Extend	68	1min/kb
Hold		4	∞

V-genes were cleaned using a QIAquick PCR Purification Kit followed by gel purification. A 2% agarose gel was used for the v-genes while the pFUSE vectors were purified in 0.8% agarose gels. The agarose gel was made by dissolving the appropriate amount of agarose in 100 mL TBE buffer with 0.5 µg/mL ethidium bromide. The gel was run in TBE buffer at 120 V for 90 minutes. V-genes and

vectors were cut from the gel and subsequently purified using QIAquick Gel Extraction Kit.

2.2.2.2. Restriction enzyme digestion

Table 2.2.5: Restriction enzyme selection

Sample	Restriction Enzyme
VH	EcoRI & NheI
VL-κ	EcoRI & BsiWI
pFUSE CHIg-hG1	EcoRI & NheI
pFUSE CLIg-κ	EcoRI & BsiWI

Sequential endonuclease digests were performed as EcoRI exhibits star activity when used with incompatible buffers. A 3 µg quantity of each DNA sequence (VL and VH) and 6 µg of each plasmid vector (VL and VH) was digested with EcoRI and one other restriction enzyme. The total digestion volume was 100 µL. The pFUSE vectors and DNA were digested with EcoRI and NEBuffer U at 37°C for 2 hours, the enzyme was then heat denatured at 65°C for 20 minutes. A 50 µL volume of the first digestion was removed and added to a new PCR tube. NEBuffer 2.1 and Nhe1 restriction enzyme were added to the VH plasmids and VH DNA. These digests were then incubated at 37°C for 2 hours with a heat denaturation of 65°C for 20 minutes. The VL DNA and plasmids were digested in NEBuffer 3.1 with BsiWI restriction enzyme at 55°C for 2 hours followed by heat inactivation at 65°C for 20 minutes. The samples were then agarose gel purified and subsequently extracted.

2.2.2.3. Ligation

Ligation was performed using NEB T4 DNA ligase. The DNA insert to plasmid vector ratio was 1:3. The ligation mixture was incubated at 16°C overnight and the following day the enzyme was heat inactivated by incubating at 65°C for 10 minutes.

2.2.2.4.Transformation

MAX EfficiencyDH5α chemically competent cells were used, genotype: F-Φ80lacZΔM15 Δ (lacZYA-argF) U169 recA1 endA1 hsdR17 (rk-, mk+) phoA supE44 λ-thi-1 gyrA96 relA1. The thermal shock method was used to insert the vector DNA into the competent cells. Prior to this the ligation reaction was diluted to 10 ng DNA/µL and 1 µL of ligated plasmid was added to the competent cells. The cells were incubated on ice for 30 minutes then removed to a water bath for heat shock at 42°C for 45 seconds and then returned to ice for 2 minutes. The cells were then grown in SOB media for 1 hour at 37°C. Following this, the cells were centrifuged and resuspended in a small volume for spreading on a selective agar plate and incubating overnight at 37°C. A plate of single colonies was chosen from the transformation plate and subsequently sent for sequencing.

2.2.3. Mammalian Cell Culture

2.2.3.1. Aseptic Technique

Howie style lab coat with disposable sleeve covers and gloves were worn at all times when working in the mammalian cell culture laboratory. All mammalian cell culture work was undertaken in a Bioair Safemate 1.2 biosafety level II laminar flow hood. Prior to use the hood air flow was equilibrated for fifteen minutes, to ensure all particulates had settled. The hood was then sprayed with 70% IPA and wiped down. All experimental equipment was sprayed with 70% IPA before being introduced in to the hood. To prevent cross contamination, one cell line was used in the laminar flow hood at a time. For work with different cell lines, fifteen minutes of air flow re-equilibration was given and the hood and associated equipment was decontaminated with 70% IPA prior to introduction of subsequent cell lines. After all cell work was complete the laminar flow hood was sprayed with 70% IPA and wiped dry. The hood was then UV irradiated for sixty minutes. Full decontamination, with 1% Virkon, of cell culture equipment was carried out on a weekly basis. Cell

lines were checked routinely for mycoplasma using a Roche MycoTool mycoplasma real-time PCR chain reaction kit with MagNa Pure LC2.0 and the Roche Lightcycler 480. No mycoplasma contamination was detected.

2.2.3.2. Cell Counting

All cell counting and viability was performed using the Invitrogen Countess Automated Cell Counter. A 10 µL culture sample was removed from the cell culture flask, aliquoted into a 1 mL, autoclaved, Eppendorf tube and removed to the work bench. Trypan blue was added to the Eppendorf in a 1:1 ratio. The sample/Trypan blue mix was incubated at room temperature for approximately 1 minute and 10 µL was applied to a Countess slide. The slide was inserted in to the instrument for count and viability determination. If cell density was too high to accurately determine the cell density the sample was diluted with fresh media.

2.2.3.3.Thawing, Routine Maintenance and Cryopreservation of HEK and CHO cell lines

All mammalian cultures were antibiotic and serum-free. The amount of media required for culturing/cryopreservation was calculated based on cell density and work to be carried out. Media was removed from refrigeration, sprayed with 70% IPA and placed in the laminar flow hood. An appropriately sized Erlenmeyer flask was sprayed with 70% IPA, moved to the laminar flow hood and labelled with cell name, date, passage number, analyst name and flask number (optional). The media was then aseptically added to a labelled Erlenmeyer flask using a sterile pipette. The flask was then placed in the Kuhner shaker incubator for approximately one hour to allow for temperature and CO2 equilibration.

2.2.3.3.1.Thawing of banked cell lines

The master and working cell banks are stored at -196°C in liquid nitrogen dewars. Prior to removal of the vial from storage, 10 mL of expression media was warmed and placed in a 50 mL Falcon tube. Upon removal of the vial from liquid nitrogen it was rapidly defrosted in a 37°C water bath. The entire contents of the cryovials

were then transferred to the 10 mL pre-warmed media and centrifuged to remove the toxic DMSO. The waste media was removed and the cell pellet was re-suspended in fresh media and transferred to a 125 mL Erlenmeyer flask containing pre-warmed media for a final culture volume of 30 mL. The cells were incubated in 8% CO_2 rotating at 125 rpm in 37°C at 80% humidity. Cell viability and density were determined 18 hours post thaw to ensure culture health.

2.2.3.3.2.Cryopreservation

Once cells had been passaged >three times, since initial defrosting, and had a density (4-6 × 10^6 cells/mL CHO, 3-5 × 10^6 cells/mL HEK) of and good viability (>95%), as per manufacturer's instructions, a cell bank was created. Freezing media was made up of 90% expression medium and 10% DMSO. Cells being banked were centrifuged and the pellet resuspended in the appropriate amount of freezing media to have a final density of 1 × 10^7 cells/mL in 1 mL of freezing media. After the cells were aliquoted into 1 mL cryovials they were placed in a Mr. Frosty freezing system which allowed a -1°C decrease in temperature per hour. The cells were initially placed in -20°C for 1-2 hours and then transferred to -80°C for 8-12 hours. Following this then the cells were moved to permanent storage in a liquid nitrogen dewar (-196°C).

2.2.3.3.3.Routine Maintenance

Expi293-F (HEK-293) and ExpiCHO-S (CHO-S) cells were maintained at 37°C with 80% humidity and 8% CO_2 shaking at 125 rpm. Subculturing took place approximately every 3-4 days or when cultures reached a certain density; 1-2 × 10^6 cells/mL for HEK 293 and 4-6 × 10^6 cells/mL for CHO-S, as per the manufacturer's guidelines. A final density after subculturing was maintained at 0.2 - 0.3 × 10^6 cells/mL for CHO-S and 0.3 - 0.5 × 10^6 cells/mL for HEK 293 ensure the cells were in the log-phase of growth. The amount of culture required to achieve a final density as stated above was removed from the flask. If the ratio of culture to new media

was less than 1:2 the culture volume was centrifuged at 182 x *g* for seven minutes. The cell pellet was resuspended and added to a new labelled Erlenmeyer flask with fresh media.

2.2.4. Microbial Cell Culture

All microbial work was carried out in an ADS Laminaire Optimale 9 biosafety level II laminar flow hood.

2.2.4.1.Selective Antibiotic Terrific Broth (TB) Media Preparation

Competent DH5α *E.coli* transformed with heavy chain plasmids were amplified using Zeocin antibiotic resistance selectivity. Trastuzumab and adalimumab light chain encoding plasmid clones were amplified using Blasticidin resistance. The media was prepared as per manufacturer's guidelines. For 200 mL, one sachet of Blas TB media was emptied into a clean, autoclaved 500 mL Duran bottle and 200 mL deionised water was added. The mix was then heated by microwave at 800 W for approximately 4 minutes until media was homogenous. The same procedure was repeated for Zeo TB media. Media was prepared on the day of use and stored in a water bath at 37°C until required.

2.2.4.2.Selective Antibiotic Lysogeny Broth (LB) Agar Preparation

LB agar was prepared as per the manufacturer's guidelines. For 200 mL, one sachet media was emptied into a clean, autoclaved 500 mL Duran bottle and 200 mL deionised water was added. The mix was then heated by microwave at 800 W for approximately 4 minutes until media was homogenous. Following preparation, the agar was carefully poured into petri dishes, ensuring no air bubbles formed, and allowed to set. Any excess plates were stored at 4°C for a maximum of three weeks. Plates to be used for culturing were labelled with mAb name, mAb region (VH/VL), date and analyst name.

2.2.4.3.DH5α cell culture

Required glycerol stock cultures were removed from -80°C freezer and placed on ice. The laminar flow hood was sprayed with 70% IPA and wiped dry. All equipment was sprayed with 70% IPA and placed in hood prior to use.

Using a pipette tip, a portion of the frozen stock was scraped off and transferred to the appropriate agar plate. The culture was then spread around the agar, the cover put in place and the plate turned upside down. The cultures were then placed in an incubator at 37°C and left to grow overnight. The following day, a secondary streak plate was created from colonies grown on the primary spread plate. This was to ensure proper clonal selection and healthy growth for inoculating into liquid media. For this, a colony, exhibiting good growth, was removed using a sterile loop and streaked in a zig zag fashion across the agar. The plate was then replaced to the incubator and left to grow overnight.

For broth cultures of 100 mL, a single colony from the secondary plate were removed using a disposable loop and the appropriate antibiotic containing broth was inoculated. The inoculated media was left overnight in Kuhner shaker 250 rpm at 37°C. Before harvest, an aliquot of culture was taken and measured in a UV-Vis spectrophotometer to ensure an OD of 2-3 at A600 was achieved. The culture broth was transferred from the shaker flask to a 50 mL centrifuge tube. The bacterial culture was then centrifuged at 2916 x *g* for 10 minutes. The waste media was removed and the cell culture pellet was resuspended in resuspension buffer, containing RNase A (as part of the Origene Maxiprep kit).

2.2.5. Plasmid DNA Purification

The DNA purification method employed with the MaxiPrep kit uses anion exchange resin to purify plasmid DNA from *E.coli* cell lysate. The anion exchange resin is equilibrated with proprietary equilibration buffer prior to loading.

The re-suspended *E.coli* cell pellet was added to an Oakridge centrifuge tube and 10 mL of cell lysis buffer was added to the suspension. The tube was then gently inverted and the suspension incubated for five minutes at room temperature. Following incubation 10 mL of neutralization buffer was added and mixed gently by inverting the tube until the solution had clearly separated into lysate and a straw coloured liquid. The lysate was centrifuged at 18,514 x *g* at room temperature for 10 minutes. The supernatant was pipetted on to the equilibrated column ensuring not to dislodge the lysate pellet. The supernatant was left to drain by gravity and the flow through discarded. The column was washed with proprietary wash buffer. A 10.5 mL volume of Isopropanol was added to a clean OakRidge tube and placed on ice. Once the wash buffer had flowed through, the column was attached to the Oakridge tube on ice. A 15 mL volume of elution buffer was added to the column and the DNA precipitated in the IPA. The DNA solution was centrifuged at 18,514 x *g* at 4°C for 30 mins. The supernatant was carefully discarded and the DNA pellet washed with 70% ethanol. The DNA was centrifuged 18,514 x *g* at 4°C for 5 mins. The ethanol was removed and the pellet left to air dry. Once dry the DNA was dissolved in 500 µL of TE buffer and transferred to an autoclaved 1 mL Eppendorf tube. The DNA concentration was then determined using a NanoDrop UV-Vis Spectrophotometer at a wavelength of 260 nm. The purity of the DNA was determined by comparing the ratio of absorbance at 260/280 nm, for RNA contamination and 260/230 nm, for organic compound contamination. A 260nm/280 nm ratio of 1.8 AU was considered ideal, values greater than this are suggestive of RNA contamination. For organic contamination a ratio between 1.5-1.8 was considered ideal.

2.2.6. Transient Transfection

The ratio of culture to DNA (mL:µg) for transfection was 1:1. For example, a 50 mL culture required 50 µg (25 µg VH & 25 µg VL) DNA. DNA was sterile filtered using

Millipore Ultrafree-MC Sterile Centrifugal Filter device and concentration was determined using a NanoDrop 2000 spectrophotometer, prior to transfection.

2.2.6.1.Expi293 Transfection

On the day prior to transfection (Day-1) cells were sub-cultured to a density of 1.5 × 10^6 – 2 × 10^6 cells/mL. A cell density, at the time of transfection, was required to be 2.5 × 10^6 cells/mL with >95% viability.

Optimem I and Expifectamine 293 reagents were warmed in a water bath at 37°C. A master mix of Optimem I and Expifectamine was prepared. For each 10 mL of culture, 0.5 mL Optimem I was mixed with 27 µL Expifectamine 293 reagent, for example; 2.5 mL Optimem I was combined with 135 µL Expifectamine for each 50 mL culture.

DNA was diluted with Optimem I in a ratio 50:1 (µL Optimem I:µg DNA) e.g. 2500 µL/50 µg. The reagents were incubated in separate tubes for 5 mins at room temperature, then combined and incubated for a further 20-30 mins at room temperature. At the end of 20-30 minute incubation, the reagents were added to the culture (Day 0). The culture was returned to the Kuhner shaker at 37°C in 80% humidity with 8% CO_2 shaking at 125 rpm. At 16-18 hours post transfection, the cell density and viability were determined to ensure cells had not died or appeared to be dying. In addition to this, transfection enhancers were added to the culture 16-18 hours post transfection.

Per 10 mL culture, 50 µL Transfection Enhancer 1 (TE1) and 0.5 mL Transfection Enhancer 2 (TE2) were added e.g. for 50 mL culture, 250 µL TE1 and 2.5 mL TE2. The transfection was incubated for 6 - 8 days; cell density and viability were recorded daily to ensure viability did not fall below 70%. If viability did reach below this limit, cell culture media was harvested to prevent unwanted modifications to mAb protein due to toxic compounds being produced by the dead cells. The cell

culture media was harvested by transferring the cultures to 50 mL Falcon tubes and centrifuging at 182 x *g* for 7 minutes. The media supernatant was removed from the cell pellet and transferred to a clean tube. The pellet was discarded and the media was clarified by sequentially filtering it through 0.45 µm and 0.22 µm filters for either storage at -80°C or continuation to Protein A purification.

2.2.6.2.ExpiCHO-S Transfection

On the day prior to transfection split cells to $3 \times 10^6 – 4 \times 10^6$ cells/mL. On the day of transfection, cell density of $7 \times 10^6 – 1 \times 10^7$ cells/mL with >95% viability was required. OptiPRO and Expifectamine reagents were used directly from refrigeration for CHO-S transfection. A master mix of OptiPRO and Expifectamine was prepared in a ratio of 11.5:1 (µL OptiPro:µL Expifectamine). For example, for each 50 mL culture, 1840 µL OptiPRO was mixed with 160 µL Expifectamine reagent. DNA was diluted with OptiPRO to a ratio of 1:40 (µg DNA : µL OptiPRO). The two reagent mixes were combined and slowly added to cell culture flask while swirling (Day 0). Return to Kuhner shaker at 37°C in 80% humidity with 8% CO_2 shaking at 125 rpm. Following 18-22 hours, post transfection, cell density and viability were determined. Transfection enhancer and feed were also added at this stage. ExpiCHO enhancer and feed were added in a ratio of 6 µL of enhancer and 300 µL of feed per 1 mL of culture. For example; per 50 mL culture 300 µL ExpiCHO Enhancer and 15 mL ExpiCHO Feed were added. The transfection was incubated for 8-10 days, recording cell density and viability daily to ensure viability did not decrease below 60-70%. Cell media harvest and clarification was performed in the same manner as the HEK transfection.

2.2.6.3. Indirect Transfection with Polyethylenimine (PEI)

Each indirect transient transfection was performed in the CHO-S cell line. Each culture was 50 mL of CHO-S culture with 80 µg plasmid DNA (40 µg VH and VL). Cell density and viability parameters were kept the same as the standard transfection protocol. Three hours prior to transfection 20 mM lithium acetate was

added to the transfection culture. The PEI: DNA ratio was 2:1. Plasmid DNA was diluted in a ratio 1:5 with OptiPRO. PEI was diluted 1:2 with OptiPRO. After a 15 minute incubation at room temperature the PEI and DNA was combined and added to the culture flask. 4 hours PT the culture was diluted 1:1 with CHO-S medium containing 2.2 mM valproic acid and 3.04mM caffeine. Cell harvest was kept constant over all CHO transfections.

2.2.6.4. Direct Transfection

Direct transient transfection procedure was the same as the indirect with two changes. No Lithium Acetate was added to the culture prior to transfection and the PEI and DNA were added separately to the culture without precomplexation. Cell harvest was the same as above.

2.2.7. Protein A Chromatography

Protein A purification chromatography was performed on a GE AKTA Start Chromatography system using 1 mL HiTrap Protein A resin columns at a flow rate of 1 $mL.min^{-1}$. The column was equilibrated with 10 column volumes of PBS prior to sample loading. Sample was delivered to the column via the sample pump port. The column was subsequently washed with 10 column volumes PBS and eluted with 100 mM Citric Acid, pH 3.2. Purified mAb was collected and the pH of the eluate tested using litmus paper. The eluate was then neutralized by adding 1 M Tris-Base until a pH of ~7 was indicated. Protein concentration was determined using the NanoDrop UV-Vis Spectrophotometer at 280 nm. Samples were then stored at -30°C for further analysis.

2.2.8. Strong Cation Exchange Chromatography

Separations were performed using an UltiMate 3000 UHPLC instrument consisting of an RS gradient pump, autosampler, column compartment and variable wavelength UV detector. A PCM 3000 pH and conductivity module was included

after the UV detector to facilitate online monitoring of eluate pH during development and optimisation experiments. Following method development, a Vanquish Flex Quaternary UHPLC instrument was used consisting of a quaternary gradient pump, autosampler, column compartment and variable wavelength detector. Both instruments were under the control of Chromeleon software, version 7.2, which was also used for data processing and analysis.

All separations were performed using either a MabPac SCX-10 column, 4 × 250 mm, 10 µm particle size or a MabPac SCX-10 RS 2.1 × 50 mm column, 5 µm particle size, both from Thermo Fisher Scientific. For pH gradient separations, the CX-1 buffers were diluted accordingly as per the manufacturers recommendations, gradient conditions were varied as described in Table 2.3.1, unless otherwise stated. The flow rate was 1 mL.min^{-1} throughout, UV detection was performed at 280 nm, and details regarding the optimal column temperature are outlined in Table 2.3.1. Separations were optimised for resolution and peak capacity using NIST mAb for all optimisation experiments. Salt gradient optimisation was performed in a similar manner to pH gradient using 60 mM NaCl in 20 mM MES, pH 5.6 as Buffer A and 300 mM NaCl in 20 mM MES, pH 5.6 as Buffer B again with a flow rate of 1 mL.min^{-1}. All samples, except where otherwise stated, were injected directly in formulation buffer to result in a final amount of 50 µg on column.

Ultrafast separations were investigated using NIST mAb and were based on the equation below, where "vol_g" is the gradient volume and "$\Delta\%B$" is the percent change in gradient[31,29].

Equation 1: $gradient\ ratio = \frac{\Delta\%B}{vol_g}$

A ratio value of 10 was considered optimum for ultrafast separations[31]. The gradient utilized for scouting experiments for the in-house produced trastuzumab biosimilar candidates was 25% B - 55% B in 3 minutes at a flow rate of 1 mL.min^{-1} with a

column temperature of 60°C. Evaluation of end-to-end process time was performed with injections of 10 µg of an in-house produced biosimilar of adalimumab using the same gradient as was used for the separation of commercially available adalimumab.

2.2.9. IdeS protease digestion and middle-up mass profiling

A total of 160 µg of each mAb in PBS or formulation buffer (2 µg/µL solution) were combined with 2 µL of FabRICATOR enzymatic digestion solution (67 units IdeS/µL in Optima grade water) and incubated at 37°C for 2 hours under agitation at 500 rpm. Disulfide bonds were reduced by incubating samples in 50 mM TCEP in 4 M guanidine hydrochloride for 45 minutes at 56°C. Following incubation, samples were dried *via* vacuum centrifugation and reconstituted in 0.1 % formic acid prior to LC-MS analysis.

LC-MS analysis was performed using a Vanquish Flex UHPLC coupled to a Q-Exactive Plus hybrid quadrupole Orbitrap mass spectrometer with extended mass Biopharma Option *via* a Heated Electrospray Ionization-II (HESI-II) probe in a standard Ion Max ion source. Table 2.2.6 shows the tune parameters used, while Table 2.2.7 shows the method settings applied.

Table 2.2.6: Tune file parameters for middle up mass profiling

Capillary Voltage (kV)	3.8
Capillary Temperature (°C)	320
Sheath Gas (AU)	25
Auxiliary Gas	10
Probe Heater Temp (°C)	150
S-Lens RF Level	60
High Mass Range mode	On
Trapping Gas Pressure Setting	1

Table 2.2.7: Method parameters for middle up mass profiling

Polarity	Positive
Microscans	5
Resolution (@ m/z 200)	240,000
Automatic Gain Control (AGC)	3e6
Maximum Inject Time (ms)	200
Mass Range (m/z)	600 – 2,400

The digested and reduced IgG domains were injected on to a 2.1 × 50 mm, 4 μm MabPac RP column. Mobile phase A was LC-MS grade water with 0.1% formic acid (v/v) and mobile phase B acetonitrile with 0.1% formic acid (v/v). The flow rate was 0.3 $mL.min^{-1}$ and the gradient parameters are shown in Table 2.2.8.

Table 2.2.8: Gradient parameters for RP-LC

Time	% Buffer B
0	25
1	25
15	32
15.5	80
16	80
16.5	25
26	25

The column and sample temperatures were 80°C and 5°C, respectively. An injection volume of 1 μL was used throughout. Following LC-MS analysis, data was processed using the Xtract algorithm in BioPharma™ Finder Version 3.0 and compared against the theoretical sequence of the mAbs studied.

Chapter 2.

2.3.Results & Discussion

2.3.1. Monoclonal Antibody Production

2.3.1.1.Molecular Biology

The sample set for this thesis was based on characterisation of 32 mAbs. To this aim 32 individual VL and VH genes were PCR amplified and ligated into commercial plasmids. The 64 plasmids were then transformed into competent *E. coli* cells. These bacterial colonies were then grown in culture to multiply the plasmid DNA, purified and subsequently transfected into mammalian cells.

The extent of time and laboratory work required to achieve this is difficult express. Following are a selection of results from each stage of the molecular experiments. Below are agarose gel images taken during PCR amplification, gel purification and restriction enzyme digestion of the VL and VH genes and respective plasmids.

Figure 2.3.1 below, illustrates the scale of the molecular work which was undertaken. This figure shows the 32 VL chains following PCR amplification.

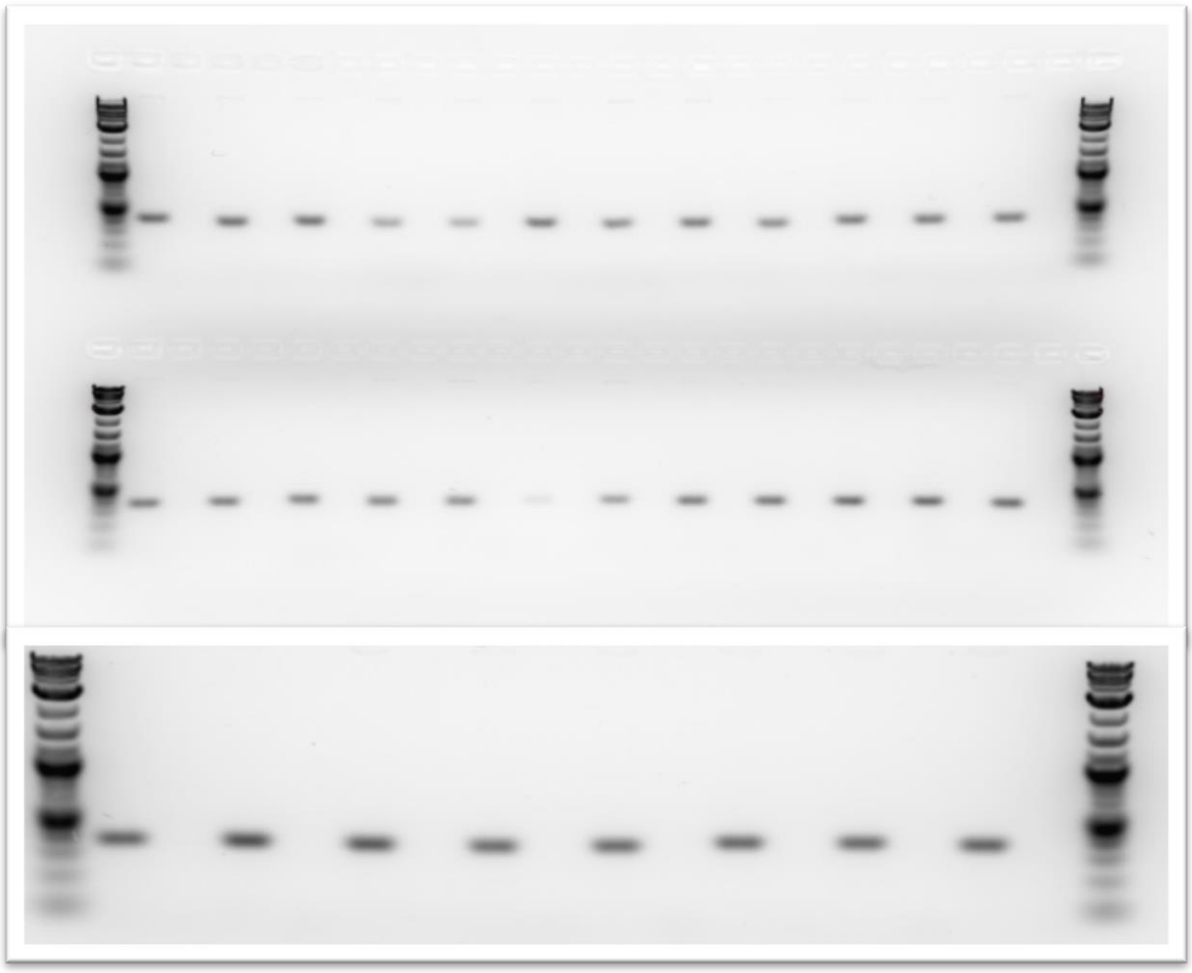


Figure 2.3.1: The 32 restriction enzyme digested VL genes before excision from 2% agarose gel.

Figure 2.3.2 (A) below depicts the agarose gels obtained following restriction enzyme digestion of the commercial plasmid. Part (B) shows the agarose gel following amplification of the VH genes. Figure 2.3.2 (C) & (D) shows the gels following restriction digestion of the kappa and lambda VL plasmids, respectively. The enzyme restrictions were considered successful by visual inspection of the molecular weight of the “drop-out” band below the main band for the plasmid. Purification of the VL and VH inserts was confirmed by visual inspection, ensuring that no other contaminant bands were present.

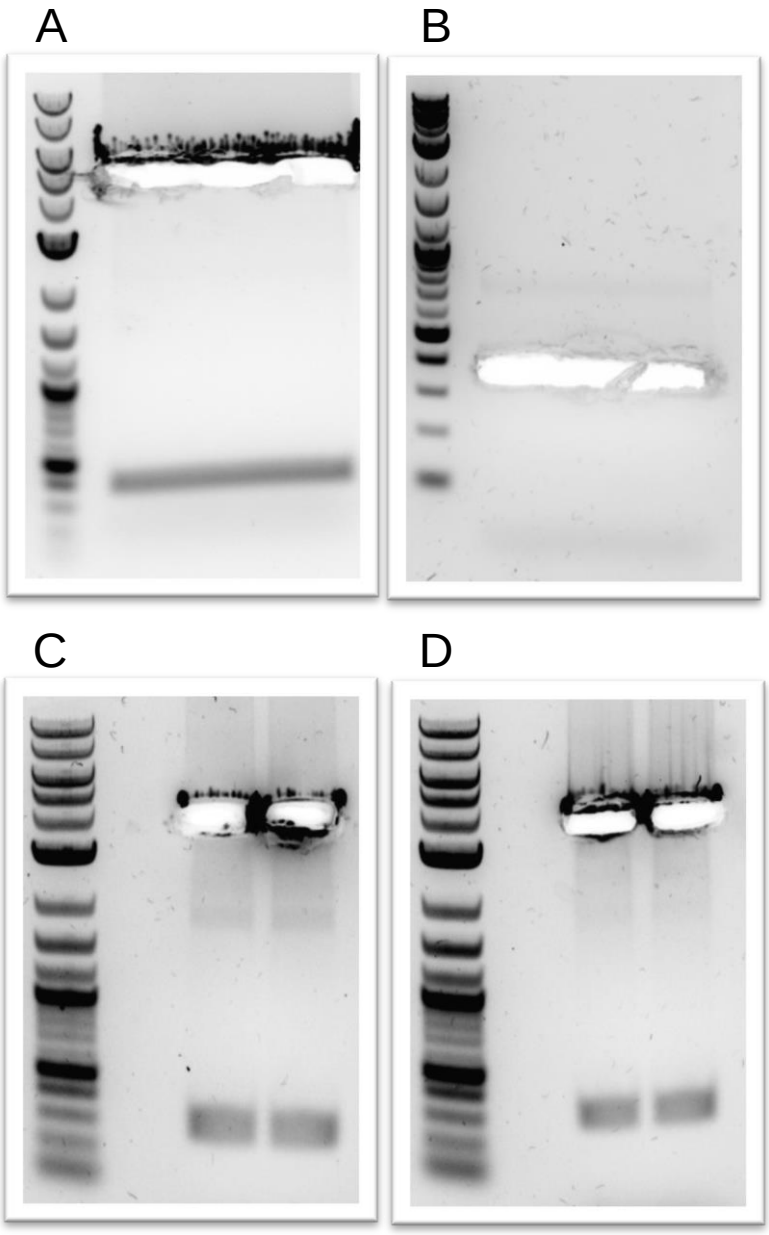


Figure 2.3.2: Agarose gel images following extraction of the pFUSE plasmid (A), the VH genes (B) and the pFUSE kappa (C) and lambda (D) light chains.

Following ligation of the VH and VL genes with pFUSE plasmid the vector was transformed into a competent *E. coli* cell line. Successful transformation was confirmed by growing the competent cells on a selective agar plate and comparing to a positive control. Following successful transformation a single colony sample was sent for conformational sequencing. After confirmation, glycerine banks were made of the transformed bacteria for long term storage in preparation for use during transfection.

Chapter 2.

2.3.1.2.Transient Transfections

The transient transfection experiments performed in the course of this body of work were largely performed using the proprietary Expifectamine kit. However, following some unsuccessful transfections, studies were undertaken into alternative, non-proprietary methods. Polyethylenimine (PEI) is a cationic polymer, which coats DNA in positive charges enabling endocytosis of the cell membrane. PEI was the carrier used as an alternative to Expifectamine. However, the proprietary kits also used “transfection enhancers” which were suspected of consisting of valproic acid and caffeine. The two methods of transfection tested with PEI were termed; direct and indirect. The main difference is the addition of Lithium Acetate prior to transfection, in the indirect method, to potentially increase the cells receptiveness to the PEI/DNA complex. These transfections were performed concurrently with a standard CHO transfection as a control. The PEI transfections failed to produce protein which could be detected following protein purification.

All subsequent transfections were performed using the proprietary kits. Table 2.3.1 below illustrates all antibodies and the total protein quantities which were obtained from those which were successful. All antibodies were transfected into HEK and CHO; however the vast majority failed to produce any detectable protein. The protein quantities for those which did are reported in Table 2.3.1.

Chapter 2.

Table 2.3.1: Full list of mAbs, concentrations of DNA obtained, protein expression quantities reported by Pfizer and total protein expressed in-house, in HEK and CHO.

Antibody	Other names	VL - DNA µg	VH - DNA µg	Expression (Pfizer reported) µg	Total Protein (mg) HEK	Total Protein (mg) CHO
Humira	Adalimumab	931.6	619.7	447		19.28
Campath	Alemtuzumab	764.5	496.5	182		
Simulect	Basiliximab	409.6	863.4	286		
Benylsta	Belimumab	1513.6	2485.1	1276		
Avastin	Bevacizumab	741.8	223.7	297		4.65
Adcetris	Brentuximab	661	882.4	1607		
Ilaris	Canakinumab	693.4	817.8	1118		
Cimzia	Certolizumab	810.3	552.3	2		
Erbitux	Cetuximab	515.8	1465.9	46	6.97	25.75
Zenapax	Daclizumab	820.1	585.5	673		
Prolia	Denosumab	997	800.8	1167		
Soliris	Eculizumab	758.9	1908.5	808		
Raptiva	Efalizumab	1584.6	1097.4	897		
Mylotarg	Gemtuzumab	631.4	635.4	501		
Simponi	Golimumab	409.6	863.4	652		
Zevalin	Ibritumomab		1209.9	311		
Remicade	Infliximab	1062.4	1943.1	ND	2.16	24.36
Yervoy	Ipilimumab	1540.3	874.6	1042		6.61
OKT3	Muromonab	356.3	433.6	259		
Tysabri	Natalizumab	629.9	656.2	864		
Gazyva	Obinutuzumab	1054.6	1189.7	1420		
Arzerra	Ofatumumab	1477	882.4	1474		
Xolair	Omalizumab	791.7	445.4	955		
Synagis	Palivizumab	1417.6	2122.6	55		
Vectibix	Panitumumab	1025.1	492.2	120		
Perjeta	Pertuzumab	1526.4	1347.7	1387		
Lucentis	Ranibizumab	626.9	863	4		
ABThrax	Raxibacumab	943.9	2065.5	1672		
Rituxan	Rituximab	287.4	1294.6	78	5.6	16.74
Actemra	Tocilizumab	398	1138.2	ND		
Herceptin	Trastuzumab	642.9	698.4	1093	14.4	15.69
Stelara	Ustekinumab	437.8	1123.5	1106		

For illustrative purposes, the recorded transfection data for the HEK cetuximab transfections are shown in Figure 2.3.3 and 2.3.4 below. Figure 2.3.3 shows the cell viabilities recorded across the 6 days PT. Figure 2.3.4, then, shows the Protein A chromatograms obtained from each replicate's purification. Table 2.3.2 shows the individual protein totals obtained for each of the 6 replicates. One replicate failed to produce any detectable protein and this is supported by the null absorbance observed during elution from the protein A column.

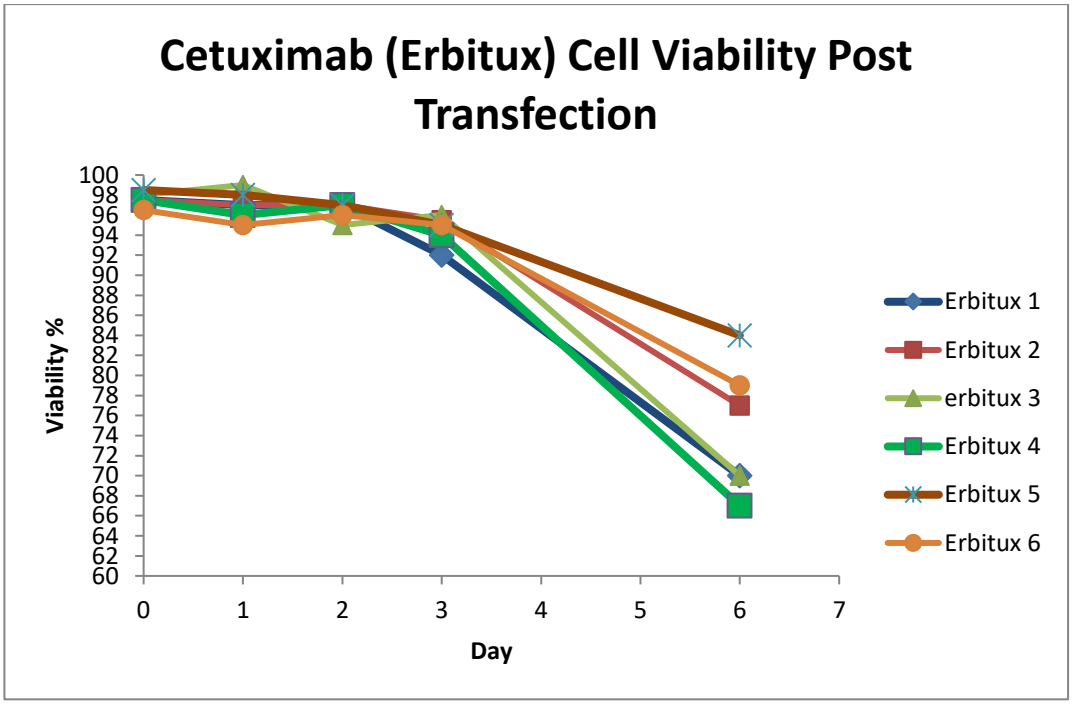


Figure 2.3.3: Cell viability post transfection – Erbitux

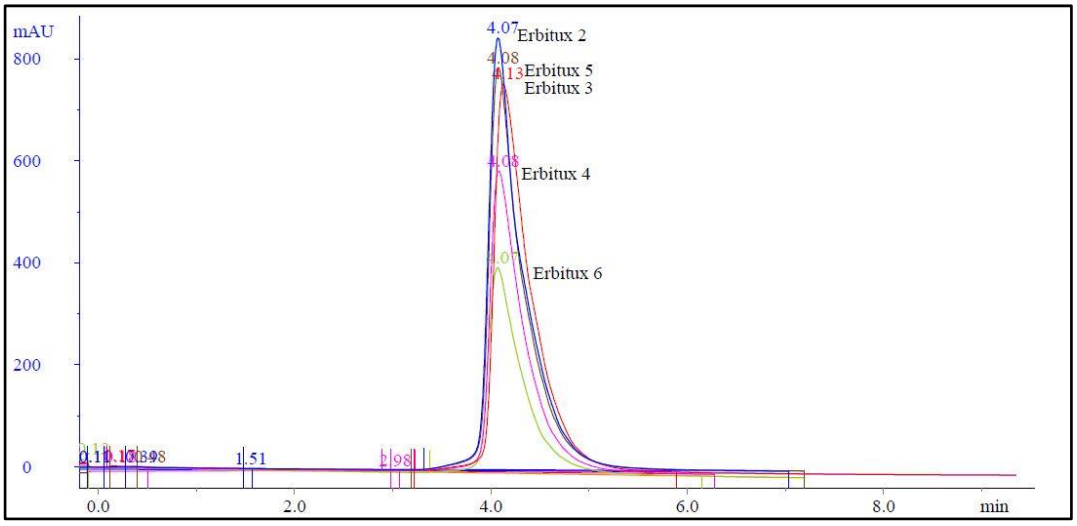


Figure 2.3.4: Protein A Chromatogram of Erbitux samples

Table 2.3.2: Total mg of Erbitux obtained from HEK transfection

Cetuximab	1	2	3	4	5	6
Total mg	0.00	2.898	1.188	0.903	1.296	0.688

Once sufficient amounts of mAbs had been obtained, roughly 6 mg per mAb, molecular work ceased and characterisation began. Prior to characterisation of in-house mAbs, characterisation methods were developed to reduce analysis time and increase

throughput. The following SCX studies were performed to firstly investigate the best elution mode for the purposes of this work and secondly to develop a separation method which could be used to analyse large sample sets.

However, future work on the antibody library will continue as it is an extremely valuable asset.

2.3.2. Method development and optimisation

An investigation into the optimum gradient slope, separation temperature and gradient length for each, pH and salt elution modes was performed. With the aim of determining the most selective and sufficiently resolved charge variant separation of NIST mAb within a reasonable analysis time. The buffers used for pH elution investigations were proprietary formulations, CX-1. The salt buffers were produced in the lab and were MES based. Initial separations generated for NIST mAb using each elution mode are shown in Figure 2.3.5 (A) using the MabPac SCX-10 4 × 250 mm column. Three parameters were used to definitively determine the optimum variable settings to be used for an optimized pH and salt separation method; peak capacity (Pc), resolution (Rs) and apparent retention factor (k_{app}). Minimum acceptance criteria included Rs ≥1.5 for the main peak and the first basic peak, and the highest achievable value for Pc. The apparent retention factor, k_{app} was used to determine appropriate gradient starting conditions to ensure complete analyte adsorption to the stationary phase and hence repeatable and robust chromatography. A k_{app} value between 4 and 18 in combination with Pc and Rs was employed, as k_{app} values <4 showed lower resolution of basic variants, while values >18 resulted in excessive peak retention and therefore created a limit on potential throughput. Experimental optimisation of these parameters for each elution mode is presented in Table 2.3.2 and overlay chromatograms are depicted in Figure 2.3.5 (B), for pH gradient elution and Figure 2.3.5 (C) for salt gradient elution.

Table 2.3.3: Optimisation of gradient conditions for pH and salt elution modes. Peak capacity (Pc.), resolution (Rs.) and apparent retention factor (k_{app}) values for pH (left) and salt (right) gradient modes for varying conditions of gradient slope (%B), time and temperature. Values in orange boxes represent the optimum parameters.

pH				Salt			
Gradient (%B)	Pc.	Rs.	k_{app}	Gradient (%B)	Pc.	Rs.	k_{app}
0-100	201.71	1.98	19.37	0-100	233.88	2.29	8.94
20-100	60.66	0.81	13.97	5-50	124.94	2.66	13.64
30-80	118.53	2.25	14.12	**10-45**	101.99	2.77	12.23
50-100	151.38	1.91	3.03	15-80	171.03	2.54	5.59
				20-60	123.29	2.60	3.93
Time (mins.)	Pc.	Rs.	k_{app}	Time (mins.)	Pc.	Rs.	k_{app}
10	88.83	1.72	6.52	10	82.77	2.32	5.70
15	105.16	2.01	8.71	15	84.95	2.35	7.51
20	109.75	2.14	10.68	20	93.15	2.60	9.16
30	118.42	2.25	14.73	**30**	101.99	2.77	12.23
40	124.08	2.40	18.14	40	100.50	2.75	14.98
Temp (°C)	Pc.	Rs.	k_{app}	Temp (°C)	Pc.	Rs.	k_{app}
20	119.49	2.27	14.89	20	95.24	2.60	11.81
30	125.92	2.48	14.09	**30**	102.01	2.77	12.23
40	125.99	2.55	13.56	40	96.44	2.70	12.63
50	124.84	2.58	13.17	50	94.10	2.65	13.27
60	129.13	2.73	12.83	60	89.20	2.54	14.04

Based on these investigations, the optimum separation conditions for pH elution were a gradient of 30% - 80% B over 30 minutes at 60°C and for salt gradient elution, a gradient of 10% - 45% B over 30 minutes at 30°C was best suited for the NIST mAb.

Chapter 2.

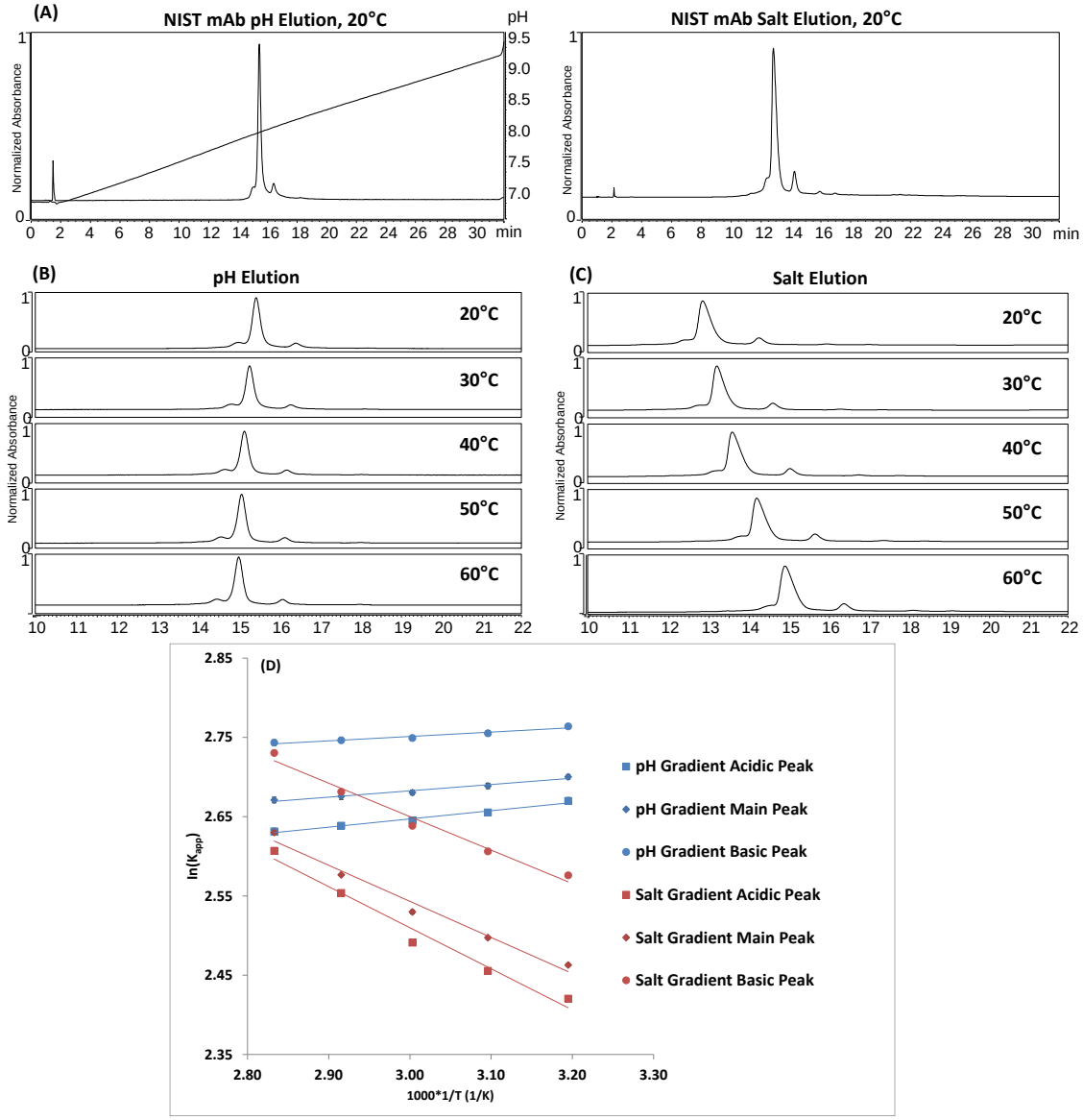


Figure 2.3.5: (A) Separation of NIST mAb by pH gradient elution using the commercially available CX-1 pH buffer system (left panel) and by salt gradient elution (right panel); (B) Overlay depicting the behavior of NIST mAb using pH gradient elution at different temperatures; (C) Overlay depicting the behavior of NIST mAb using salt gradient elution at different temperatures; (D) Van't Hoff plots generated from the experimental data in (B) and (C) above for both, pH and salt gradient elution.

Van't Hoff plots were prepared for the main peak, acidic and basic peaks for the NIST mAb, to investigate the effect of temperature on the separation. Examination of the resulting Van't Hoff plots, depicted in Figure 2.3.5 (D), revealed that temperature could not be used to increase selectivity; however, the effect of temperature on the two separation modes was different. The change in adsorption enthalpy (ΔH) resulting from the transfer of the protein between the mobile and stationary phases was estimated from the slope of the fitted linear regression line in the Van't Hoff plots[21,32].

For both the salt and pH gradient elution approaches, a linear relationship between k_{app} and temperature for the main NIST mAb peak, along with the acidic and basic variant peaks was observed. Interestingly, for salt gradient elution k_{app} was noted to increase with increasing temperature with the inverse being observed for pH gradient elution. Salt gradient elution was determined to exhibit positive ΔH values suggesting that adsorption is an endothermic process. As can be seen from Figure 2.3.5 (C), retention increased with increasing temperature. The opposite was observed for pH gradient elution of NIST mAb although the effect was less pronounced; see Figure 2.3.5 (B). The minor effect of temperature on pH gradient elution may be attributed to a slight decrease in the viscosity of the eluent with increasing temperature which caused a slight leftward shift in the pH gradient as monitored using post column pH and conductivity detection. The behavior of NIST mAb under these conditions is consistent, with elution occurring when the pH of the mobile phase reaches the molecules apparent pI value resulting in spontaneous desorption and elution from the column due to alteration of the chemical environment within the column that prevents re-adsorption to the stationary phase. The observation of increased retention with increasing temperature using salt gradient elution may be attributed to a number of factors including an increase in the mAbs diffusion coefficient and a decrease in the conductivity of the mobile phase with increasing temperature. Previous studies, wherein micro-calorimetric investigations of protein binding to ion exchange resins have suggested that the types of binding forces contributing to adsorption; electrostatic or hydrophobic, were dependent on salt concentration at certain temperatures[33,34,35]. In such cases, ΔH decreased with increasing salt concentration due to NaCl creating a shielding effect on the column. The behavior of the acidic and basic variant peaks was consistent with that of the main peak isoform during all investigations when using both elution modes.

Comparing the elution modes, pH gradient elution was superior to salt based elution methods for the chromatographic parameters evaluated. Further investigation of the

elution modes was next evaluated for the development of ultrafast CEX separations for mAb charge variant profiling.

2.3.3. Ultrafast pH gradient CEX method development and validation

To enable the development of an ultrafast method, the ratio of percentage change in the chromatographic gradient versus the volume of the gradient was determined (Eq.1),[31,29]. A ratio of between eight and ten was recommended as an optimum value[31]. For NIST mAb, the optimized pH gradient was 30% - 80% B over time at 60°C, with a Δ%B of 50. For salt gradient elution, the optimized parameters were 10% - 45% B over time at 30°C, with a Δ%B of 35. The gradient time was initially set to five minutes at 1mL.min^{-1} (gradient volume = 5mL). This resulted in gradient ratios of 10 and 7, for pH and salt gradient elution, respectively. Adaptation to an ultrafast method was problematic for salt based elution, the short gradient time was not sufficient to provide an acceptable separation of the various NIST mAb variants, with elution of the antibody during the column flushing step. A partial charge variant separation was seen wherein a main peak with various shoulders was observed.

The overall aim however, was to create a charge variant separation within three minutes using the MabPac SCX-10 RS 2.1 × 50 mm column. Hence, the optimized gradients were further narrowed, maintaining a flow rate of 1mL.min^{-1} and gradients of 30% - 60% B (gradient ratio 10) for pH and 10% - 40% B (gradient ratio 10) for salt elution mode. Salt elution mode failed to deliver an acceptable charge variant separation within the gradient time, whereas the pH elution mode could maintain selectivity but with a dramatic reduction in the analysis time, see Figure 2.3.6. Compared to the longer gradient separation, the chromatographic selectivity of the ultrafast separation was maintained with the exception of the most abundant acidic peak, which began to co-elute with the main peak and appeared as a shoulder. Peak integration as separate species was still obtainable and an example of the integration approach is illustrated in Figure 2.3.6 (B)

Chapter 2.

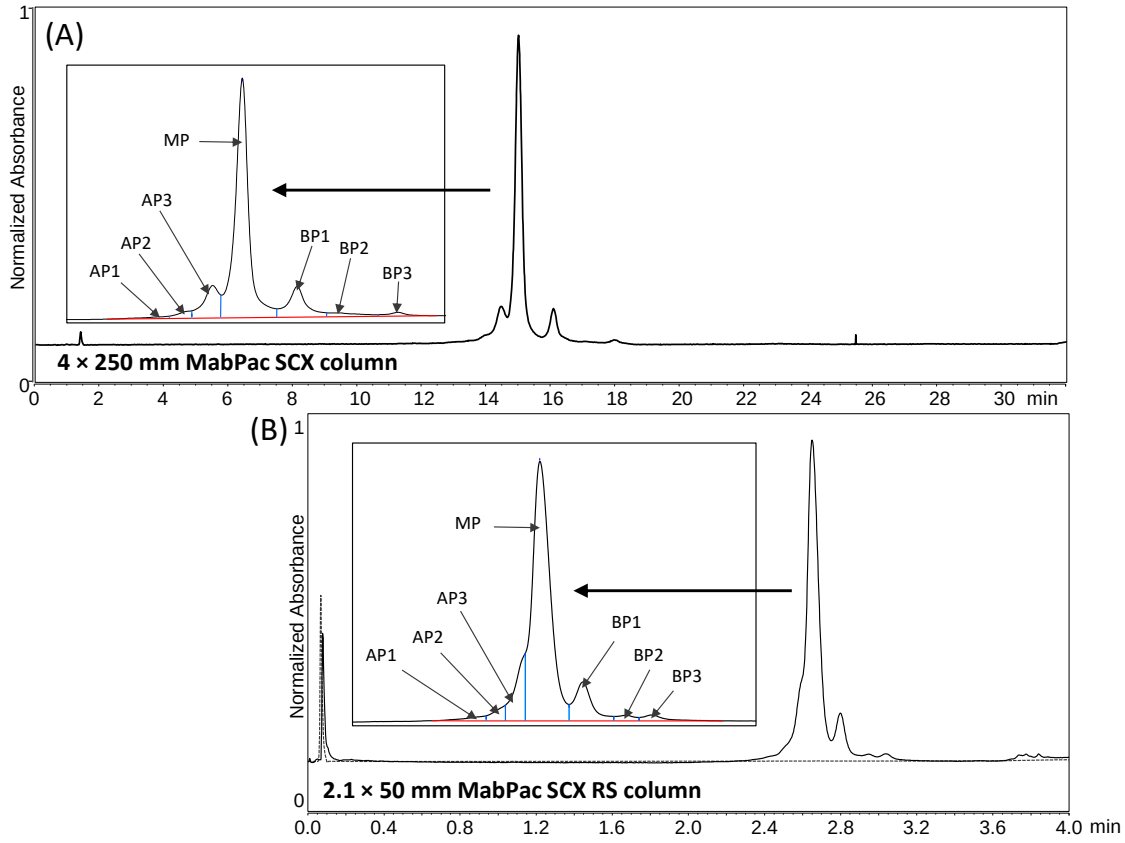


Figure 2.3.6: (A) Optimized pH gradient elution of NIST mAb from a MabPac SCX-10 4 ×250 mm cation exchange column. (B) Ultrafast separation of NIST mAb using same chromatographic conditions with smaller column and narrower elution range.

The ultrafast pH method was validated for linearity, repeatability and sensitivity as per ICH guidelines[36]. The resulting linearity plot had a correlation coefficient (R^2) of >0.99, indicating an acceptable level of linearity. The limit of detection (LOD) and limit of quantitation (LOQ) were determined using the slope of the regression line (S) and the standard deviation of the response (σ), LOD was 3.3 σ and LOQ was 10 σ and were determined to be 1.9 and 6.4 µg/µL, respectively. Repeatability was evaluated by six repeat injections of NIST mAb. The percentage relative standard deviation (%RSD) of the peak area and retention time of the main peak were determined and were less than 1% and 0.1%, respectively, indicating that the method is highly repeatable. The calibration curve used to derive the LOD and LOQ is shown in figure 2.3.7 below. The data table used to populate the curve is shown in Table 2.3.4 with the standard deviation and %RSD for the peak area average.

Chapter 2.

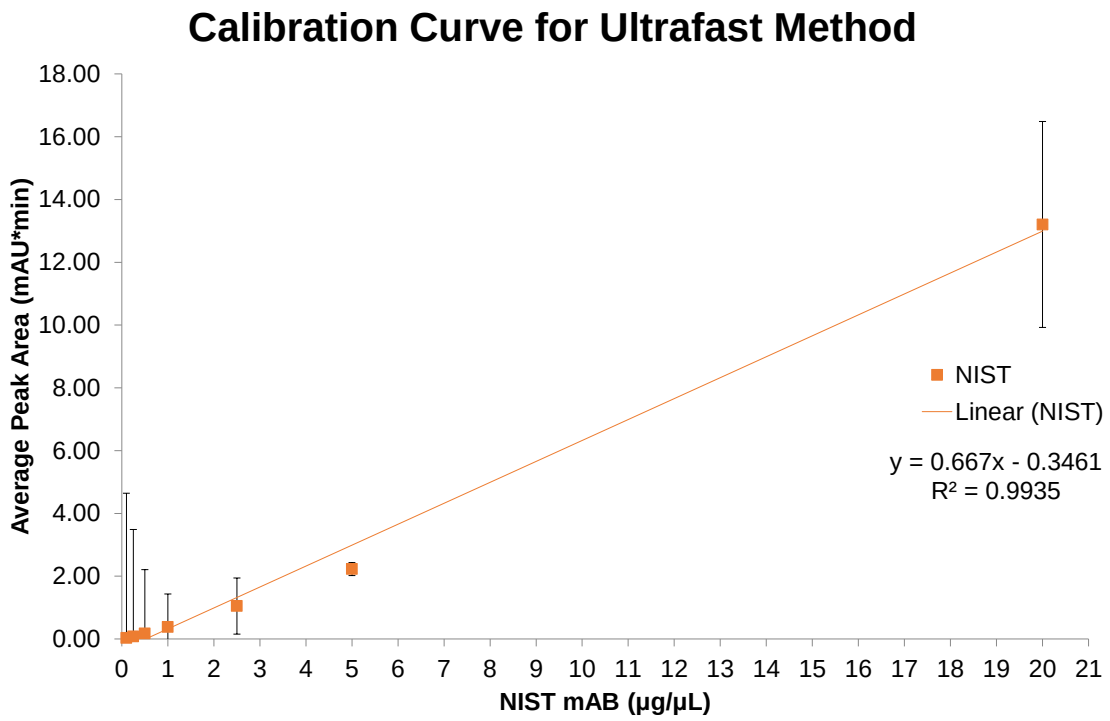


Figure 2.3.7: Calibration curve for the ultrafast method validation. The average area of the main peak (n=3) was plotted against increasing concentrations from 0.1 µg/µL to 20 µg/µL. The equation of the fitted trend-line was then used to establish LOD and LOQ for the method. Error bars are fitted against %RSD for the average peak area.

Table 2.3.4: Calibration curve data

Conc. µg/µL	Average Peak Area (mAU*min)	Stdev	%RSD
0.10	0.0301	0.0014	4.6114
0.25	0.0818	0.0028	3.4077
0.50	0.1747	0.0036	2.0358
1.00	0.3840	0.0040	1.0471
2.50	1.0499	0.0094	0.8931
5.00	2.2275	0.0046	0.2087
20.00	13.2049	0.4334	3.2824

Table 2.3.5 below shows the results of the repeatability test where 6 repeat injections were analysed and the average retention time and peak area, of the main peak, were recorded. The standard deviation and %RSD are low showing high repeatability.

Table 2.3.5: Repeatability data

n	RT (mins)	Peak Area (mAU*min)
1	2.892	2.314
2	2.894	2.324
3	2.893	2.289
4	2.893	2.328
5	2.892	2.290
6	2.893	2.306
Mean	2.893	2.308
Max	2.894	2.328
Min	2.892	2.289
Std. Dev	0.001	0.015
% RSD	0.020	0.659

To determine the suitability of the ultrafast method for in-process screening, repeat injections of the in-house variant of adalimumab were performed (n = 45), Figure 2.3.8 (A) shows the resulting overlay. The mAb sample was injected in the protein A elution buffer, neutralised citric acid formulation (~pH 7).Figure 2.3.8 (B) shows a single sample charge variant separation. Three minor acidic variants followed by the main variant and three minor basic variants were observed

Chapter 2.

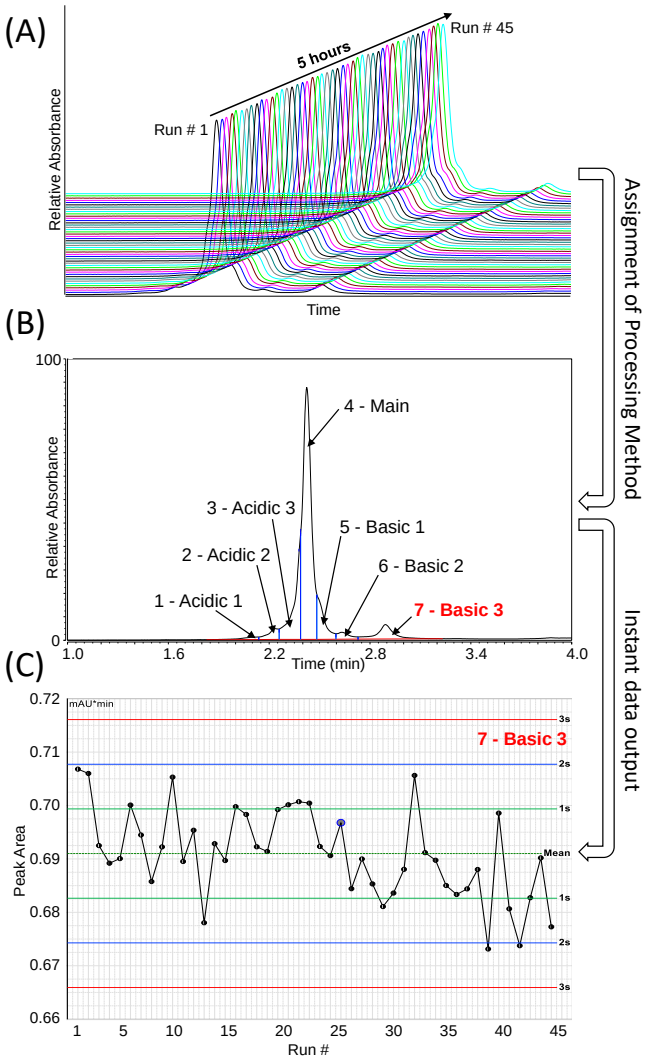


Figure 2.3.8: (A) Comparison of consecutive injections of the adalimumab biosimilar using the ultrafast method. (B) Processing method integration of single injection. Basic peak three is highlighted in red, this peak was used as a quality criterion which was monitored throughout all injections. (C) Line graph showing the peak area of basic peak three over the course of all injections. Green, blue and red lines represent one, two and threefold standard deviations.

High repeatability over the 45 runs is demonstrated in Figure 2.3.8 (A) as charge variant patterns appear to be highly similar throughout all injections. This suggests full compatibility of this separation system with typical protein A elution buffers, used in PAT frameworks. This also supports the application of this method for product and process development, as a high sampling numbers obtained during downstream processing can be screened within a reduced processing time.

CVA traces of samples can be automatically acquired with simultaneous assignment of a processing method. This enables the analyst to obtain in depth chromatographic information in one click. Using the Cobra Wizard in Chromeleon 7.2, a suitable

processing method was created in less than five minutes which was used for all subsequent runs in the sequence. The area of basic peak three (peak number 7) was used as a reference peak to ensure consistency throughout.

Chromeleon allows the user to create standard deviation charts, based on peak area, as a report, Figure 2.3.8 (C) shows the chart generated for this study. This feature allows the user to track out-of-specification events over the course of a series of injections. It took five hours, for the complete measurement and analysis of this sample set. This highlights the useful-ness of this approach for product and process development.

2.3.4. Rapid charge variant scouting of in-house biosimilar candidates

Following validation of the method, the CV profiles of in-house biosimilar candidates were evaluated. Initial scouting runs with a twelve minute gradient of 0% to 60% B at a flow rate of 0.5 $mL.min^{-1}$ (gradient ratio 10) at 60°C were performed for each of the samples. The sample set consisted of; Expi293 and ExpiCHO produced transiently transfected trastuzumab, a stably transfected CHO trastuzumab and trastuzumab drug product as a reference. Trastuzumab drug product is produced in CHO cells[37] and so it was hypothesized that the in-house produced CHO mAbs, transiently and stably expressed, should exhibit highest similarity. Figure 2.3.5 shows the chromatograms obtained for all scouting runs.

Trastuzumab is a well characterized mAb and literature reports have previously described the sources of posttranslational modifications leading to the generation of the charge variant profile [30] [38]. Isoforms in the acidic region have been attributed to deamidation of asparagine 30 in the light chains of the mAb [39]. The predominant glycans present on trastuzumab drug product include the standard bi-antennary glycans bearing core fucosylation and between zero to two galactose residues. The degree of sialylation on trastuzumab is low and therefore unlikely to make any significant contribution to the

charge variant profile observed [38]. Isoforms in the basic region of the drug product profile are attributed to isomerization of aspartic acid and C-terminal lysine clipping [39].

The trastuzumab drug product chromatogram features four visible variants in the acidic region, Figure 2.3.9 (A1 and A2). One variant is featured in acidic region 1, two are clear peaks and one appears as a minor peak in the shoulder of the larger acidic variant peak (acidic 2). There is one clearly visible basic variant peak and one which appears as a shoulder of the main isoform. From the literature a broad identification can be assigned to regions in the chromatogram and tentatively extrapolate these identifications to the in-house produced trastuzumab variants[39].

The most striking charge variant profile obtained was for the Expi293 produced transiently transfected mAb, which features a high content of acidic species with four prominent peaks, likely be attributable to posttranslational modifications previously reported in the literature[39]. The transiently transfected ExpiCHO variant, in contrast, contains lower amounts of acidic variants but still slightly more than the reference drug product. Another observation is that unlike the drug product, all in-house molecules show basic variants in the basic region two (B2) which indicates the presence of additional basic variants. However, a single step purification of the in-house expressed material was performed rather that a complete downstream purification and polishing process as was the case for the reference drug product.

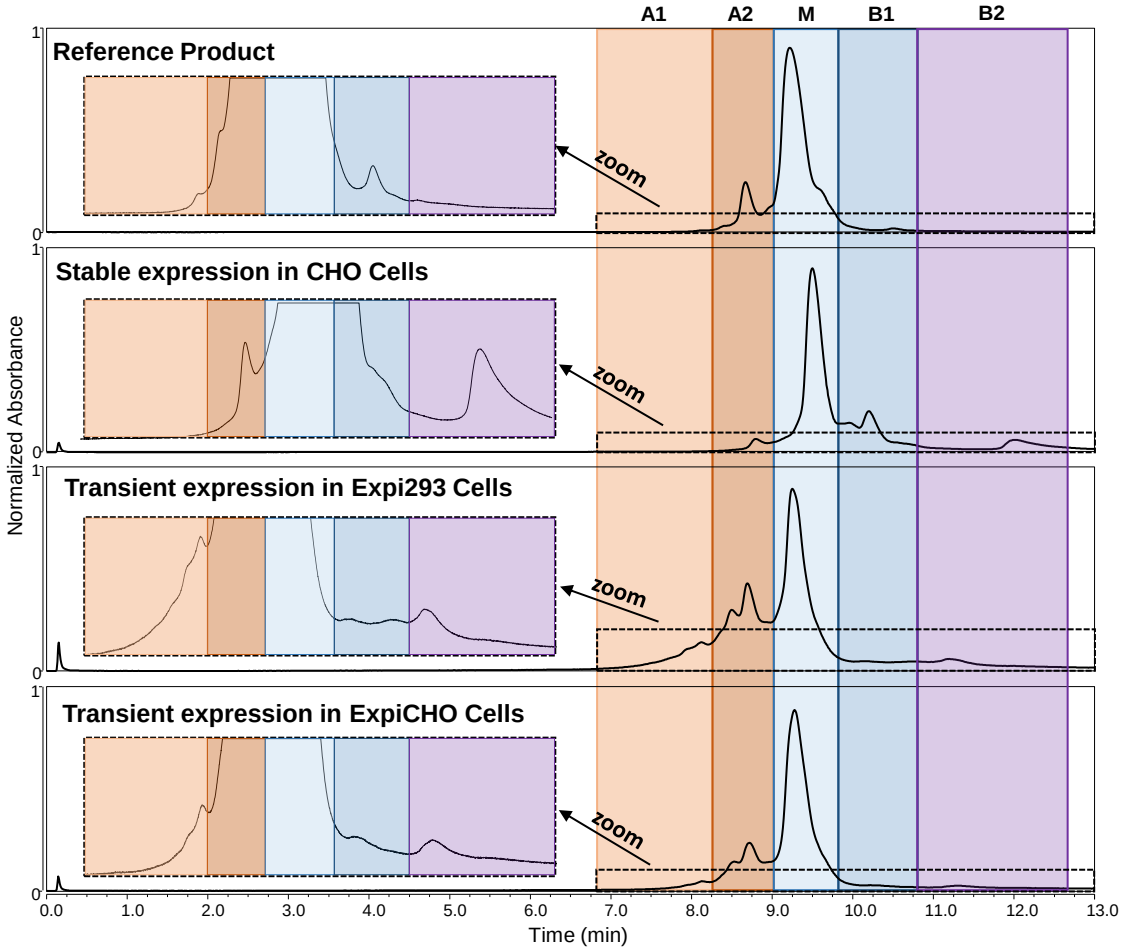


Figure 2.3.9: Scouting separations of the trastuzumab sample set consisting of the reference drug product and in-house produced candidate biosimilar material using transient transfection of Expi293 and ExpiCHO cells and a stably transfected CHO cell line. All separations were performed using pH gradient elution on a MabPac SCX-10 RS 2.1 × 50 mm cation exchange column.

Interestingly, the stably transfected CHO produced mAb had a CV profile considerably different to the transiently produced mAbs and the drug product, eluting later, at a more basic pH, indicating a more basic pI. This mAb had an identical purification process and was produced on the same scale, using the same transfection reagents. However, unlike the transiently produced mAb, the stably produced mAb had a different transfection procedure and the inclusion of a clonal selection workflow. To further understand the reason for the different behavior of the mAb derived from the stably expressed CHO line, a middle-up LC-MS analysis was performed on the material following digestion with IdeS protease and reduction. The light chain, Fc/2 and Fd fragments were separated using reversed-phase liquid chromatography coupled to high resolution Orbitrap mass

spectrometer. The data files obtained were analyzed with the Sliding Window deconvolution feature in BioPharma Finder 3.0 and are shown in Figure 2.3.10. Deconvolution revealed an alteration in the light chain, wherein incomplete clipping of the signal peptide sequence was observed, resulting in the presence of five additional N-terminal amino acids. This could cause a difference in chromatographic behavior on the SCX stationary phase and the overall shift in retention time.

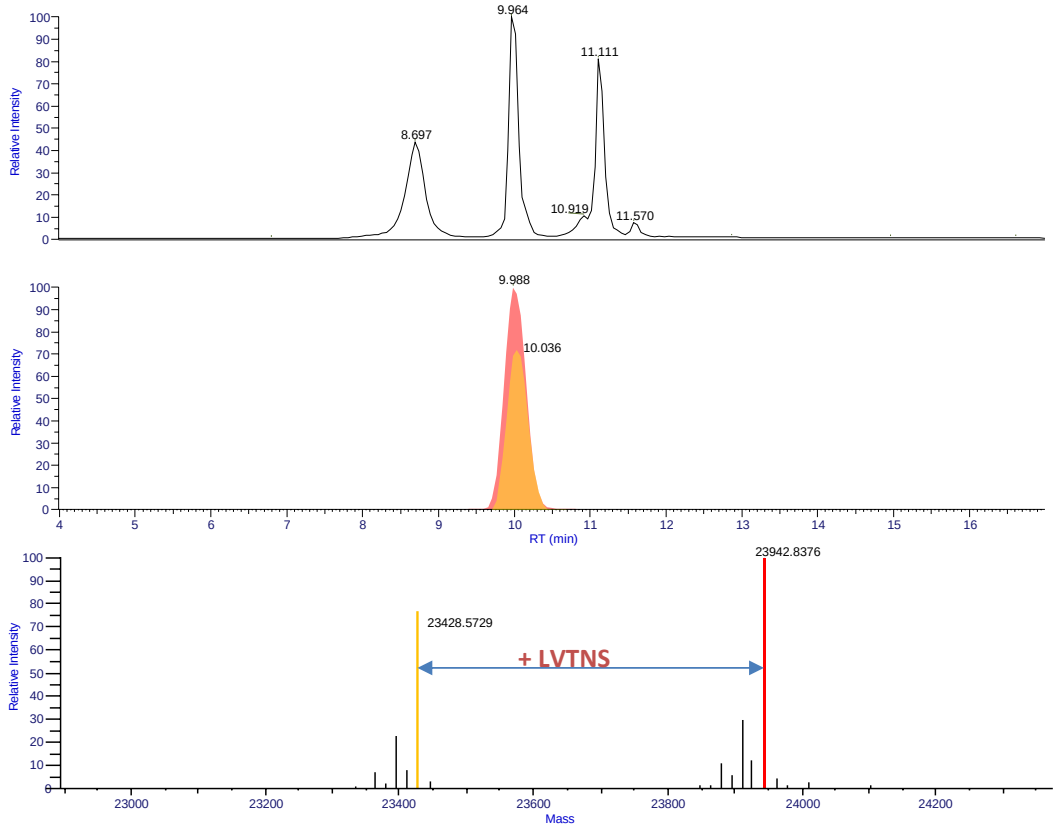


Figure 2.3.10: Top trace: Reversed-phase LC-MS separation of candidate trastuzumab biosimilar expressed using the in-house generated, stably transfected CHO cell line following digestion with IdeS protease and reduction of disulfide bonds; Middle trace: generated chromatogram following deconvolution depicting the presence of two dominant species in the chromatographic peak at 9.964 minutes that correspond to the light chain; Lower trace: deconvoluted mass spectra showing that the different species detected correspond to the light chain bearing an additional five amino acids resulting from incomplete cleavage of the signal peptide.

Once the main features of each mAb CV profile were noted, the ultra-fast method was applied and the outcome was evaluated for comparability to what had been observed in the scouting runs. The chromatograms using the ultrafast pH gradient elution method for all trastuzumab samples are depicted in Figure 2.3.11.

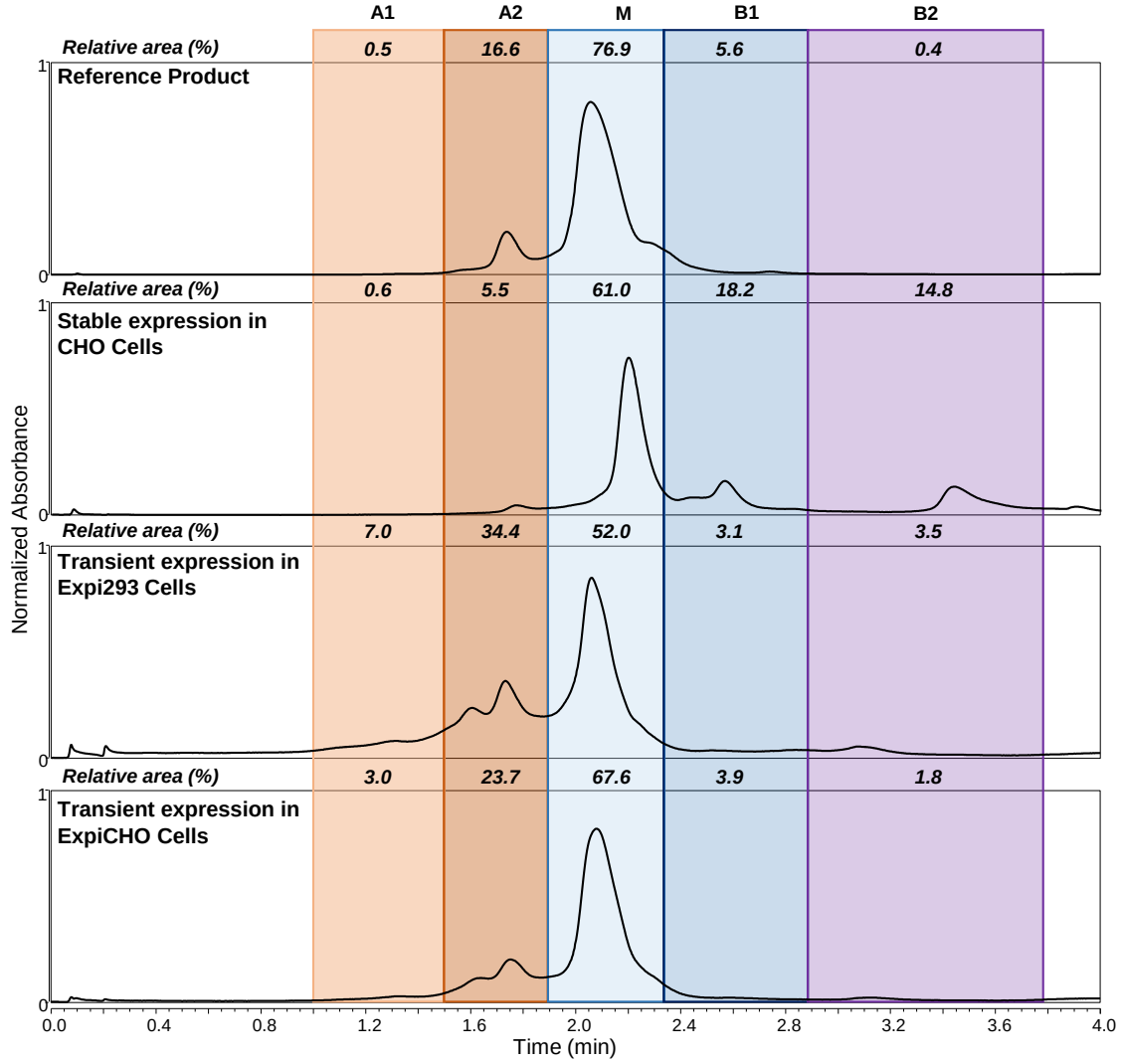


Figure 2.3.11: Ultrafast 3 minute separations of the trastuzumab sample set consisting of the reference drug product and in-house produced candidate biosimilar material using either transient transfection of Expi293 and ExpiCHO cells and a stably transfected CHO cell line.

Separation of the main CV isoforms for the trastuzumab sample set showed excellent comparability to the longer scouting gradient. All major peaks featured in the chromatograms shown in Figure 2.3.10 were also observable in those obtained for the ultra-fast method. Peak capacity, resolution and retention factors were determined for both scouting and ultrafast runs. A reduction in these metrics was expected, as using the ultrafast method requires a trade off in performance for analytical speed. Resolution was generally ≥1.5 with peak capacity values in the range 22 – 30, sufficient to facilitate rapid profiling for molecular triage or for potential PAT applications for process monitoring. Peak areas of the acidic, main and basic region windows have been outlined

for comparison purposes and are shown in Figure 2.3.11 above each chromatogram. Peak analysis shows, higher proportions of acidic and basic variants of the biosimilars compared to the drug product. Whereas, the stably expressed trastuzumab exhibits a high content of basic variants, reaching approximately 18 and 15% of the total area for both, B1 and B2 regions. The transiently transfected antibodies have a high acidic variant content, increasing to almost 35 and 24%, for the HEK and CHO respectively. These results demonstrate the applicability of the method for fast, highly customisable and in-depth comparison of mAb samples.

Chapter 2.
2.4.Conclusion and future work

In this chapter the initial molecular work has been outlined which resulted in a handful of mAbs being produced in useable quantities. The DNA amplification was successful however transfection into two mammalian cell lines remained difficult. The mAbs that were produced in high enough quantities are used throughout this thesis for comparison with their respective commercial forms. Following the molecular portion of this study, comprehensive analytical methods were developed.

In this chapter, an evaluation of pH gradient elution versus salt gradient elution in SCX chromatography using the NIST mAb standard reference material on a MabPac SCX-10 column was undertaken. This study revealed that pH gradient elution resulted in superior chromatographic performance for profiling the charge variant pattern of the NIST mAb when compared to salt gradient elution. Generated Van't Hoff plots showed that temperature could not be used to tune selectivity using both pH and salt gradient elution, but also revealed different behavior of the two elution mechanisms on the same stationary phase following temperature changes. pH gradient elution was nearly unaffected by increasing the separation temperature, whereas increased retention of NIST mAb was observed when using salt gradient elution which might be due to a reduction in the conductivity of the eluent. The pH gradient elution was further developed and validated into an ultrafast method to facilitate molecular triage based on the charge variant profile during lead candidate and cell line development. Analysis of an in house variant of adalimumab, directly derived from protein A purification, showed the applicability of method to in-process samples. Repeat injections of the adalimumab biosimilar, with a total end-to-end processing time of five hours were demonstrated. This included data acquisition and automated analysis by the assignment of a customised processing method.

The ultrafast method was then applied to profiling trastuzumab drug product and a series of trastuzumab variants produced using transient transfection in different expression

systems and a stably transfected CHO cell line. Differences in the charge variant patterns were detected and with reference to the drug product higher levels of acidic variants were present in the transiently expressed mAbs.

Additionally, the stable CHO antibody was determined to exhibit longer retention on the MabPac SCX-10 column. Further investigation using a middle-up LC-MS approach following IdeS protease digestion and reduction of disulfide bonds revealed the presence of five additional amino acids on the N-terminus of the light chain, resulting from incomplete cleavage of the signal peptide.

The ultrafast method showed high repeatability, high versatility and has proven applicability for rapid screening and comparison of mAb samples.

Future work on this method will be on-line coupling in a large scale bioprocess for proof of concept.

This investigation provided an excellent grounding in the theoretical and practical aspects of analytical chromatography. The resulting paper has been well received and will contribute to the field of high throughput separations.

2.5. Author Contributions

Adapted from Trappe *et al.* Rapid Charge Variant Analysis of Monoclonal Antibodies to Support Lead Candidate Biopharmaceutical Development. In publication

Research study devised by Anne Trappe, Florian Füssl and Jonathan Bones; research work performed by Anne Trappe, Florian Füssl and Sara Carillo ; Chapter written by Anne Trappe and reviewed by Jonathan Bones.

Chapter 2.
2.6.References

1.Philippidis, A. The Top 15 Best-Selling Drugs of 2016 *Genetic Engineering & Biotechnology News* [Online], 2017. https://www.genengnews.com/the-lists/the-top-15-best-selling-drugs-of-2016/77900868.

2.Trefis Team. Roche's Blockbuster Oncology Drugs Losing Steam As They Approach Patent Expiry *Forbes* [Online], 2017. https://www.forbes.com/sites/greatspeculations/2017/10/12/roches-blockbuster-oncology-drugs-losing-steam-as-they-approach-patent-expiry/.

3.Sands, A. Top 10 patent losses to watch out for *Total BioPharma* [Online], 2014. http://www.totalbiopharma.com/2014/06/06/top-10-patent-losses-watch/.

4.Mukherjee, S. The World's Best-Selling Drug Just Lost a Key Patent Battle *Fortune* [Online], 2017. http://fortune.com/2017/05/18/humira-abbvie-patent-biotech/.

5.European Medicines Agency, Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues (revision 1). 2014.

6.Food and Drug Administration, Quality Considerations in Demonstrating Biosimilarity of a Therapeutic Protein Product to a Reference Product. 2015.

7.International Conference on Harmonisation, Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products. 1999; Vol. Q6B.

8.International Conference on Harmonisation, Comparability of Biotechnological/Biological Products Subject to Changes in Their Manufacturing Process. 2004; Vol. Q5E.

9.Reusch, D.; Haberger, M.; Maier, B.; Maier, M.; Kloseck, R.; Zimmermann, B.; Hook, M.; Szabo, Z.; Tep, S.; Wegstein, J.; Alt, N.; Bulau, P.; Wuhrer, M., Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles - Part 1: Separation-based methods. *mAbs* **2014,** *7* (1), 167-179.

10.Parr, M. K.; Montacir, O.; Montacir, H., Physicochemical characterization of biopharmaceuticals. *Journal of Pharmaceutical and Biomedical Analysis* **2016,** *130* (Supplement C), 366-389.

11.Leblanc, Y.; Ramon, C.; Bihoreau, N.; Chevreux, G., Charge variants characterization of a monoclonal antibody by ion exchange chromatography coupled on-line to native mass spectrometry: Case study after a long-term storage at +5°C. *Journal of Chromatography B* **2017,** *1048* (Supplement C), 130-139.

12.Fekete, S.; Beck, A.; Veuthey, J. L.; Guillarme, D., Ion-exchange chromatography for the characterization of biopharmaceuticals. *J Pharm Biomed Anal* **2015,** *113*, 43-55.

13.Dakshinamurthy, P.; Mukunda, P.; Prasad Kodaganti, B.; Shenoy, B. R.; Natarajan, B.; Maliwalave, A.; Halan, V.; Murugesan, S.; Maity, S., Charge variant analysis of proposed biosimilar to Trastuzumab. *Biologicals* **2017,** *46*, 46-56.

14.Rea, J. C.; Freistadt, B. S.; McDonald, D.; Farnan, D.; Wang, Y. J., Capillary ion-exchange chromatography with nanogram sensitivity for the analysis of monoclonal antibodies. *Journal of Chromatography A* **2015,** *1424*, 77-85.

15.Salas-Solano, O.; Kennel, B.; SungAe, S. P.; Roby, K.; Sosic, Z.; Boumajny, B.; Free, S.; Reed-Bogan, A.; Michels, D.; McElroy, W.; Bonasia, P.; Hong, M.; He, X.; Ruesch, M.; Moffatt, F.; Kiessig, S.; Nunnally, B., Robustness of iCIEF methodology for the analysis of monoclonal antibodies: An interlaboratory study. *Journal of Separation Science* **2012,** *35* (22), 3124-3129.

16.Tebbey, P. W.; Varga, A.; Naill, M.; Clewell, J.; Venema, J., Consistency of quality attributes for the glycosylated monoclonal antibody Humira (R) (adalimumab). *Mabs* **2015,** *7* (5), 805-811.

17.Ahrer, K.; Jungbauer, A., Chromatographic and electrophoretic characterization of protein variants. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2006,** *841* (1-2), 110-122.

18.Vanam, R. P.; Schneider, M. A.; Marlow, M. S., Rapid quantitative analysis of monoclonal antibody heavy and light chain charge heterogeneity. *mAbs* **2015,** *7* (6), 1118-1127.

19.Du, Y.; Walsh, A.; Ehrick, R.; Xu, W.; May, K.; Liu, H., Chromatographic analysis of the acidic and basic species of recombinant monoclonal antibodies. *mAbs* **2012,** *4* (5), 578-585.

20.Fekete, S.; Beck, A.; Fekete, J.; Guillarme, D., Method development for the separation of monoclonal antibody charge variants in cation exchange chromatography, Part I: Salt gradient approach. *Journal of Pharmaceutical and Biomedical Analysis* **2015,** *102* (Supplement C), 33-44.

21.Fekete, S.; Beck, A.; Fekete, J.; Guillarme, D., Method development for the separation of monoclonal antibody charge variants in cation exchange chromatography, Part II: pH gradient approach. *Journal of Pharmaceutical and Biomedical Analysis* **2015,** *102* (Supplement C), 282-289.

22.Lingg, N.; Berndtsson, M.; Hintersteiner, B.; Schuster, M.; Bardor, M.; Jungbauer, A., Highly linear pH gradients for analyzing monoclonal antibody charge heterogeneity in the alkaline range: Validation of the method parameters. *Journal of Chromatography A* **2014,** *1373* (Supplement C), 124-130.

23.Talebi, M.; Nordborg, A.; Gaspar, A.; Lacher, N. A.; Wang, Q.; He, X. Z.; Haddad, P. R.; Hilder, E. F., Charge heterogeneity profiling of monoclonal antibodies using low ionic strength ion-exchange chromatography and well-controlled pH gradients on monolithic columns. *Journal of Chromatography A* **2013,** *1317* (Supplement C), 148-154.

24.Farnan, D.; Moreno, G. T., Multiproduct High-Resolution Monoclonal Antibody Charge Variant Separations by pH Gradient Ion-Exchange Chromatography. *Analytical Chemistry* **2009,** *81* (21), 8846-8857.

25.Zhang, L.; Patapoff, T.; Farnan, D.; Zhang, B., Improving pH gradient cation-exchange chromatography of monoclonal antibodies by controlling ionic strength. *Journal of Chromatography A* **2012,** *1272*, 56-64.

26.Rozhkova, A., Quantitative analysis of monoclonal antibodies by cation-exchange chromatofocusing. *Journal of Chromatography A* **2009,** *1216* (32), 5989-5994.

27.Kang, X.; Frey, D. D., High-performance cation-exchange chromatofocusing of proteins. *Journal of Chromatography A* **2003,** *991* (1), 117-128.

28.Patel, B. A.; Pinto, N. D. S.; Gospodarek, A.; Kilgore, B.; Goswami, K.; Napoli, W. N.; Desai, J.; Heo, J. H.; Panzera, D.; Pollard, D.; Richardson, D.; Brower, M.; Richardson, D. D., On-Line Ion Exchange Liquid Chromatography as a Process Analytical Technology for Monoclonal Antibody Characterization in Continuous Bioprocessing. *Analytical Chemistry* **2017,** *89* (21), 11357-11365.

29.Joshi, V.; Kumar, V.; Rathore, A. S., Rapid analysis of charge variants of monoclonal antibodies using non-linear salt gradient in cation-exchange high performance liquid chromatography. *Journal of Chromatography A* **2015,** *1406* (Supplement C), 175-185.

30.Wagner-Rousset, E.; Fekete, S.; Morel-Chevillet, L.; Colas, O.; Corvaïa, N.; Cianférani, S.; Guillarme, D.; Beck, A., Development of a fast workflow to screen the charge variants of therapeutic antibodies. *Journal of Chromatography A* **2017,** *1498*, 147-154.

31.Cook, K.; Steiner, F.; Pra, M. D. *Development of Ultra-fast pH-Gradient Ion Exchange Chromatography for the Separation of Monoclonal Antibody Charge Variants*; ThermoFisher Scientific: 2014.

32.Roselin, L. S.; Lin, M.-S.; Lin, P.-H.; Chang, Y.; Chen, W.-Y., Recent trends and some applications of isothermal titration calorimetry in biotechnology. *Biotechnology Journal* **2009,** *5* (1), 85-98.

33.Bowes, B. D.; Traylor, S. J.; Timmick, S. M.; Czymmek, K. J.; Lenhoff, A. M., Insights into Protein Sorption and Desorption on Dextran-Modified Ion-Exchange Media. *Chemical Engineering & Technology* **2011,** *35* (1), 91-101.

34.Lin, F.-Y.; Chen, C.-S.; Chen, W.-Y.; Yamamoto, S., Microcalorimetric studies of the interaction mechanisms between proteins and Q-Sepharose at pH near the isoelectric

point (pI): Effects of NaCl concentration, pH value, and temperature. *Journal of Chromatography A* **2001,** *912* (2), 281-289.

35.Silva, G. L.; Marques, F. S.; Thrash, M. E.; Dias-Cabral, A. C., Enthalpy contributions to adsorption of highly charged lysozyme onto a cation-exchanger under linear and overloaded conditions. *Journal of Chromatography A* **2014,** *1352*, 46-54.

36.International Conference on Harmonisation, Validation of Analytical Procedures: Text and Methodology. 2005; Vol. Q2 R1.

37.European Medicines Agency, Scientific Discussion of Herceptin. 2005.

38.Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianférani, S., Characterization of Therapeutic Antibodies and Related Products. *Analytical Chemistry* **2013,** *85* (2), 715-736.

39.Harris, R. J.; Kabakoff, B.; Macchi, F. D.; Shen, F. J.; Kwong, M.; Andya, J. D.; Shire, S. J.; Bjork, N.; Totpal, K.; Chen, A. B., Identification of multiple sources of charge heterogeneity in a recombinant antibody. *Journal of Chromatography B: Biomedical Sciences and Applications* **2001,** *752* (2), 233-245.

Chapter 3.

3. Comprehensive Characterisation of the Heterogeneity of Adalimumab via Charge Variant Analysis Hyphenated On-Line to Native High Resolution Orbitrap Mass Spectrometry

3.1. Introduction

Biopharmaceutical development and manufacture is a highly regulated environment. Like traditional small molecule production, regulatory agencies have detailed guidelines on the levels of heterogeneity which can be deemed safe and effective[1]. MAbs are subject to a wide variety of modifications which result in heterogeneity. Glycosylation, C-terminal lysine clipping, oxidation, deamidation, disulphide bond scrambling, fragmentation, aggregation, glycation, isomerization and pyroglutamate formation are modifications which are typically monitored[2]. The levels of post translational modification can be influenced by media components and the cell culture environment, during production[3]. To assure product quality, it is mandatory to monitor critical quality attributes, using adequate analytical techniques which utilise differences in physicochemical properties of the analytes such as size, surface charge or hydrophobicity.

The surface charge of a mAb can be altered by glycosylation, lysine clipping, deamidation and pyroglutamate formation[4]. Cation exchange chromatography (CEX) utilises the heterogeneous charge to separate a protein into isoform fractions. CEX can be used in either salt gradient or pH gradient elution mode, which is the chromatographic equivalent to isoelectric focusing, as discussed in the previous chapter[5]. Both modes have been applied successfully and were subject to a series of comparison studies with diverse outcomes[5,6,7]. In the previous chapter an in-depth exploration of the merits of each elution mode was undertaken. It was found that pH elution was superior by analysing the peak capacity, resolution and apparent retention factor of CV separation of the NIST mAb.

Chapter 3.

CEX is most commonly used in comparability studies as the identification of observed variants requires peak fraction collection and multi-step sample preparation. This means that in-depth analysis of charge variants is often time-consuming and costly[9]. More importantly, these procedures are a possible source for sample preparation induced modifications[10].

Coupling of CVA to MS has been the subject of numerous studies[11,12,13]. Stoll *et al*. (2015) hyphenated 2D LC separation to ToF-MS[14]. In this study the authors used CEX separation followed by RP separation as CEX buffers are not compatible with MS analysis. This problem is widely known and investigations in to volatile buffers have been undertaken[11,12,15,16]. Füssl *et al.*(2018) developed a tool for on-line CVA-MS using volatile buffers which has been shown to provide a stable pH gradient, hence facilitating CVA of multiple, commercially available mAb drug products using Orbitrap based MS detection[17]. However, the presence of adducts caused by the nature of the buffer system and the ESI-source parameters applied, required further investigation.

In this chapter, the buffer constituents used in the Füssl paper are used with a wider pH range. These buffers consisted of ammonium bicarbonate and ammonium hydroxide. In the previous chapter proprietary CX-1 buffers were used but are not MS compatible, therefore the volatile buffer system described by Füssl was employed to facilitate MS analysis of mAbs. The modified buffer system enabled the analysis of highly basic mAbs like rituximab or the NIST reference antibody. It was found that data acquisition benefits from higher resolution settings, when discrimination of co-eluting, near isobaric substances is required. Adalimumab was recently subjected to batch-to-batch and originator-to-biosimilar comparison experiments and demonstrated a complex charge variant pattern[15,5,16]. In these studies, glycan profiling indicated a dominance of FA2 and FA2G1 glycans and the most frequent glycoforms were demonstrated to be FA2/FA2, FA2/FA2G1 and FA2G1/FA2G1 or FA2/FA2G2. CVA performed in a batch-to-batch comparison, showed the presence of high abundant basic species, which were attributed

to different C-terminal lysine forms[15]. The same analysis also showed the presence of minor acidic species which were unidentified.

In this chapter, sixteen of the major isoforms of adalimumab are identified. Among them were; proteoforms modified by different levels of C-terminal lysine clipping, deamidation, aspartic acid isomerisation, glycation and fragmentation. To investigate whether host cell proteins (HCP's) might be involved in the potential fragmentation of adalimumab, HCP analysis using LC-MS was performed. Several HCP's were detected, among them the protease Cathepsin L, which could contribute to protein fragmentation. Carboxypeptidase B (CpB) digestion and peptide mapping experiments were performed as orthogonal methods to confirm the observations on the intact protein level.

Chapter 3.

3.2. Materials & Methods

3.2.1. Reagents

Water (Optima™, LC-MS grade, Catalogue No 10505904), 0.1% formic acid in water (Optima, LC-MS grade, Catalogue No 10188164) and 0.1% formic acid in acetonitrile (Optima, LC-MS grade, Catalogue No 10118464) were provided by Fisher Scientific. DL-Dithiothreitol (Catalogue No D0632-1G), acetic acid (ACS reagent grade, ≥ 99.7%, Catalogue No A0808), ammonium bicarbonate (BioUltra, ≥ 99.5%, Catalogue No 09830) and ammonium hydroxide solution (BioUltra, 1M in H_2O, Catalogue No 09859) were purchased from Sigma-Aldrich.

The commercially available monoclonal antibodies trastuzumab, infliximab, bevacizumab, rituximab and cetuximab were provided by the Hospital Pharmacy Unit of the University Hospital of San Cecilio in Granada, Spain. Adalimumab was provided by St. Vincent's University Hospital in Dublin, Ireland and the NIST monoclonal antibody reference material (Catalogue No RM8671) was purchased from the National Institute of Standards and Technology.

Carboxypeptidase B (Catalogue No 10103233001) was obtained from Hoffmann-La Roche AG. The SMART Digest trypsin kit, magnetic bulk resin option (Catalogue No 60109-101-MB), KingFisher Deepwell 96 well plates and KingFisher Duo 12-tip combs were obtained from Thermo Scientific™. Vivaspin® 500 spin filters with 10 kDa cut-off were provided by Sartorius. Hi3 Phos B Standard (Catalogue No 186006011) for label free Hi3 quantification of HCP's was obtained from Waters.

3.2.2. Accelerated Aging Study (AAS)

Two hundred and ten µg of adalimumab in formulation buffer were incubated in triplicate for 2, 5, 11, 16, 21, 27 and 32 days at 55 °C in an Incubating Mini Shaker from VWR. Time points were chosen, according to the Arrhenius equation, to be equivalent to theoretical storage, at 4 °C for 2, 6, 12, 18, 24, 30 and 36 months, respectively, using

an aging factor of 2.0. Sample condensation during elevated temperature incubation was counteracted by daily centrifugation of the samples at 1,000 rcf for 1 min. After incubation, samples were stored at -20 °C.

3.2.3. Sample preparation for LC-MS analyses

Carboxypeptidase digestion of the intact antibody was performed by the addition of the enzyme to obtain a substrate-enzyme ratio of 5:1 (w/w). Digestion was performed at 37 °C for 2 hours with agitation at 450 rpm. For peptide mapping, 60 µg of adalimumab was buffer exchanged to dH_2O to a final concentration of 0.83 mg mL^{-1} using VIVASPIN spin filters with a cut-off size of 10,000 Da. For peptide mapping of the stressed samples, 20 µg of protein, per replicate, were merged to obtain a final protein amount of 60 µg for each accelerated aging time point.

Digestion was carried out with a KingFisher Duo Prime Purification System under operation of the BindIt software, version 4.0. Samples were diluted with SMART buffer in a KingFisher Deepwell 96 well plate to obtain a final concentration of 0.3 mg mL^{-1}. Magnetic beads were picked up with KingFisher Duo 12-tip combs and were washed for 1 min in SMART buffer, diluted with water in a ratio of 1:4 (v/v) before they were moved into the sample wells. Digestion was performed at 70 °C for 40 min; afterward tip combs together with the magnetic beads were discarded. Digested samples were reduced by addition of a 100 mM DTT solution to reach a final DTT concentration of 5 mM followed by incubation at 37 °C for 30 min. Following reduction, samples were directly analysed *via* LC-MS/MS. Quantitation of HCP's was accomplished by spiking Waters Hi3 PhosB standard, consisting of an equimolar mixture of six synthetic peptides of rabbit Phosphorylase B proteins, to obtain a final injection amount of 25 pmol per peptide (150 pmol total). For this experiment, the limits of column capacity were deliberately exceeded, by injecting a protein amount approximately thirtyfold higher than manufacturer's recommendation, in order to detect HCP which is normally at a lower concentration relative to the RDP.

Chapter 3.

3.2.4. CVA with UV detection for gradient optimisation

For mAb gradient optimization, 25 mM ammonium bicarbonate and 30 mM acetic acid in LC-MS grade water (pH 5.3) were used as buffer A. Ten mM ammonium hydroxide in LC-MS grade water (pH 10.9) was used as buffer B. Separations were performed on a Thermo Scientific Ultimate 3000 HPLC system, equipped with a Variable Wavelength Detector, a Split-Loop Auto sampler, a Dual Gradient Pump, a Column Compartment and an Ultimate 3000 PCM-3000 pH and Conductivity Monitor. A MAbPac SCX-10 RS column with dimensions of 2.1 × 50 mm and 5 µm particle size was used. The column oven was held at 25 °C, a flow rate of 0.4 $mL.min^{-1}$ was applied and detection was performed *via* UV absorption at 280 nm.

For gradient optimization 50 µg of antibody were injected per run. Gradients for all mAbs, after optimization are shown in Table 3.2.1. The gradient length was 25 minutes consisting of; 10 minutes gradient time, 3 minutes column flushing at 100% buffer B and 12 minutes re-equilibration time.

Table 3.2.1 Optimised gradients for seven commercially available mAbs

mAb	Time (min)	% buffer B	Curve
Infliximab	0	30	5
	10	55	
Bevacizumab	0	35	5
	10	60	
Cetuximab	0	30	5
	0.5	40	
	3.5	42	7
	7	55	5
	10	100	
Adalimumab	0	40	5
	10	100	
Trastuzumab	0	40	5
	10	100	
NIST mAb	0	90	5
	10	100	
Rituximab	0	85	5
	10	100	

3.2.5. CVA-MS analysis of adalimumab

Separations were performed on a Thermo Scientific Vanquish UHPLC system equipped with a Fluorescence Detector F, a Split Sampler HT, a Binary Pump H and Column Compartment. The chromatographic column used, as well as the buffers, temperature, flow rate and the gradient applied were the same as described for gradient optimisation. Fluorescence detection was employed using excitation and emission wavelengths of 280 nm and 360 nm, respectively, prior to MS detection. One hundred µL were injected per run and analysis was performed on a Q-Exactive Plus mass spectrometer hybrid quadrupole Orbitrap mass spectrometer with extended mass Biopharma Option. Data acquisition of unstressed and stressed adalimumab samples were performed in triplicate. LC and MS systems were hyphenated *via* a Heated Electrospray Ionization-II (HESI-II) probe in a standard Ion Max ion source. The MS tune parameters used are shown in Table 3.2.2. The MS method parameters are shown in Table 3.2.3.

Table 3.2.2: Tune file parameters

Capillary Voltage (kV)	3.6
Capillary Temperature (°C)	275
Sheath Gas (AU)	20
Auxiliary Gas	5
Probe Heater Temp (°C)	275
S-Lens RF Level	200
High Mass Range mode	On
Trapping Gas Pressure Setting	1

Table 3.2.3: Method parameters

Polarity	Positive
Collision Induced Dissociation (eV)	150
Microscans	10
Resolution (@ m/z 200)	17,500; 35,000; 70,000
Automatic Gain Control (AGC)	3e6
Maximum Inject Time (ms)	200
Mass Range (m/z)	2,500 – 8000

Data analysis was performed in BioPharma Finder 2.0 software using the ReSpect algorithm for deconvolution. Deconvolution of spectra of main lysine variants was performed using the Sliding Window deconvolution feature. All other analyses were based on deconvolution after manual peak integration. The average molecular masses obtained after deconvolution were compared to the theoretical adalimumab sequence mass considering FA2/FA2, FA2/FA2G1 and FA2G1/FA2G1 or FA2/FA2G1 glycoforms as well as theoretical masses of modifications such as: single and double lysine loss, glycation, single and double deamidation, succinimide formation, Asp loss and fragmentation. Details on the BioPharma Finder 2.0 setting, for the manually selected area are summarised in Table 3.2.4 below. The settings for the sliding window deconvolution are in Table 3.2.5.

Table 3.2.4: Deconvolution setting for manual area selection

Parameter	Intact mAb	Fragments
m/z range	5,000 – 7,000	2,500 – 5,000
Deconvolution Algorithm	ReSpect	ReSpect
Model Mass Range	145,000 – 152,000	40,000 – 60,000 (LMWF) 90,000 – 110,000 (HMWF)
Mass Tolerance	10 ppm	10 ppm
Charge State Range	20 – 30	10 -20 (LMWF) 20 – 30 (HMWF)
Minimum Adjacent Charges	3 - 3	3 - 3

Table 3.2.5: Parameters for sliding window deconvolution

Parameter	
m/z range	5,000 – 7,000
Deconvolution Algorithm	ReSpect
Model Mass Range	145,000 – 152,000
Mass Tolerance (ppm)	10
Charge State Range	20 – 30
Minimum Adjacent Charges	3 - 3
Retention Time Range	3 – 12
Target Avg. Spectrum Width (min)	0.3
Target Avg. Spectrum Offset (%)	28%
Merge Tolerance	10
Max RT gap (min)	0.166
Minimum number of detected intervals	3

3.2.6. Peptide Mapping

3.2.6.1. Separation and collection of acidic and main adalimumab peaks for peptide mapping

Separation was performed on a Thermo Fisher Scientific Vanquish Flex UHPLC with diode array detection. All LC parameters were the same as those used for gradient optimisation. The charge variant peaks of five separate CVA runs were collected and each replicate peak merged. A total of 100 µg of sample was injected per run.

3.2.6.2. Mass Spectrometry for peptide mapping

Peptide mapping experiments were performed on a Thermo Fisher Scientific Vanquish Flex UHPLC coupled to a Q-Exactive Plus mass spectrometer. Separations were performed on an Acclaim VANQUISH C18 column with dimensions of 250 × 2.1 mm and a particle size of 2.2 µm. Separation was at 25 °C with a flow rate of 0.3 ml min^{-1}. Gradient buffer A was LC-MS grade 0.1% formic acid in water and buffer B was 0.1%

formic acid in acetonitrile. A gradient of 2 – 40% buffer B in 45 min was used. Table 3.2.6, below, shows the tune settings used, while table 3.2.7 shows the MS method settings.

Table 3.2.6: Tune settings for peptide mapping

Capillary Voltage (kV)	3.8
Capillary Temp (°C)	320
Sheath Gas (AU)	40
Auxiliary Gas	10
Probe Heater Temp (°C)	400
S-Lens RF Level	50

Table 3.2.7: MS method settings for peptide mapping

Polarity	Positive
Collision Induced Dissociation (eV)	0
Default Charge State	2
Full MS	
Resolution (@ m/z 200)	70,000
Automatic Gain Control target (AGC)	3e6
Maximum Inject Time (ms)	100
Mass Range (m/z)	200 – 2,000
MS^2	
Resolution (@ m/z 200)	17,500
AGC target	1e5
Maximum IT (ms)	200
Loop counts	5
Isolation window (m/z)	2
Scan range (m/z)	200 – 2,000
(N)CE	28
Peptide match	Preferred
Dynamic exclusion (s)	7.0

3.2.6.3. Host cell protein identification

Peptide mapping of the drug product was performed using 5 µg of injected protein. For the deep search of HCP, 160 µg of protein was injected. LC-UV-MS data was analysed using BioPharma Finder 2.0. For peptide identification, the adalimumab light chain and heavy chain sequences were searched with the following modifications; C-terminal Lys loss, deamidation of Gln and Asn, oxidation of Met, Trp and Cys, glycation of Lys and succinimide on Asp. Search results were filtered for a confidence score of ≥ 0.8. Peptide mapping experiments of drug product were performed n = 3. Peptide mapping of charge variant peaks was performed n = 3 or n = 2.

For HCP analysis, data files obtained, were searched against the *Cricetulus Griseus* (Chinese hamster) database from Uniprot using PEAKS studio 7.5. The parent mass error tolerance was set to 10.0 ppm, fragment mass error tolerance was 0.01 Da. Trypsin was chosen as the digestion enzyme, one missed cleavage and no non-specific cleavages were specified. Oxidation and deamidation were included as variable modifications and a false discovery rate cut-off of 1% was applied. Protein annotation, data visualisation and quantification were performed in Progenesis QI for Proteomics.

Chapter 3.

3.3. Results & Discussion

3.3.1. Adjustment of the CVA-MS buffer system

The CVA-MS method developed by Füssl *et al.* (2018) was expanded to further enhance the capabilities of CVA-MS[17]. Low ionic strength eluents, which are required for successful MS detection, exhibit relatively low buffering capacity. In the aforementioned study, the buffering effect of the strong cation exchange (SCX) stationary phase, especially in the higher pH range, resulted in a limitation of the analytes which could be studied.

Buffer B was originally prepared with a pH of 10.18, however the pH delivered on column did not exceed 9.7, when measured with an on-line pH meter. To investigate potential expansion of the operational pI range, two mAbs with high isoelectric points were included into the sample set, the NIST reference antibody (pI = 9.2) and rituximab (pI = 9.4).

CVA-MS analysis of these two mAbs resulted in only partial elution of charge variants. This was due to the reported buffer system being limited to mAbs with pI values of < 9.2. Consequently, the pH of buffer B was adjusted from ~10.2 to 10.9 which resulted in a pH range of 5.3 to 10.2 covered. Figure 3.3.1 (A) shows the pH trace across the applied gradient.

Using a higher pH elution buffer, resulted in a longer equilibration time, compared to the lower pH buffer with a reported elution time of 10 minutes[17]. This was overcome by using a two-step equilibration; where the pH was rapidly shifted by flushing the column with 100% buffer A, for 5 minutes and then adjusted to the % B required for equilibration for 8 minutes. Figure 3.3.1 (B) shows the dramatic change in pH using the two-step equilibration.

Chapter 3.

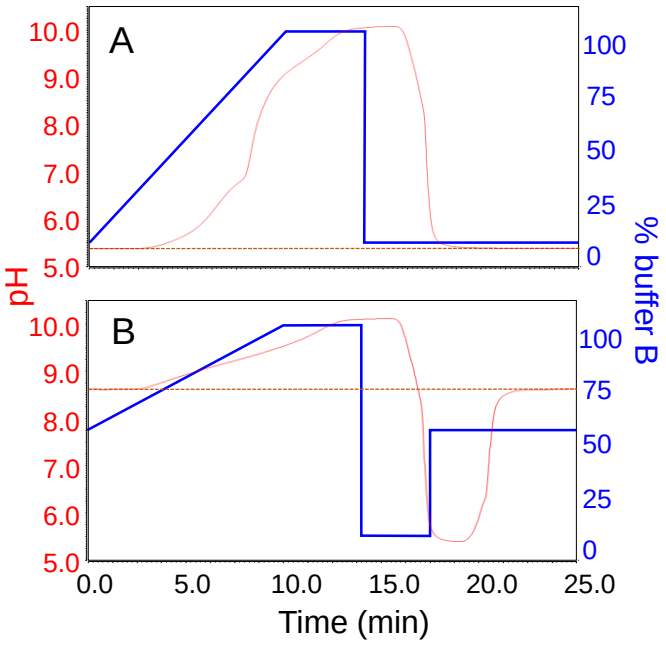


Figure 3.3.1: (A) Gradient from 0 – 100% buffer B (pH 10.9) in 10 minutes followed by 4 minutes of column flushing and 12 minutes of column equilibration. pH of the eluate was monitored in real time and the trace is shown in red. (B) Two-step re-equilibration with 100 % buffer A (pH 5.3) for 4 minutes followed by 50% buffer B for 12 minutes.

To investigate the applicability of the method to a range of pIs, seven commercially available mAbs with pIs ranging from 7.6 to 9.4 were analysed. For the majority of the mAbs studied, a method was established within three chromatographic injections. The majority of the mAbs tested showed a marked increase in the number of distinguishable charge variant peaks. Figure 3.3.2 shows the improvement in resolution achieved when gradient optimisation had been performed. The pI values shown in Figure 3.3.2 have previously been determined *via* imaged capillary isoelectric focusing by Goyon *et al*. (2017)[18]

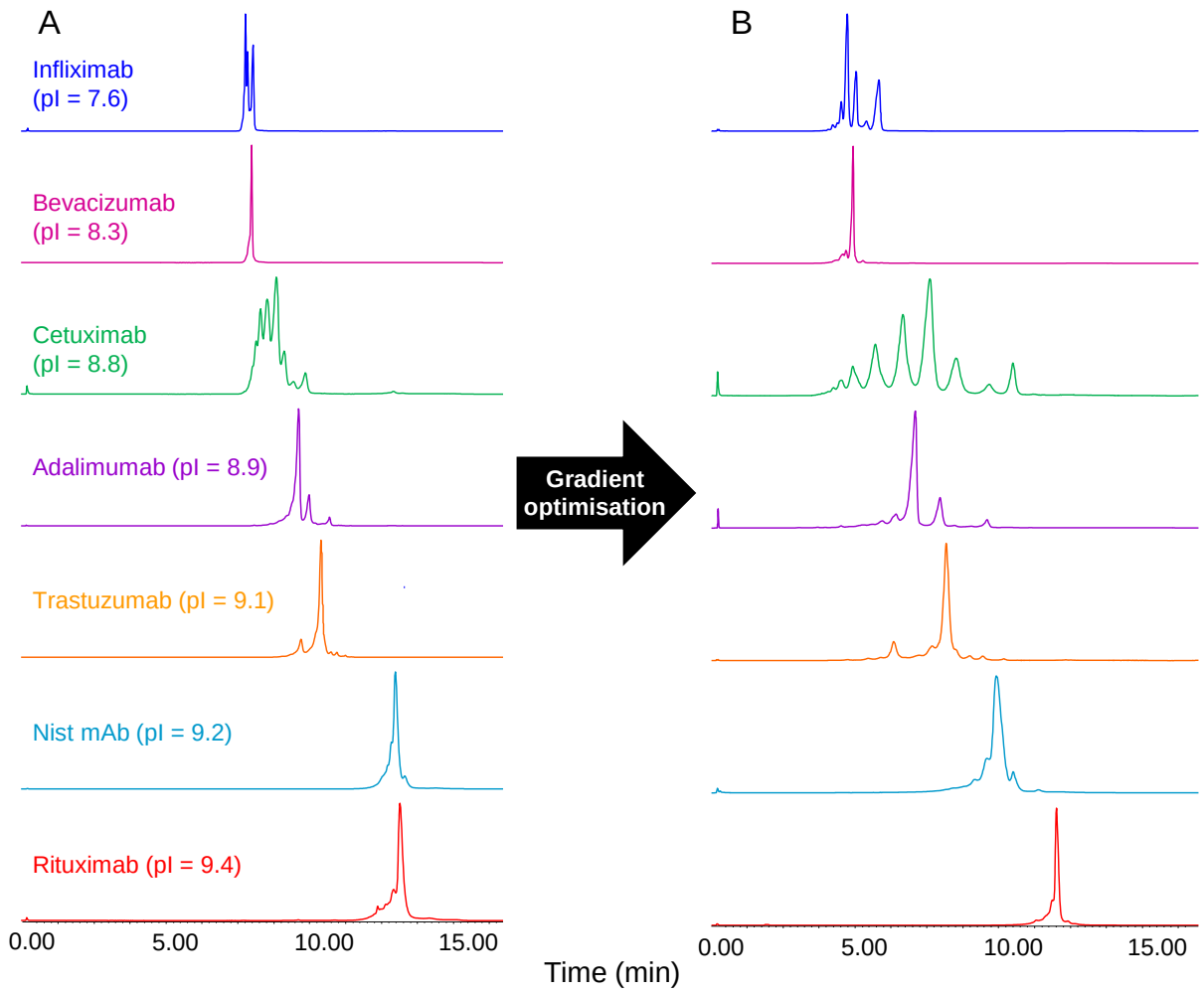


Figure 3.3.2: (A) Charge variant separation of seven monoclonal antibodies using a gradient from 0-100% buffer B in 10 min. (B) Charge variant separation after gradient optimisation.

The method buffer components and method developed by Füssl *et al.* required no further optimisation, beyond broadening of the pI range for coupling to MS. For most of the mAbs analysed, the UHPLC method provided enhanced selectivity and resulted in a variety of fully and partially resolved peaks.

Adalimumab exhibited a relatively complex charge variant pattern with a high number of low abundant peaks. The chromatogram obtained by method optimisation was analysed in more detail, and is shown in Figure 3.3.3. Magnification of the low abundant peaks shows; nine acidic, six basic and the main variant. Adalimumab was chosen for further intact mass analysis, as the heterogeneity of the mAb and the high quality of the

separation provided a good opportunity to test the methods effectiveness for peak annotation based on retention time and associated high resolution MS data.

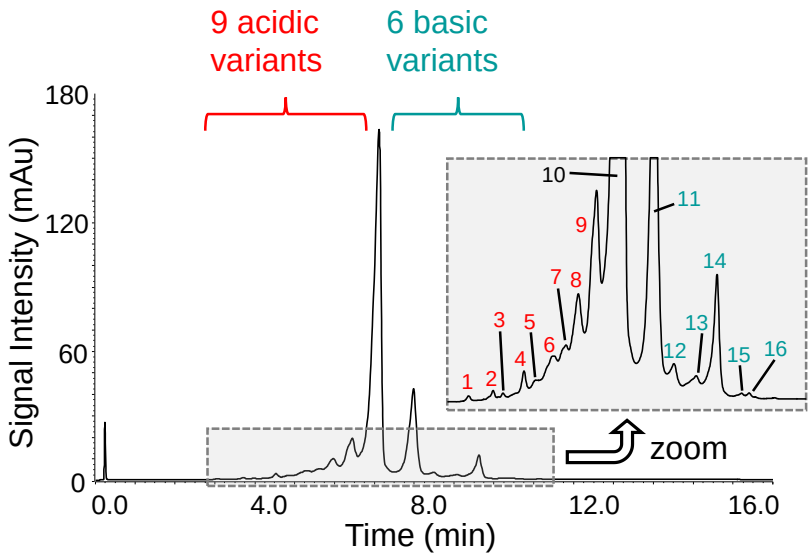


Figure 3.3.3: UV-chromatogram highlighting the charge variant separation of adalimumab. Data was acquired using the optimized gradient for adalimumab (Table 3.2.4.1).

3.3.2. Impact of MS resolution settings

One hundred µg of adalimumab were analysed with three different MS resolution settings: 17,500, 35,000 and 70,000. Figure 3.3.4 shows a comparison of the MS base peak chromatograms (BPC's) and the averaged spectra obtained for the main adalimumab charge variant. The BPC's generated with all three resolution settings look highly similar and closely resemble the chromatogram acquired with fluorescence detection.

The optimal resolution setting depends on several physical parameters: the molecular mass, charge state and collisional cross section of the ions. The gas pressure within the C-trap and Orbitrap at the time of trapping and mass analysis also has a significant effect on the optimisation of the resolution setting. These factors impact the stability of the ion current as by increasing kinetic energy and shortening the lifetime of analyte ions in the Orbitrap resulting in faster signal decay[19]. As large and heavily modified protein ions

cannot be isotopically resolved, it is advisable to start with the lowest available resolution setting, which translates into a short transient detection time and maximum sensitivity. Increasing the resolution setting by one or two steps may help to resolve adducts or near-isobaric species. However, increasing the resolution setting will also result in decreased signal intensity. Thus, the lowest resolution setting results in the maximum intensity and sensitivity, whereas higher resolution settings will result in greater resolved species, close in m/z, but at the cost of decreased intensity.[20]

Füssl *et al*. showed, that the use of a higher resolution setting can benefit correct mass assignment with co-eluting near isobaric species, such as unmodified and adducted species of the same isoform[17].

In this study, a higher resolving power was used, to investigate the differentiation of two near isobaric substances which co-elute from the SCX stationary phase.

Figure 3.3.4 (A) shows the BPCs of adalimumab acquired at three different resolution settings. The main adalimumab CVA peak, highlighted in Figure 3.3.4 (A), corresponds to the antibody charge variant missing both C-terminal lysine residues of the Fc region. Mass spectra of the main variant peak are shown in (B), different charge states are colour coded and labelled, the charge state of the main isoform is +26, and (C) shows a magnification of the spectra for this charge state.

The two most dominant glycoforms were determined to be FA2/FA2 and FA2/FA2G1 (based on mass difference from empirical mAb). The lower abundant, FA2/FA2G1 glycoform, in the spectral magnification in Figure 3.3.4 (C) (m/z = 5,702.9), is near-isobaric with a co-eluting species which could not be sufficiently resolved at a resolution setting of 17,500. The insufficient resolving power affects the m/z value of this variant, as well as the average mass obtained after deconvolution, with a 56 ppm mass deviation from the theoretical mass. Increasing the resolution setting to 35,000 starts to resolve the peak pattern but with a small shoulder remaining, resulting in an improved mass

accuracy of 16 ppm. Resolution of 70,000 results in fully baseline resolved peaks and best mass accuracy of 4 ppm. Identification of the unknown, co-eluting isoform was further investigated, and is discussed in section 3.3.4.

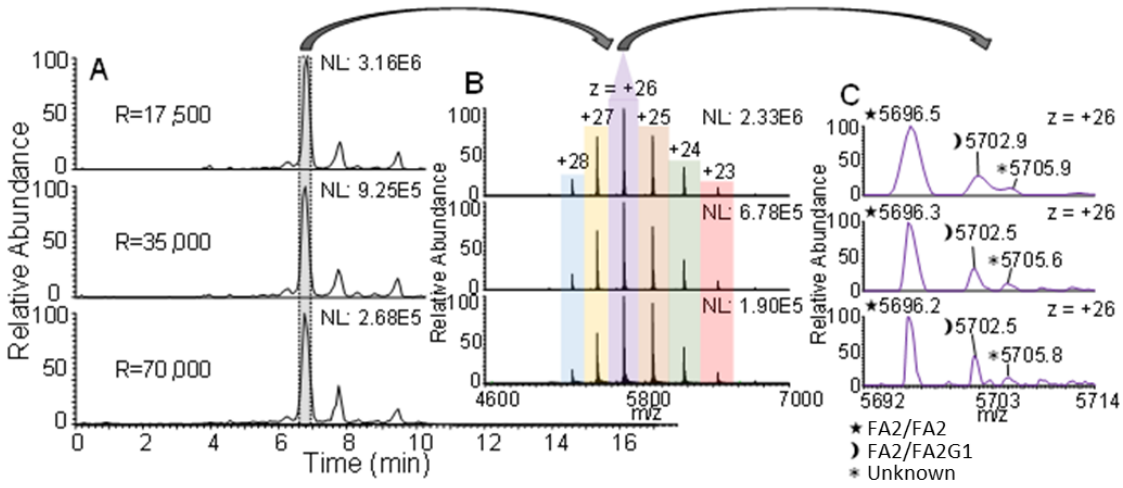


Figure 3.3.4: (A) BPC's of adalimumab at resolution settings 17,500, 35,000 and 70,000. (B) Mass spectra of the main variant peak. (C) Magnification of charge state +26, with labels of the three most abundant isoforms.

Conventional SCX would not have the resolving power to differentiate the co-eluting species seen here. The benefit of MS coupling to SCX is evident, particularly for the evaluation of unknown isoforms. Furthermore, it is shown that optimal method parameters for MS data acquisition are dependent on the sample and spectral complexity.

3.3.3. C-terminal Lysine clipping

The charge variant pattern of adalimumab is dominated by three different lysine variant isoforms[16]. C-terminal lysine isoforms derive from incomplete CpB cleavage in upstream processing and are among the most commonly observed basic variants in mAbs and related products[4]. As lysine is a basic residue, incomplete lysine clipping results in the generation of basic variants showing increased retention on a SCX stationary phase.

Using the optimised separation parameters, baseline resolution was achieved, of the three adalimumab isoforms, corresponding to the presence of zero, one and two C-terminal lysine residues. Figure 3.3.5 (A) shows the BPC of this separation. The Sliding Window deconvolution feature of the BioPharma Finder 2.0 software was used to deconvolute and annotate the three most abundant glycoforms of all three lysine variants. Mass deviations of less than 10 ppm were achieved and the Sliding Window feature also enabled a relative quantification of all antibody variants annotated, which can be observed in Figure 3.3.5 (C).

The BPC of a CpB digested sample is shown in Figure 3.3.5 (B). Enzyme digestion supported the annotations derived from intact molecular mass analysis relative to retention time. Following digestion with CpB there appeared to be a reduction in peak area of the two basic variants, however, small peaks remained. Interestingly, comparison with the intact mass data confirmed that both peaks were not attributable to unclipped C-terminal lysine but to the co-elution of different basic adalimumab variants. The identification of these species was achieved following more in-depth investigation of minor species, see section 3.3.4.

Additionally, when analysing the three most prominent peaks, an unknown variant was observed which co-elutes with the main species. This variant had been detected in the resolution study, Figure 3.3.5, at the same m/z value of 5705.5, which results in an average mass of 148,322.3 Da. A more detailed outline of the identification and quantification of lysine variants can be found in Table 3.3.1 below.

Chapter 3.

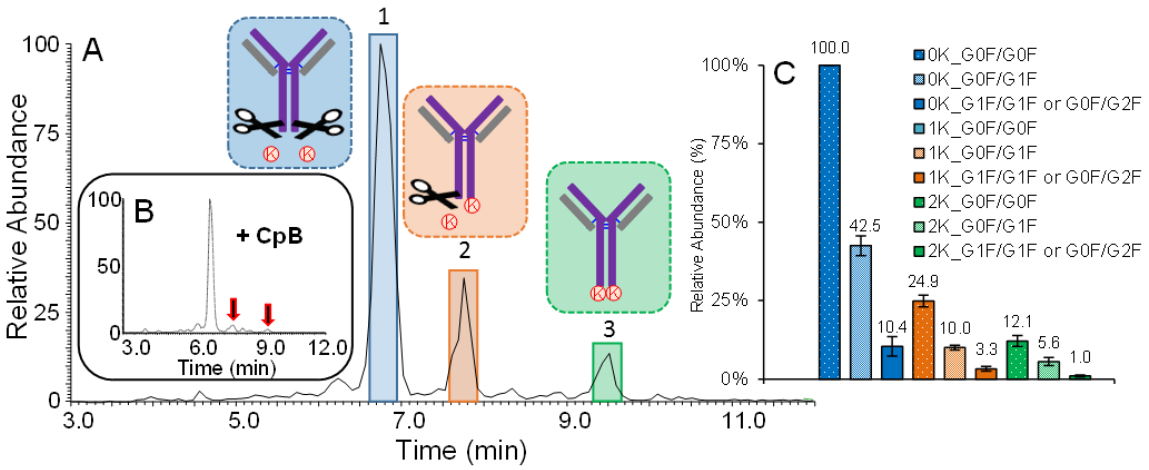


Figure 3.3.5: (A) BPC of adalimumab acquired at a resolution setting of 70,000. The main peak in blue corresponds to adalimumab without both C-terminal lysine residues, orange represents the variant containing one C-terminal lysine and green the variant carrying both C-terminal lysine residues. (B) BPC of CpB digested adalimumab. The arrows mark the positions where the lysine variants were present. (C) Relative abundance of the 3 most intense glycoforms of all 3 lysine variants identified.

Chapter 3.

Table 3.3.1: Summary of the abundance of the three major lysine variants with a combination of the three most common glycoforms.

Peak #	Run #	Variant	Theoretical Mass (Da)	Average mass (Da)	Δm (ppm)	Rel. Abundance (%)	RT Range (min)
1	1	0K FA2/FA2	148080.1	148079.8	-2.0	100.0	5.396 - 7.367
	2	0K FA2/FA2	148080.1	148079.5	-4.1	100.0	5.227 - 7.695
	3	0K FA2/FA2	148080.1	148080.5	2.7	100.0	5.395 - 7.440
	1	0K FA2G1/FA2G1	148242.2	148241.9	-2.0	44.4	5.311 - 7.283
	2	0K FA2/FA2G1	148242.2	148242.1	-0.7	44.2	5.564 - 7.619
	3	0K FA2/FA2G1	148242.2	148241.7	-3.4	38.9	4.974 - 7.440
	1	0K FA2G1/FA2G1	148404.3	148405.2	6.1	9.0	6.070 - 7.702
	2	0K FA2G1/FA2G1	148404.3	148404.6	2.0	14.0	6.070 - 8.266
	3	0K FA2G1/FA2G1	148404.3	148404.7	2.7	8.3	6.318 - 7.102
2	1	1K FA2/FA2	148208.2	148208.3	0.7	24.6	7.199 - 8.270
	2	1K FA2/FA2	148208.2	148208.7	3.4	26.9	7.366 - 8.182
	3	1K FA2/FA2	148208.2	148208.3	0.7	23.1	7.018 - 8.169
	1	1K FA2/FA2G1	148370.4	148369.4	-6.7	10.5	6.490 - 8.186
	2	1K FA2/FA2G1	148370.4	148370.7	2.0	10.4	7.451 - 8.097
	3	1K FA2/FA2G1	148370.4	148370.0	-2.7	9.2	6.933 - 8.085
	1	1K FA2G1/FA2G1	148532.5	148532.1	-2.7	4.2	7.030 - 8.101
	2	1K FA2G1/FA2G1	148532.5	148533.1	4.0	2.8	7.113 - 8.182
	3	1K FA2G1/FA2G1	148532.5	148531.8	-4.7	2.8	7.102 - 7.916
3	1	2K FA2/FA2	148336.4	148337.0	4.0	13.5	8.945 - 10.038
	2	2K FA2/FA2	148336.4	148336.2	-1.3	12.6	9.025 - 9.948
	3	2K FA2/FA2	148336.4	148336.5	0.7	10.2	9.097 - 9.846
	1	2K FA2/FA2G1	148498.5	148499.5	6.7	5.9	9.029 - 9.785
	2	2K FA2/FA2G1	148498.5	148498.9	2.7	6.7	9.025 - 9.863
	3	2K FA2/FA2G1	148498.5	148498.5	0.0	4.2	8.169 - 9.677
	1	2K_FA2G1/FA2G1	148660.7	148661.7	6.7	1.4	9.114 - 9.701
	2	2K FA2G1/FA2G1	148660.7	148662.1	9.4	0.7	9.194 - 9.695
	3	2K_FA2G1/FA2G1	148660.7	148660.1	-4.0	1.0	9.181 - 9.761

Chapter 3.

3.3.4. Other basic variants

Lysine content is not the only contributor to basic variation in the adalimumab CV profile. Figure 3.3.6 (A) shows a magnification of the adalimumab BPC, focusing on minor basic variants which are highlighted in either a blue or yellow box. For peaks 4, 5 and 6, blue boxes, deconvolution revealed a repetitive mass difference of +128 Da, in an increasingly basic order, similar to the elution pattern of the three most abundant peaks. This indicates the presence of three C-terminal lysine forms of an adalimumab variant already carrying a basic modification. This assumption was further supported by the disappearance of the two latter peaks upon CpB digestion, marked by the red "X" in the BPC of Figure 3.3.6 (B).

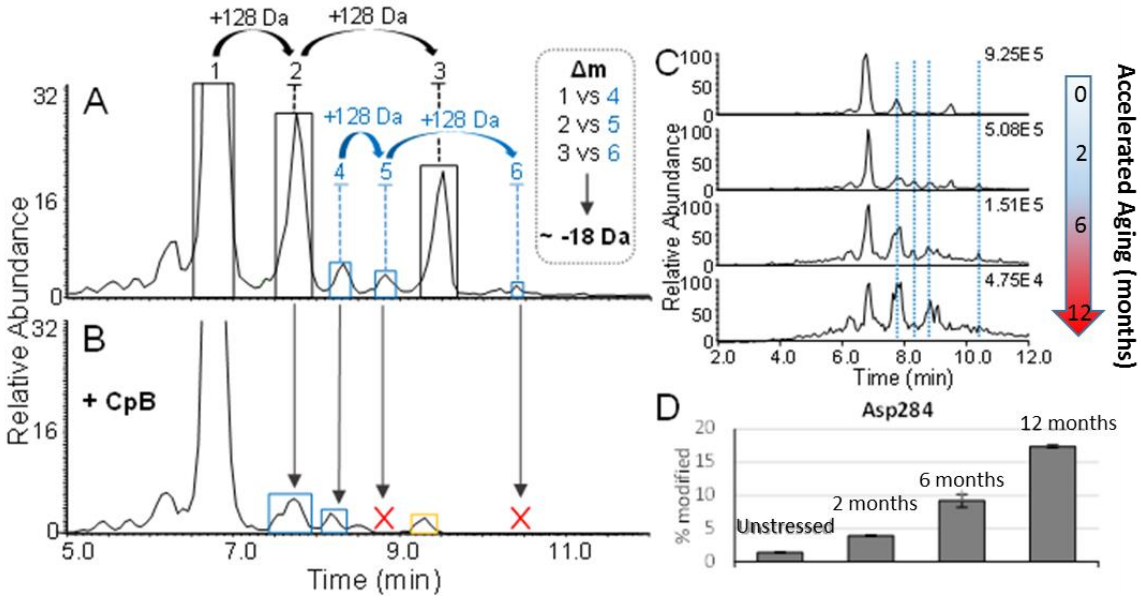


Figure 3.3.6: (A) Magnification of the BPC of an undigested and unstressed adalimumab sample. The peaks labelled from 1 – 3 are the main lysine variants, peaks 4, 5 &6 correspond to the succinimide Asp containing forms. (B) BPC of a CpB digested sample. Peak positions of CpB digested variants are indicated by a red "X". The co-eluting substance, suspected of being a basic variant derived by N-terminal Asp loss is indicated in yellow. (C.) BPC's of samples which were subject to accelerated aging. (D) Bar chart showing the increase of succinimide formation of Asn284 over time.

Interestingly, it was observed that; the antibody variant remaining, at the peak 2 position, after CpB digestion, was close in mass to the successive peak, labelled 4, blue box. Peak annotation was performed using the two most abundant glycoforms of each peak,

as the low abundance of these species prohibited reliable mass determination of the FA2G1/FA2G1 glycoform.

In comparison to the main lysine variants, which deviated by +128 Da, per Lys, it was found that the three-minor species deviated by -18 Da. Taking retention time into consideration, these species are likely to be three C-terminal lysine variants of an antibody, which underwent succinimide formation of an Asp residue. Table 3.3.2 shows the variant assignment, theoretical mass, average mass and Δ ppm for each peak, over triplicate injections. Subtraction of the average mass for each peak, 4, 5 & 6 from the average mass of their earlier eluting forms, 1, 2 & 3 respectively gives a difference of, on average -18 Da.

Table 3.3.2: Annotation of Asp succinimide containing adalimumab forms based on the two most abundant glycoforms per peak. Mass deviations >20 ppm are attributed to low signal intensity or near isobaric, co-eluting substances.

Peak #	Run #	Variant	Theoretical Mass (Da)	Average mass (Da)	Δm (ppm)	RT Range (min)
4	1	0K_FA2/FA2	148062.0	148063.2	8.1	8.187 - 8.386
	2	0K_FA2/FA2	148062.0	148063.7	11.5	8.218 - 8.398
	3	0K_FA2/FA2	148062.0	148063.3	8.8	8.212 - 8.424
	1	0K_FA2/FA2G1	148224.2	148224.6	2.7	8.187 - 8.386
	2	0K_FA2/FA2G1	148224.2	148227.1	19.6	8.218 - 8.398
	3	0K_FA2/FA2G1	148224.2	148224.9	4.7	8.212 - 8.424
5	1	1K_FA2/FA2	148190.2	148191.6	9.4	8.674 - 8.933
	2	1K_FA2/FA2	148190.2	148190.1	-0.7	8.717 - 8.945
	3	1K_FA2/FA2	148190.2	148191.7	10.1	8.689 - 9.000
	1	1K_FA2/FA2G1	148352.4	148352.0	-2.7	8.674 - 8.933
	2	1K_FA2/FA2G1	148352.4	148352.4	0.0	8.717 - 8.945
	3	1K_FA2/FA2G1	148352.4	148355.6	21.6	8.689 - 9.000
6	1	2K_FA2/FA2	148318.4	148322.3	26.3	10.276 - 10.410
	2	2K_FA2/FA2	148318.4	148318.8	2.7	10.326 - 10.478
	3	2K_FA2/FA2	148318.4	148320.6	14.8	10.330 - 10.519
	1	2K_FA2/FA2G1	148480.5	148479.1	-9.4	10.276 - 10.410
	2	2K_FA2/FA2G1	148480.5	148479.5	-6.7	10.326 - 10.478
	3	2K_FA2/FA2G1	148480.5	148480.5	0.0	10.330 - 10.519

Succinimide formation in mAbs has previously been reported to result in an increase in protein pI. It has also been positively correlated to exposure of the molecule to certain environmental conditions such as; acidic pH and elevated temperature[21,22,23]. This hypothesis is supported by the AAS experiments. After accelerated aging, an increase in succinimide content, on the intact level with ongoing incubation time was observed, this is illustrated in Figure 3.3.6 (C). Peptide mapping of the AAS samples revealed an increase in succinimide formation on multiple Asp residues over time. This is illustrated graphically in 3.3.6 (D), with the example of Asp284. Asp284, in particular, appears to be prone to succinimide formation with increasing storage duration. After simulated storage of 24 months at 4 °C almost 20% of the residue was modified. This not only suggests correct peak annotation of the succinimide forms on the intact level, but also Asp succinimide formation as a dominant degradation pathway of adalimumab under extended storage durations.

The identity of the species, co-eluting with the +2K C-terminal lysine variant, marked in yellow, in the basic region of Figure 3.3.6 (B), was further investigated. This peak is dominated by two species which are 112 Da – 113 Da lower than two of the main glycoform peaks, FA2/FA2 and FA2/FA2G1, the third glycoform FA2G1/FA2G1, with this observed difference, could not be identified.

Hydrolysis of Asp at the C-terminal, is one of the most commonly occurring mAb degradation pathways [24]. The mass shift observed, could be indicative of an N-terminal Asp loss which would cause a mass decrease of ~115 Da. The complete loss of an acidic amino acid, could also explain the relatively late elution. Succinimide formation can be as low as 1 Da, the 2-3 Da discrepancy makes an unambiguous assignment, based on intact mass, difficult. Peptide mapping data indicates the presence of a low abundant peptide which underwent N-terminal Asp loss, Figure 3.3.4.2. This is deduced by mass difference of y3 and y1 fragments (Δ 172.0488) less the mass of G (57.0214 Da).

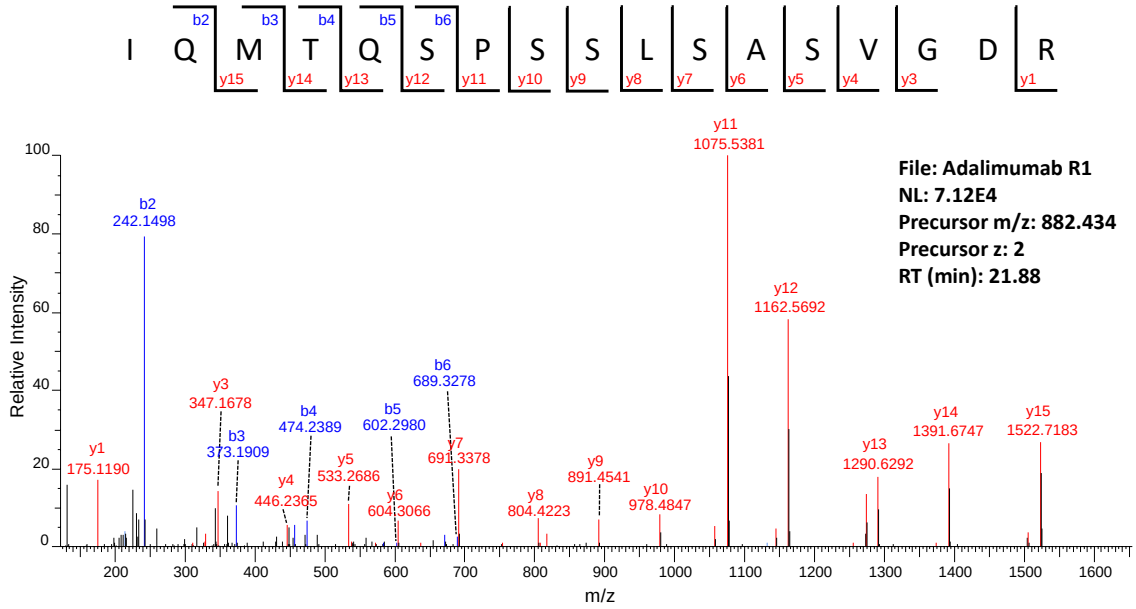


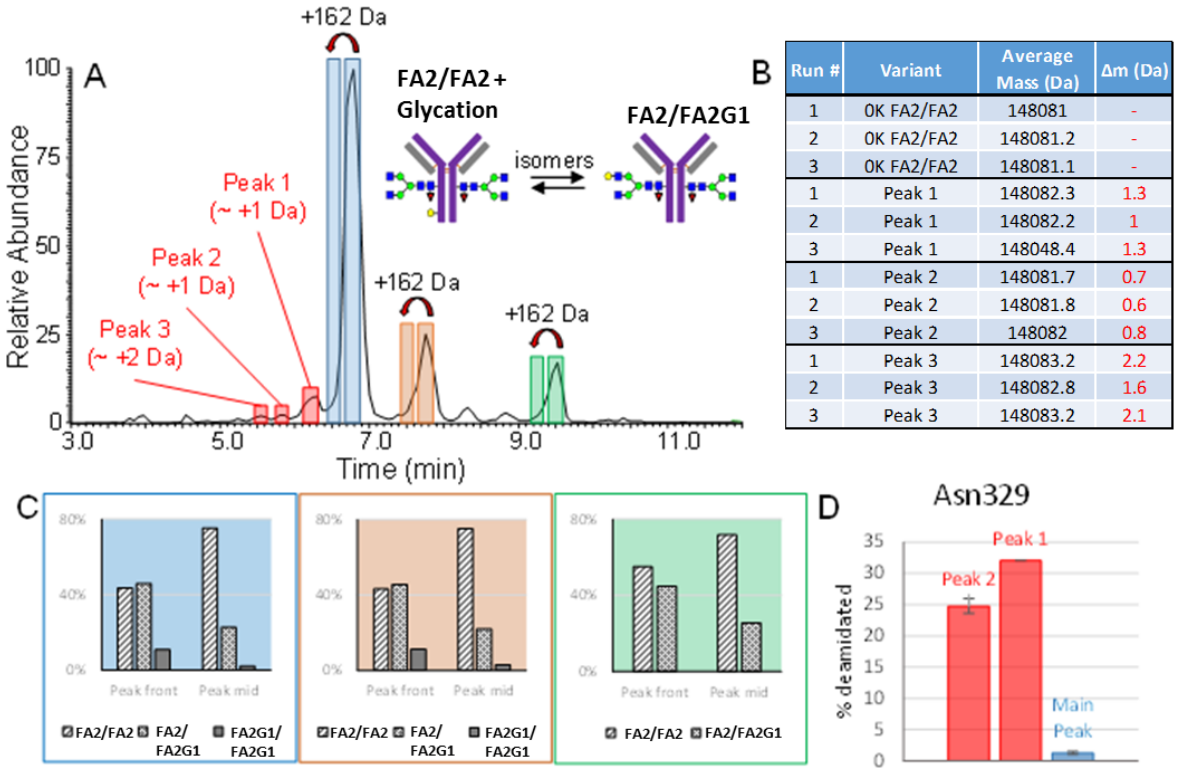
Figure 3.3.7: MS^2 spectrum of N-terminal peptide, from the antibody light chain, after aspartic acid loss.

3.3.5. Deamidation & isomerisation

Deamidation, next to sialylation, glycation and fragmentation are frequently occurring modifications on mAbs, leading to the generation of acidic species[4]. Deamidated and non-modified species can be distinguished from each other, by using adequate charge-sensitive separation techniques[22,17,25,26,27]. The standard method for identification of deamidated forms is peak collection after separation, followed by proteolytic digestion and peptide mapping [17].

In this study, chromatographic selectivity of adalimumab acidic species was high, which resulted in the clear resolution of three peaks, in the acidic region, in front of the main species. These peaks, highlighted in red in Figure 3.3.8 (A), are indicative of several acidic adalimumab charge variants, all of which, were found to be closely isobaric to the main species (Δm = +0.6 – 2.2 Da). Accounting for the relatively low abundance of the peaks, calculations are based on triplicate measurements of only the most abundant glycoform.

Chapter 3.



Run #	Variant	Average Mass (Da)	Δm (Da)
1	0K FA2/FA2	148081	-
2	0K FA2/FA2	148081.2	-
3	0K FA2/FA2	148081.1	-
1	Peak 1	148082.3	1.3
2	Peak 1	148082.2	1
3	Peak 1	148048.4	1.3
1	Peak 2	148081.7	0.7
2	Peak 2	148081.8	0.6
3	Peak 2	148082	0.8
1	Peak 3	148083.2	2.2
2	Peak 3	148082.8	1.6
3	Peak 3	148083.2	2.1

Figure 3.3.8: (A) BPC of adalimumab. The presence and retention times of glycated antibody forms is indicated by blue, orange and green labels, in peak fronts of all three major lysine variants. Antibody variants that were found to carry deamidation are indicated in red. (B) Table showing the mass differences between the main glycoform of the three acidic species and the major lysine variant (no C-terminal lysine residues and G0F/G0F glycoform). The average masses were experimentally observed. C.) Ratios of the three most abundant glycoforms, of all three major lysine variants, of the peak centres compared to peak fronts. D.) Asn329 deamidation of the main peak compared to acidic peaks 1 & 2. Mass difference tolerance was set at 5 Da in the deconvolution algorithm

Peak one in Figure 3.3.8 (A), was found to contain a species, which, on average is 1.2 Da larger than the main variant. This indicates the presence of a singly deamidated form. Peak two shows a similar mass, with an average difference of +0.7 Da, whereas peak three is +2.2 Da from the main variant. An overview of the mass differences observed is given in Figure 3.3.8 (B).

These results suggest the presence of singly and also possibly doubly deamidated adalimumab variants. The presence of two peaks, indicating single residue deamidation, could be explained by the formation of both, aspartate and isoaspartate which have previously been shown to be distinguishable *via* charge sensitive separation methods[26,22,23]. The presence of deamidated forms is not only suggested on the intact

level but also confirmed *via* peptide mapping of the two most abundant acidic variants after preparative peak collection. Compared to the main peak, these charge variants showed a total incident of Asn329 deamidation of ~31% and ~25%, respectively. The graphical representation of this data is shown in Figure 3.3.8 (D).

3.3.6. Glycation

Protein glycation is a non-enzymatic addition of a hexose unit onto a lysine residue within the protein sequence and occurs when reducing sugars interact with the protein[28,29,30]. Glycation does not directly impact drug potency, but it has been linked to an increased incidence of aggregation[31,28]. In a study where a mixture of glycated and non-glycated mAb, as well as the separated fractions were analysed by CEX chromatography, it was shown that glycation resulted in only a slight retention time shift towards the acidic region. The lack of selectivity may not be sufficient to distinguish glycated and main variants as two distinct peaks[31]. Using CVA-MS, however, it was possible to partially overcome this limitation, by integration and deconvolution of different regions of a single peak.

The ratios between the three main glycoforms; FA2/FA2, FA2/FA2G1 and FA2G1/FA2G1 were found to change from peak front to peak mid, from 10:10:3 to 10:4:1, the latter, would be the only ratio expected to be seen, across all prominent charge variant peaks. Figure 3.3.8 (C) shows the variation in abundance of these glycoforms across the three highlighted peaks. The presence of an additional galactose residue on an N-glycan does not impact protein pI, and hence should not result in retention time shifts on a CEX stationary phase. Due to the alteration in the ratio of glycoforms, over the chromatographic peak it is more probable that the shift in retention is due to the presence of glycated species, of the three major C-terminal lysine variants. Table 3.3.3, provides more detail on the change in abundance of glycoforms, within the three lysine variant peaks.

Table 3.3.3: Comparison of the relative abundance of the three most abundant glycoforms of the peak fronts and peak centres of the three lysine variants.

	Variant	Theoretical Mass (Da)	Average Mass (Da)	Δm (ppm)	Relative Abundance (%)	Fractional Abundance (%)	RT Range (min)
Peak Mid	0K FA2/FA2	148080.1	148081.2	7.4	100.0	61.2	6.650 - 6.950
	0K FA2/FA2G1	148242.2	148240.0	-14.8	30.1	18.4	6.650 - 6.950
	0K FA2G1/FA2G1	148404.3	148409.3	33.7	2.8	1.7	6.650 - 6.950
	1K FA2/FA2	148208.2	148210.2	13.5	100.0	50.0	7.615 - 7.907
	1K FA2/FA2G1	148370.4	148371.9	10.1	29.2	14.6	7.615 - 7.907
	1K FA2G1/FA2G1	148532.5	148537.8	35.7	3.8	1.9	7.615 - 7.907
	2K FA2/FA2	148336.4	148337.6	8.1	100.0	48.2	9.372 - 9.603
	2K FA2/FA2G1	148498.5	148498.2	-2.0	35.6	17.1	9.372 - 9.603
Peak Front	0K FA2/FA2	148080.1	148081.3	8.1	95.0	23.1	6.403 - 6.617
	0K FA2G1/FA2	148242.2	148242.6	2.7	100.0	24.3	6.403 - 6.617
	0K FA2G1/FA2G1	148404.3	148408.9	31.0	23.2	5.6	6.403 - 6.617
	1K FA2/FA2	148208.2	148209.8	10.8	94.7	16.5	7.330 - 7.586
	1K FA2G1/FA2	148370.4	148372.7	15.5	100.0	17.4	7.330 - 7.586
	1K FA2G1/FA2G1	148532.5	148531.1	-9.4	24.7	4.3	7.330 - 7.586
	2K FA2/FA2	148336.4	148337.7	8.8	100.0	20.8	9.078 - 9.314
	2K FA2/FA2G1	148498.5	148499.4	6.1	81.4	16.9	9.078 - 9.314

3.3.7. Degradation Products

The main focus of this study was to investigate mAb heterogeneity on the intact protein level by native on-line LC-MS. The acquisition parameters employed for MS analysis were chosen to fully capture the charge envelope of a native mAb, utilising the mass range of up to 8,000 m/z provided by the MS instrumentation.

To also visualize potentially present lower molecular weight species, such as antibody sub-units, the data files acquired were filtered, post-acquisition, to only display ions within a mass range of m/z 2,500 – 5,000. Figure 3.3.9 (A) shows a BPC resulting from

application of the mass filter. This represents the same data shown in previous figures but with a mass filter applied to exclude all ionic species of the intact antibody. Three peaks labelled in red, in Figure 3.3.9 (A) were found to contain four charge state distributions, typical for proteins. Deconvolution of the mass spectra revealed that the peaks contain species of either ~48 Kilodaltons (kDa) or ~100 kDa, suggesting the presence of antibody fragments. The outset of (A) is the filtered mass spectra obtained upon averaging of each of the three peaks, the charge state envelopes of each can be clearly seen. Mab sub-units occurring due to the same degradation events are illustrated by matching colour (Peak 1 & 3; blue and peak 2 purple). The appearance of the charge envelope of the intact mAb, in the averaged spectrum of peak 3, can be explained by the simultaneous elution of the sub-unit and acidic, intact antibody variants, which elute relatively early.

Accelerated aging of the drug product resulted in extensive degradation. Figure 3.3.9 (C) compares the BPC of a sample at 0 months compared to a sample aged to 36 months. The sub-units of 48 kDa and 100 kDa, peaks 1 & 3, decrease over time until they have completely disappeared, what is indicated by the red “X”. However, new isoforms occurred with mass shifts of -18 Da indicating repeating isoforms with extensive succinimide formation. The retention time shift also supports the mass evidence observed, as the formation of succinimide results in a more basic pI. Succinimide formation on sub-units after accelerated aging studies has been reported before[32].

Chapter 3.

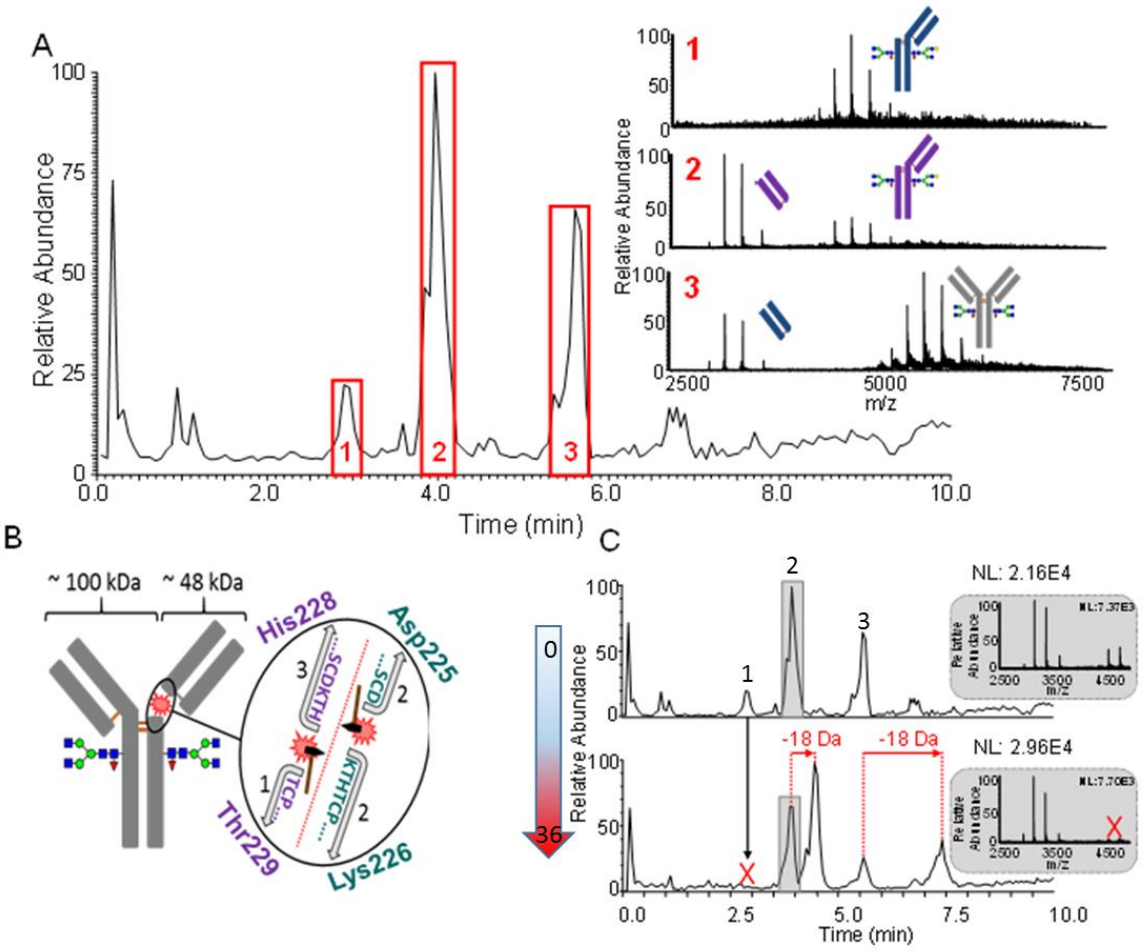


Figure 3.3.9: (A) Adalimumab BPC acquired at a resolution setting of 35,000 adjusted to a mass range of m/z 2,500 to 5,000. (B) Fragmentation sites and resulting species are shown on the intact mAb as well as by a magnification of the upper hinge region. (C) BPC's showing the mass range of m/z 2,500 – 5,000 of the unstressed drug product and the drug product after accelerated aging.

Monoclonal antibody degradation can occur at different stages of the production process, such as upstream and downstream processing, formulation or storage. Numerous possible pathways for non-enzymatic protein cleavage have been described in the literature[24]. Mechanisms for protein cleavage have increasingly been studied, in recent years, especially in relation to residual host cell proteins causing degradation of biopharmaceuticals[33,34,35,36]. Cleavage of the molecule can result in loss of potency, and represents a critical quality attribute.

Utilising molecular mass information after deconvolution, it was found that the major fragment masses, when combined, make up the total antibody mass of ~148,080 Da, indicating two distinct cleavage sites in adalimumab. The size ranges of the sub-units observed, provided valuable information regarding annotation of the cleavage sites. Breaking of disulphide bonds could not account for the cleavage observed as masses resulting from such would be ~25 kDa, ~75 kDa or ~125 kDa in mass and so were ruled out. Accounting for the fact that the higher molecular weight species were found as different glycoforms, it was hypothesised that cleavage probably occurs in the upper hinge region, as illustrated in Figure 3.3.9 (B). By comparing theoretical and observed sub-unit masses, the two cleavage sites were distinguished to be C-terminal to Asp225 and His228. Table 3.3.4, below, summarises the masses calculated for higher molecular weight sub-units (HMWF) and lower molecular weight fragments (LMWF).

Table 3.3.4: Adalimumab sub-unit assignment.

Variant	Theoretical Mass (Da)	Average Mass (Da)	Δm (ppm)	RT Range (min)
HMWF1 (0K FA2/FA2)	100417.3	100417.9	6.0	2.798 - 3.064
HMWF2 (0K FA2/FA2)	100783.8	100784.2	4.0	3.806 - 4.177
LMWF1	47314.3	47314.1	-4.2	3.806 - 4.177
LMWF2	47680.7	47680.4	-6.3	5.333 - 5.736

The upper hinge region has repeatedly been described as a region prone to non-enzymatically driven degradation events[37,38,39,32]. Cleavage has been shown to be mediated either; by direct hydrolysis or by β-elimination. Specific cleavage sites have been shown to be heavily influenced by the pH of the protein environment[39]. The two distinct cleavage sites, which are reported in this study, have previously been detected on an IgG1 antibody and reported by Tao et.al.[32]. This finding suggests that these sites may be prone to hydrolysis in IgG1 molecules, in general. To investigate whether cleavage at these two distinct sites could potentially be related to protease activity, a peptide mapping experiment on the drug product was performed. In this experiment, the

column was overloaded, to enable the detection of low abundant HCP's. Fifty-one different CHO HCP's were identified, six of them by two or more unique peptides. Table 3.3.5 shows the six most probable HCP's identified in this study. The specificity of each is given in the description.

Table 3.3.5: List of CHO proteins found by at least two unique peptides in a peptide mapping based HCP search.

Accession	# of Unique Peptides	Confidence Score	Description	ppm (ng HCP/mg therapeutic protein)
tr|G3GR64	2	112.8	Inter-alpha-trypsin inhibitor heavy chain H5 OS=Cricetulus Griseus GN=I79_000007 PE=4 SV=1	3.2
tr|G3HUA1	2	110.59	Cationic trypsin-3 OS=Cricetulus Griseus GN=I79_014509 PE=4 SV=1	32.7
tr|G3HUU6	3	286.82	Protein S100-A11 OS=Cricetulus Griseus GN=I79_014714 PE=4 SV=1	4.5
tr|G3IHL5	2	110.57	Zinc finger and SCAN domain-containing protein 2 OS=Cricetulus Griseus GN=I79_023307 PE=4 SV=1	0.8
tr|G3IMJ0	2	228.5	Putative uncharacterized protein (Fragment) OS=Cricetulus Griseus GN=I79_025130 PE=4 SV=1	136.5
tr|G3INC5	3	207.4	Cathepsin L1 OS=Cricetulus Griseus GN=I79_025440 PE=3 SV=1	3.9

Cathepsin L1 was chosen for further investigation as it shows specificity for one of the cleavage sites reported herein, according to the MEROPS database[40,41] Three unique peptides were found for Cathepsin L, the associated MS^2 spectra with assigned b- and y-ions are shown in Figure 3.3.10.

Chapter 3.

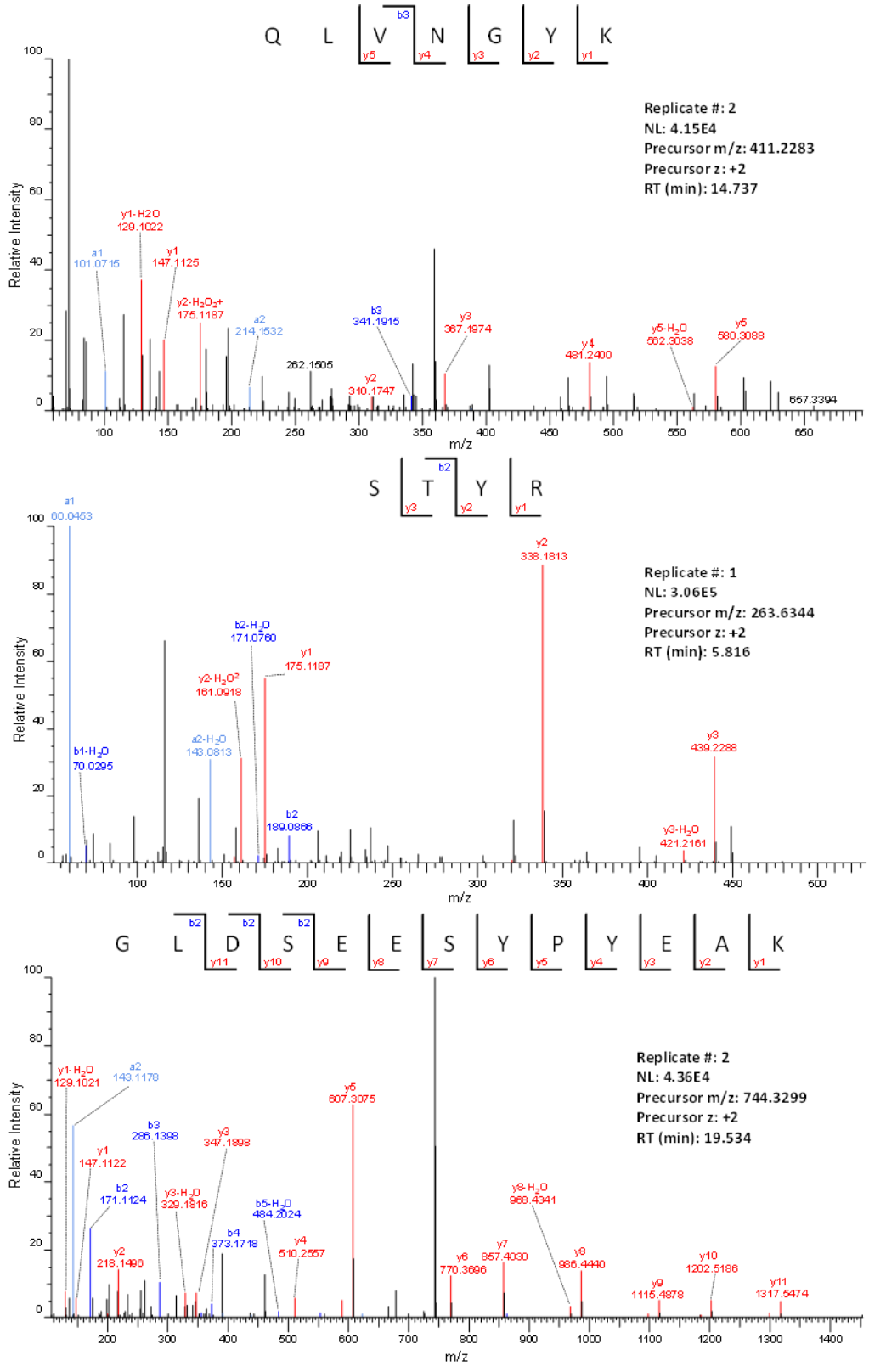


Figure 3.3.10: Fragment spectra of Cathepsin L, found in adalimumab.

The shortest peptide ‘STYR’ was found to also be present in the sequence of adalimumab; however tryptic digestion does not yield this peptide. Therefore, it should only occur due to HCP digestion and not from the drug substance. Using label free quantitation, the concentration of Cathepsin L in the drug product as found to be in the range of ~4 ppm (ng HCP per mg therapeutic protein). Considering its specificity for one of the cleavage sites identified, the presence of detectable amounts of Cathepsin L could be crucial for drug stability. This may be the case especially under non-ideal storage conditions where the enzymatic activity of the protease can be dramatically increased.

A majority of studies on antibody degradation cited here were based on peptide mapping or middle up approaches after peak collection by SEC. In comparison to these approaches the advantage of the CVA-MS method for sub-unit characterisation is a comparable depth of information, while avoiding multiple analyses and sample preparation procedures. This makes CVA-MS an incredibly powerful method for fast and in depth analysis of monoclonal antibodies.

Chapter 3.

3.4. Conclusion & Future directions

The work by Füssl *et al.* provided a strong basis for the study undertaken here. In the aforementioned study, the main C-terminal lysine variants and glycoforms were detected for a number of commercial drug products. The underlying contributors to charge variants had not previously been investigated using this method. However, in this chapter, the capabilities of the method were further tested using a combination of post-acquisition analysis and further MS analysis. A multitude of physical aspects of mAbs that would have previously required extensive sample preparation and numerous procedures can now be detected using one method, one injection with no sample prep. Building on this method, the formation of adducts was lessened and the universal applicability was broadened by optimisation of the buffer system and the separation procedure. Adalimumab was chosen as a model mAb to show that gradient optimisation can result in a plethora of charge variant peaks chromatographically resolvable. The superiority of the method used in this chapter is evident upon comparison with other reported methods of IEC-MS hyphenation. Talebi *et al*. (2013) had successful produced a volatile buffer system; however, the chromatographic resolution was not as sufficient the method reported here. Only one major glycoform was identified.

MS resolution settings can be selected to best suit experimental requirements, which is evident from the obtained MS data of even very low abundant charge variant species. Using molecular mass and retention time information, more than 16 different adalimumab charge variants were identified, some of which were not sufficiently chromatographically distinguishable but were detectable by mass spectrometric means. Using orthogonal analysis such as; CpB enzyme digestion, peptide mapping and forced degradation experiments confident annotation could be achieved. Among the modifications identified were; various levels of lysine truncation, glycation and succinimide formation of the differently truncated forms, deamidation, isomerisation as well as double deamidation, N-terminal Asp loss and fragmentation at two distinct sites.

The utility of CVA-MS for the analysis of degraded antibody samples was also evident, as well as the identification of the degradation pathways by piecing together the wealth of information provided by this method. It was found that, fragmentation and Asp succinimide formation to be major degradation pathways of adalimumab over long term storage. Additionally, highly compelling leads for the presence of low amounts of Cathepsin L within the drug product were observed, likely a serious contributor to antibody degradation under non-ideal storage conditions. Muneeruddin *et al.* (2015) demonstrated the usefulness of IEC-MS for analysing intact and reduced protein samples. In that work, the analysts completely reduced a protein sample prior to analysis and could accurately detect them, as they eluted earlier. In this chapter, the antibody sub-units found were not expected and came from thorough data analysis and high quality native separations. From a single injection without any sample preparation, significant amounts of data can be obtained, which would normally require multiple modes of analysis with several sample preparation techniques. These marks CVA-MS as a true multi attribute monitoring (MAM) technique.

Future work with this method will aim to achieve further characterisation of antibody variants, by extensive characterisation, through traditional means of in-house produced biosimilars and reference drug products. Using the information acquired, by this extensive testing, the MS data interpretation of the CVA-MS method will be fully informed on the identity of variants.

3.5. Author Contributions

Adapted from Füssl *et al.,* Comprehensive Characterisation of the Heterogeneity of Adalimumab *via* Charge Variant Analysis Hyphenated On-Line to Native High Resolution Orbitrap Mass Spectrometry. Under review for publication

Research study devised by Florian Füssl, Jonathan Bones and Anne Trappe; Sample preparation performed by Anne Trappe; research performed by Florian Füssl; data

analysis performed by Florian Füssl; Chapter adapted and written by Anne Trappe and reviewed by Jonathan Bones.

Chapter 3.

3.6. References

1.Berkowitz, S. A.; Engen, J. R.; Mazzeo, J. R.; Jones, G. B., Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. *Nature Reviews Drug Discovery* **2012,** *11*, 527.

2.Liu, H.; Gaza-Bulseco, G.; Faldu, D.; Chumsae, C.; Sun, J., Heterogeneity of monoclonal antibodies. *J Pharm Sci* **2008,** *97* (7), 2426-47.

3.Torkashvand, F.; Vaziri, B., Main Quality Attributes of Monoclonal Antibodies and Effect of Cell Culture Components. *Iranian biomedical journal* **2017,** *21* (3), 131-141.

4.Du, Y.; Walsh, A.; Ehrick, R.; Xu, W.; May, K.; Liu, H., Chromatographic analysis of the acidic and basic species of recombinant monoclonal antibodies. *MAbs* **2012,** *4* (5), 578-85.

5.Fekete, S.; Beck, A.; Veuthey, J. L.; Guillarme, D., Ion-exchange chromatography for the characterization of biopharmaceuticals. *J Pharm Biomed Anal* **2015,** *113*, 43-55.

6.Fekete, S.; Beck, A.; Fekete, J.; Guillarme, D., Method development for the separation of monoclonal antibody charge variants in cation exchange chromatography, Part I: Salt gradient approach. *J.Pharm.Biomed.Anal.* **2015,** *102*, 33-44.

7.Fekete, S.; Beck, A.; Fekete, J.; Guillarme, D., Method development for the separation of monoclonal antibody charge variants in cation exchange chromatography, Part II: pH gradient approach. *J.Pharm.Biomed.Anal.* **2015,** *102*, 282-289.

8.Patel, B. A.; Pinto, N. D. S.; Gospodarek, A.; Kilgore, B.; Goswami, K.; Napoli, W. N.; Desai, J.; Heo, J. H.; Panzera, D.; Pollard, D.; Richardson, D.; Brower, M.; Richardson, D. D., On-Line Ion Exchange Liquid Chromatography as a Process Analytical Technology for Monoclonal Antibody Characterization in Continuous Bioprocessing. *Analytical Chemistry* **2017,** *89* (21), 11357-11365.

9.Griaud, F.; Denefeld, B.; Lang, M.; Hensinger, H.; Haberl, P.; Berg, M., Unbiased in-depth characterization of CEX fractions from a stressed monoclonal antibody by mass spectrometry. *MAbs* **2017,** *9* (5), 820-830.

10.Hao, P.; Ren, Y.; Alpert, A. J.; Sze, S. K., Detection, Evaluation and Minimization of Nonenzymatic Deamidation in Proteomic Sample Preparation. *Molecular & Cellular Proteomics : MCP* **2011,** *10* (10), O111.009381.

11.Muneeruddin, K.; Bobst, C. E.; Frenkel, R.; Houde, D.; Turyan, I.; Sosic, Z.; Kaltashov, I. A., Characterization of a PEGylated protein therapeutic by ion exchange chromatography with on-line detection by native ESI MS and MS/MS. *Analyst* **2017,** *142* (2), 336-344.

12.Muneeruddin, K.; Nazzaro, M.; Kaltashov, I. A., Characterization of intact protein conjugates and biopharmaceuticals using ion-exchange chromatography with online detection by native electrospray ionization mass spectrometry and top-down tandem mass spectrometry. *Anal Chem* **2015,** *87* (19), 10138-45.

13.Talebi, M.; Nordborg, A.; Gaspar, A.; Lacher, N. A.; Wang, Q.; He, X. Z.; Haddad, P. R.; Hilder, E. F., Charge heterogeneity profiling of monoclonal antibodies using low ionic strength ion-exchange chromatography and well-controlled pH gradients on monolithic columns. *Journal of chromatography. A* **2013,** *1317*, 148-54.

14.Stoll, D. R.; Hannes, D. C.; Danforth, J.; Wagner, E.; Guillarme, D.; Fekete, S.; Beck, A., Direct Identification of Rituximab Main Isoforms and Subunit Analysis by Online Selective Comprehensive Two-Dimensional Liquid Chromatography-Mass Spectrometry. *Analytical Chemistry* **2015,** *87* (16), 8307-8315.

15.Tebbey, P. W.; Varga, A.; Naill, M.; Clewell, J.; Venema, J., Consistency of quality attributes for the glycosylated monoclonal antibody Humira (R) (adalimumab). *Mabs* **2015,** *7* (5), 805-811.

16.Liu, J.; Eris, T.; Li, C.; Cao, S.; Kuhns, S., Assessing Analytical Similarity of Proposed Amgen Biosimilar ABP 501 to Adalimumab. *BioDrugs* **2016,** *30* (4), 321-38.

17.Fussl, F.; Cook, K.; Scheffler, K.; Farrell, A.; Mittermayr, S.; Bones, J., Charge Variant Analysis of Monoclonal Antibodies using Direct Coupled pH Gradient Cation Exchange Chromatography to High Resolution Native Mass Spectrometry. *Anal Chem* **2018**.

18.Goyon, A.; Excoffier, M.; Janin-Bussat, M. C.; Bobaly, B.; Fekete, S.; Guillarme, D.; Beck, A., Determination of isoelectric points and relative charge variants of 23 therapeutic monoclonal antibodies. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2017,** *1065-1066*, 119-128.

19.Shaw, J. B.; Brodbelt, J. S., Extending the Isotopically Resolved Mass Range of Orbitrap Mass Spectrometers. *Analytical Chemistry* **2013,** *85* (17), 8313-8318.

20.Makarov, A.; Denisov, E., Dynamics of ions of intact proteins in the Orbitrap mass analyzer. *J Am Soc Mass Spectrom* **2009,** *20* (8), 1486-95.

21.Chu, G. C.; Chelius, D.; Xiao, G.; Khor, H. K.; Coulibaly, S.; Bondarenko, P. V., Accumulation of succinimide in a recombinant monoclonal antibody in mildly acidic buffers under elevated temperatures. *Pharm Res* **2007,** *24* (6), 1145-56.

22.Ponniah, G.; Nowak, C.; Neill, A.; Liu, H., Characterization of charge variants of a monoclonal antibody using weak anion exchange chromatography at subunit levels. *Anal Biochem* **2017,** *520*, 49-57.

23.Harris, R. J.; Kabakoff, B.; Macchi, F. D.; Shen, F. J.; Kwong, M.; Andya, J. D.; Shire, S. J.; Bjork, N.; Totpal, K.; Chen, A. B., Identification of multiple sources of charge heterogeneity in a recombinant antibody. *Journal of Chromatography B: Biomedical Sciences and Applications* **2001,** *752* (2), 233-245.

24.Vlasak, J.; Ionescu, R., Fragmentation of monoclonal antibodies. *MAbs* **2011,** *3* (3), 253-63.

25.Ponniah, G.; Kita, A.; Nowak, C.; Neill, A.; Kori, Y.; Rajendran, S.; Liu, H., Characterization of the acidic species of a monoclonal antibody using weak cation exchange chromatography and LC-MS. *Anal Chem* **2015,** *87* (17), 9084-92.

26.Neill, A.; Nowak, C.; Patel, R.; Ponniah, G.; Gonzalez, N.; Miano, D.; Liu, H., Characterization of Recombinant Monoclonal Antibody Charge Variants Using OFFGEL Fractionation, Weak Anion Exchange Chromatography, and Mass Spectrometry. *Anal Chem* **2015,** *87* (12), 6204-11.

27.Gandhi, S.; Ren, D.; Xiao, G.; Bondarenko, P.; Sloey, C.; Ricci, M. S.; Krishnan, S., Elucidation of degradants in acidic peak of cation exchange chromatography in an IgG1 monoclonal antibody formed on long-term storage in a liquid formulation. *Pharm Res* **2012,** *29* (1), 209-24.

28.Banks, D. D.; Hambly, D. M.; Scavezze, J. L.; Siska, C. C.; Stackhouse, N. L.; Gadgil, H. S., The effect of sucrose hydrolysis on the stability of protein therapeutics during accelerated formulation studies. *J Pharm Sci* **2009,** *98* (12), 4501-10.

29.Fischer, S.; Hoernschemeyer, J.; Mahler, H. C., Glycation during storage and administration of monoclonal antibody formulations. *European journal of pharmaceutics and biopharmaceutics : official journal of Arbeitsgemeinschaft fur Pharmazeutische Verfahrenstechnik e.V* **2008,** *70* (1), 42-50.

30.Gadgil, H. S.; Bondarenko, P. V.; Pipes, G.; Rehder, D.; McAuley, A.; Perico, N.; Dillon, T.; Ricci, M.; Treuheit, M., The LC/MS analysis of glycation of IgG molecules in sucrose containing formulations. *J Pharm Sci* **2007,** *96* (10), 2607-21.

31.Quan, C.; Alcala, E.; Petkovska, I.; Matthews, D.; Canova-Davis, E.; Taticek, R.; Ma, S., A study in glycation of a therapeutic recombinant humanized monoclonal antibody: where it is, how it got there, and how it affects charge-based behavior. *Anal Biochem* **2008,** *373* (2), 179-91.

32.Xiang, T.; Lundell, E.; Sun, Z.; Liu, H., Structural effect of a recombinant monoclonal antibody on hinge region peptide bond hydrolysis. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2007,** *858* (1-2), 254-62.

33.Gao, S. X.; Zhang, Y.; Stansberry-Perkins, K.; Buko, A.; Bai, S.; Nguyen, V.; Brader, M. L., Fragmentation of a highly purified monoclonal antibody attributed to residual CHO cell protease activity. *Biotechnol Bioeng* **2011,** *108* (4), 977-82.

34.Robert, F.; Bierau, H.; Rossi, M.; Agugiaro, D.; Soranzo, T.; Broly, H.; Mitchell-Logean, C., Degradation of an Fc-fusion recombinant protein by host cell proteases:

Chapter 3.

Identification of a CHO cathepsin D protease. *Biotechnol Bioeng* **2009,** *104* (6), 1132-41.

35.Wang, W.; Ignatius, A. A.; Thakkar, S. V., Impact of residual impurities and contaminants on protein stability. *J Pharm Sci* **2014,** *103* (5), 1315-30.

36.Bee, J. S.; Tie, L.; Johnson, D.; Dimitrova, M. N.; Jusino, K. C.; Afdahl, C. D., Trace levels of the CHO host cell protease cathepsin D caused particle formation in a monoclonal antibody product. *Biotechnology progress* **2015,** *31* (5), 1360-9.

37.Cohen, S. L.; Price, C.; Vlasak, J., Beta-elimination and peptide bond hydrolysis: two distinct mechanisms of human IgG1 hinge fragmentation upon storage. *J Am Chem Soc* **2007,** *129* (22), 6976-7.

38.Cordoba, A. J.; Shyong, B. J.; Breen, D.; Harris, R. J., Non-enzymatic hinge region fragmentation of antibodies in solution. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2005,** *818* (2), 115-21.

39.Gaza-Bulseco, G.; Liu, H., Fragmentation of a recombinant monoclonal antibody at various pH. *Pharm Res* **2008,** *25* (8), 1881-90.

40.Rawlings, N. D.; Barrett, A. J.; Bateman, A., MEROPS: the peptidase database. *Nucleic Acids Res* **2010,** *38* (Database issue), D227-33.

41.Rawlings, N. D., Peptidase specificity from the substrate cleavage collection in the MEROPS database and a tool to measure cleavage site conservation. *Biochimie* **2016,** *122*, 5-30.

Chapter 4.

4. In-depth characterisation of in-house produced cetuximab variants using multiple analytic approaches

4.1. Introduction

Cetuximab is a human/mouse chimeric antibody used for the treatment of metastatic cancers. It received market approval in 2004 for the treatment of metastatic colorectal cancer and in 2006 it received approval for the treatment of head and neck cancer[1]. Epidermal growth factor receptor (EGFR) is the binding target for cetuximab. Cetuximab was the first chimeric anti-EGFR antibody to receive market approval.

EGFR is a member of the erbB family and is related to other receptor tyrosine kinases such as HER2. EGFR is a transmembrane protein, which, when bound to the EGF ligand activates tyrosine kinase activity, which results in the phosphorylation of Tyr at the C-terminal of EGFR, this signals downstream activation of a number of proteins which control DNA synthesis, such MAPK and JNK[2]. The activity of EGFR ultimately results in the uncontrolled proliferation of cells which have a phenotype that readily migrates. One of the benefits of cetuximab is a ligand binding affinity 5 – 10 times that of EGFR's natural binding partners[3]. This is often referred to as cetuximab's "internal" activity[4]. However, it has been observed that the signalling cascade which occurs via EGFR binding can occur independently through the mutation of the KRAS gene, which also regulates cell proliferation. When KRAS functions normally it provides a molecular on/off switch for proliferation. However, in almost 40% of colorectal cancer patients there is a mutation[5]. Therefore, genetic typing prior to cetuximab treatment is required.

Cetuximab has a second activity or its "external" activity[4]. As an IgG mAb it has antibody dependent cell-mediated cytotoxicity (ADCC) effector function on Natural Killer (NK) cells in the immune system. Upon binding to the EGFR receptor on the tumour cell the Fc portion of cetuximab faces out to the micro-environment of the tumour. NK cells have Fc receptors; typically CD16 or FcγRIII, which bind to cetuximab's Fc fragment and elicit the NK cell to secrete cytotoxic compounds[4]. It is this secondary mechanism which is a main

focus of biosimilar development[6]. As almost 40% of targeted patients cannot receive cetuximab enhancing the ability of the mAb to elicit ADCC activity is of high importance.

Fc glycosylation has been shown to have a significant impact on the elicitation of ADCC activity. In de-glycosylated isoforms there is little to no ADCC activity[7]. With glyco-engineering becoming more prevalent in biopharma development, biosimilar developers are looking at making glyco-better forms of cetuximab[6]. However, this is not solely due to ADCC activity. Approximately 3.4% of patients receiving cetuximab developed human anti chimeric antibodies (HACA) against the drug which resulted in an anaphylactic rash[3]. The cause of the HACA was found, to be due to, the murine glycoform α1-3 galactose, α-Gal[8]. This discovery highlighted the importance of glycosylation as a PTM for both function and immunogenicity. Cetuximab has three potential glycosylation sites; Asn88 and Asn 299 on the heavy chain, variable and constant regions, respectively and the third, Asn41 on the light chain variable region. Only Asn88 and 299 are reported to be glycosylated[9,10,1].

In this chapter, the glycosylation profiles of two in house produced cetuximab variants are determined. The mAbs were produced in HEK293 and CHO cells, two of the widely used cell lines for biopharmaceutical production, with HEK predominantly in pilot scale and CHO for larger, stable transfection. These cell lines used produce human like glycosylation and are grafted on to human IgG structures, only having the cetuximab CDR sequence in common with the actual reference product.

This study involves the full characterisation of these mAbs; including charge variant analysis by SCX, glycan profiling by HILIC-MS, site specific glycan mapping and determination of PTM's by peptide mapping. The resulting product profiles were then compared to data generated by the same methods, using cetuximab drug product as a reference compound.

The charge variant profiles of these molecules exhibited more heterogeneous isoforms than anticipated, HEK produced mAb exhibited a pattern similar to the reference product while the CHO produced mAb was less complex. As previously stated, (Chapter 2), the charge variant profile is constituted by isoforms exhibiting differing levels of C-terminal lysine clipping, N-terminal pyroglutamate formation and glycoforms with an abundance of terminal sialylation.

Chapter 4.

4.2. Materials & Methods

Reagents

All molecular biology reagents and mammalian cell culture reagents were as previously described in Chapter 2.

Acetone HPLC grade, acetonitrile LC-MS grade , ammonium hydroxide , CX-1 pH gradients buffers, dimethyl sulfoxide (DMSO) , formic acid , formic acid (0.1% v/v) in water LC-MS grade, glacial acetic acid, trichloroacetic acid and water LC-MS grade were all purchased from ThermoFisher Scientific (Dublin, Ireland).

Ammonium bicarbonate, anthranilic acid (2-AA), DL-Dithiothreitol (DTT), iodoacetamide (IAA), sodium cyanoborohydride, trisodium citrate dihydrate and Urea were purchased from Sigma Aldrich (Wicklow, Ireland)

Promega sequencing grade modified Trypsin was purchased from MyBio (Kilkenny, Ireland). CarboClip (PNGase F) purchased from Asparia Glycomics (Gipuzkoa, Spain). New England Biolabs α1-3,4 Fucosidase (AMF), α1-2,4,6 Fucosidase O (BKF), β 1-4 Galactosidase S (SPG), Prozyme Sialidase A (ABS) and β-N-Acetylhexosaminidase (GUH) purchased from Europa Bioproducts Ltd. (Cambridge, United Kingdom). Fabricator (IdeS) enzyme was purchased from Genovis (Lund, Sweden).

Methods

All molecular biology and mammalian cell culture methods are the same as described, in Chapter 2, with the exception of cetuximab primer sequence and PCR temperature settings. Table 4.2.1 below describes these.

4.2.1.1. Polymerase Chain Reaction (PCR) - cetuximab

Table 4.2.1: Primer sequence, T_m, GC content and annealing temperature for Cetuximab encoding V-genes.

Antibody	Primer For-Rev (5'-3')	Tm (°C)	GC%	Ann. Temp
Heavy Chain V-gene	For-GCACTTGTCACGAATTCGCAGGTGCAGCTGAAG	73.8	54.55	72
	Rev-TGGGCCCTTGGTGCTAGCTGCAGAGACAGTGAC	77.3	60.61	
Light Chain V-gene	For-GCACTTGTCACGAATTCGGATATTCTGATGACGCAAAGCC	72.1	47.5	72
	Rev-TGGTGCAGCCACCGTACGTTTCAGTTCCAGTTTAGTACCTGC	75.5	52.38	

4.2.2. Glycoprotein Deglycosylation

A 200 µg aliquot of each biological replicate and the commercial reference product (n=7) were buffer exchanged by trichloroacetic acid (TCA) precipitation.

An equal volume of 40% w/v TCA was added to each sample to give a final TCA concentration of 20%. The samples were then incubated for 30 minutes on ice and centrifuged in an Eppendorf 5424R microfuge set at 4ºC at 6,100 rpm (8,000g) for 20 minutes. As pellets were difficult to discern from the plastic tubing, the tubes were always orientated in the centrifuge with the hinge of the tube facing outward. Following centrifugation the supernatant was removed and a 300 µL ice-cold acetone wash was added. The samples were then centrifuged for a further five minutes at 4ºC rotating at 6,100 rpm (18,000 × g). The supernatant was removed and the pellets air dried. The pellet was then re-suspended in 380 µL 8 M urea. A volume of 20 µL of 100 mM DTT was added to the sample, to reduce the disulphide bonds present on the IgG, to achieve a final DTT concentration of 10 mM. The reduction reaction was left to incubate at 20°C for 30 minutes shaking at 500 rpm using an Eppendorf Thermomixer FP. The samples were then alkylated, to prevent the reformation of disulphide bonds, by adding 150 µL of 200 mM IAA stock to the tubes and adding urea to make a final volume of 600 µL and a

final IAA concentration of 50 mM. The reaction was then incubated in darkness at 20°C for 30 minutes at a shaking speed of 500 rpm.

The samples were then buffer exchanged by a second TCA precipitation following the procedure as described above. The samples were re-suspended in 50 µL 50 mM ammonium bicarbonate (ABC) and 1 µL of CarboClip PNGase F was added. The digestion complexes were incubated for 16 hours at 37°C.

Sartorius Vivaspin 500 filters were used to separate glycosylamines from the protein solution. The samples were spun for 10 minutes at 12,200 rpm (16,000 × g) at 10°C. A 200 µL volume of ABC was used to wash remaining sample from the reaction tube and the washings added to the filter to ensure good sample recovery. The wash was then spun for a further 10 minutes. The filtrate was collected and transferred to a new Eppendorf tube and dried *via* vacuum centrifugation. The protein remaining on the filter was collected and stored -30°C for peptide analysis.

To enable fluorescent labelling and therefore detection of N-glycans the non-reducing glycosylamines were treated with 1% formic acid for 20 minutes at room temperature. The *N*-acetylglucosamine at the terminal end of the glycan structure has a hemiacetal group which must be hydrolysed to form an aldehyde group to enable fluorescent labelling. The samples were dried down after the incubation and 5 µL of anthranilic acid (2AA) label was added and incubated at 65°C for 5 hours.

The 2-AA labelling solution combines the fluorescent amine compound, anthranilic acid, with glacial acetic acid and sodium cyanoborohydride. The labelling reaction is a reductive amination reaction which forms a stable, fluorescently labelled compound. In the process of a reductive amination reaction, the carbonyl group of the formed aldehyde is converted to an amine through an imine intermediate. This process is induced by the carboxylic acid group of acetic acid and the amine group of 2AA to form a hemiacetal species which is then hydrolysed by the acidic environment created by the glacial acetic

acid. The sodium cyanoborohydride, then, reduces the imine to an amine with the addition of the fluorescent label. Figure 4.2.2 below shows the reaction mechanism of the main species involved in conversion of a glycosylamine to a labelled glycan.

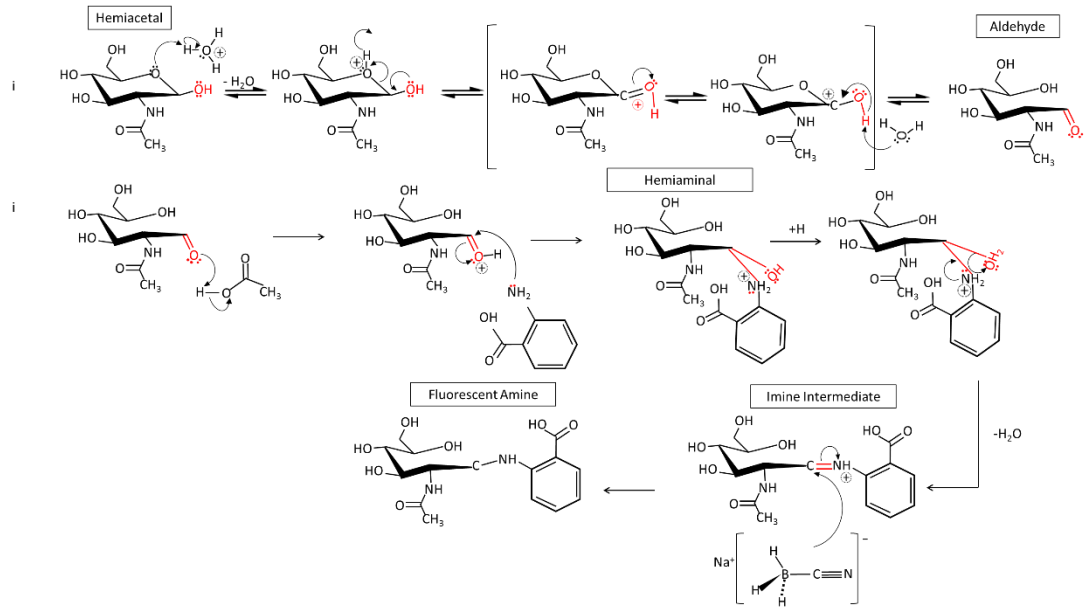


Figure 4.2.1: Conversion of hemiacetal group to aldehyde and further hydrolysis to a fluorescently labelled glycan.

4.2.2.1. Purification by Hydrophilic Interaction Chromatography (HILIC)

Purification of the labelled N-Glycans was performed on a ThermoScientific Ultimate 3000 UHPLC with fluorescent detection. Fraction collection was carried out on a Waters Fraction Collector III. The column used for purification was a ThermoScientific Accucore 150 Amide HILIC 2.1 × 50mm, particle size 2.6 µm. Mobile phases were 50 mM ammonium formate, pH 4.5 (A), LCMS grade acetonitrile (B) and deionised (DI) water, 18.2 MΩ/cm resistivity (C). The gradient conditions were as follows:

Table 4.2.2: Gradient conditions for HILIC purification

Time	Flow (mL.min^{-1})	%B	%C
0.00	0.5	80.0	0.0
2.50	0.5	80.0	0.0
2.51	0.5	20.0	80.0
5.00	0.5	20.0	80.0
5.01	0.5	80.0	0.0
8.00	0.5	80.0	0.0

Fluorescent excitation/emission wavelengths were 350/425 for 2AA. Samples were diluted to a final volume of 20 µL with DI water and 80 µL acetonitrile was added for a final injection volume of 100 µL. Samples eluted at approximately 4.18 minutes whereby fraction collection began and continued for two minutes. The fractionated samples were then dried by vacuum centrifuge and stored at -30°C until further analysis.

4.2.3. Strong Cation Exchange Chromatography (SCX)

A ThermoScientific Vanquish Flex UHPLC system with variable wavelength UV-Vis detector was used with a ThermoScientific MabPac SCX-10 RS column of dimensions 4.6 × 250mm, 5 µm particle size. The ThermoScientific CX-1 buffer system was used to establish a pH gradient based SCX separation. Initial separations were performed to estimate the pH of mAb elution and the retention time of all charge variants. Table 4.2.6 below shows the gradient methods used for each cell line and reference product.

Table 4.2.3: Optimised SCX gradients for cetuximab variants sample set

Sample	Flow Rate ($mL.min^{-1}$)	Time (minutes)	Initial %B	Final %B
HEK	0.5	10	20	70
CHO	0.5	10	20	70
RP	0.8	10	0	80

Following these scouting runs 150 µg of intact mAb were injected in triplicate, for a total of 450 µg of each sample (n=7), and each charge variant fraction was collected. The charge variant fractions were then deglycosylated as per the above procedure and stored at -30°C for further analysis.

4.2.4. Middle down enzyme digestion and Size Exclusion Chromatography (SEC) separation

Size Exclusion Chromatography was carries out on a ThermoScientific Vanquish Flex UHPLC with variable wavelength UV-Vis detector. A MabPac SEC-1 Size Exclusion column of dimensions 7.8 × 300 mm, particle size 5 µm was used for separation with

isocratic conditions. The chromatography buffer consisted of 50 mM NaH_2PO_4 with 300 mM NaCl. The flow rate used was 0.8 $mL.min^{-1}$ at 30°C for 15 minutes. The monomer peak eluted at roughly 10 minutes. A 100 µg quantity of each sample (n=7) were digested with 100 units of enzyme. A 2 µg/ µL concentration was injected (50 µg in a volume of 25 µL) and the subunits separated. After IdeS digestion the Fc and $F(ab')_2$ regions eluted at 11 minutes and 12 minutes respectively each fraction was collected and deglycosylated.

4.2.5. Hydrophilic Interaction Chromatography

HILIC was used to determine the glycosylation profile of the collected charge variant and mAb subunit glycans. HILIC was performed on a ThermoScientific Vanquish Flex UHPLC with a fluorescence detector. An Accucore 150 Amide HILIC LC column was used with dimensions 2.1 × 150mm, particle size 2.6 µm. Mobile phase A consisted of 50 mM ammonium formate, pH 4.4 and mobile phase B was LC-MS grade Acetonitrile. Separation took place at temperature of 50°C, fluorescent excitation/emission wavelengths were 350/425 nm respectively. The gradient conditions were as follows:

Table 4.2.4: HILIC gradient parameters for glycan analysis

Time (minutes)	Flow rate ($mL.min^{-1}$)	%B
-20	0.4	75
0	0.4	75
30	0.4	50
30.1	0.4	45
32	0.4	45

4.2.6. Exoglycosidase digests and Mass Spectral analysis

A panel of exoglycosidase enzymes were used to determine the exact structure of the N-glycans of the cetuximab biosimilars and the reference product. A 400 µg quantity of mAb was digested as per section 4.2.8. The glycan profile was then screened by HILIC chromatography before exoglycosidase digest. The released glycans were resuspended

in 10 µL DI water and aliquoted in to five PCR tubes (n=5 × 7 (samples) = 35). The samples were digested at 37°C for 16 hours. The following day the samples were purified by ethanol precipitation by adding 90 µL ice cold ethanol to the reaction tubes and set at -30°C for 30 minutes. The samples were then centrifuged for 10 minutes at 10°C at a speed 8,800 rpm (16000 × g). The supernatant was removed and the samples left to air dry, then resuspended 70% ACN for a final injection volume of 20 µL. The exoglycosidase panel consisted of; ABS, Sialidase A which has specificity for α(2,3)-, α(2,6)-, α(2,8)-, and α(2,9)-linked N-acetylneuraminic acid (Sialic acid), AMF, α1-3,4 Fucosidase which has specificity for α1-3 and α1-4 fucose but not terminally linked fucose, BKF or α1-2,4,6 Fucosidase O cleaves terminal fucose, SPG also known as β 1-4 Galactosidase S cleaves β linked galactose but not α linked and GUH, β-N-Acetylhexosaminidase (GlcNAc) which cleaves all β linked non reducing N-acetylglucosamine but does not cleave bisecting GlcNAc.

Table 4.2.5: Exoglycosidase enzyme digestion

Enzyme	ABS (µL)	AMF (µL)	BKF (µL)	SPG (µL)	GUH (µL)	H_2O (µL)	Buffer (µL)	Sample (µL)	Total (µL)
ABS	1	-	-	-	-	6	1	2	10
AMF	1	1	-	-	-	5	1	2	10
BKF	1	1	1	-	-	4	1	2	10
SPG	1	1	1	2	-	2	1	2	10
GUH	1	1	1	2	2	-	1	2	10

4.2.6.1. MS Analysis

A ThermoScientific Vanquish Flex UHPLC with FLD was couple to a Q Exactive Plus Hybrid Quadrupole Orbitrap Mass Spectrometer. The HILIC chromatographic conditions were as previously described. Mass spectral analysis was performed on undigested glycans released from HEK and CHO produced Cetuximab and the reference product. To ensure accurate determination of the glycan structures analysis of the exoglycosidase panel were also analysed by MS. All samples were injected in 70% ACN/H_2O.

Table 4.2.6: MS parameters for Glyco profiling

Polarity	Negative
Capillary Voltage (kV)	3.5
Capillary Temperature (°C)	320
Probe Heater Temperature (°C)	400
Sheath Gas (AU)	50
HESI source	
Collision Induced Dissociation (eV)	20
Resolution	70,000
Automatic Gain Control (AGC)	3e6
Maximum Inject Time (ms)	50
Mass Range (m/z)	400 – 3000

Data processing of the separated glycans was performed using XCalibur software and GlycoWorkbench 2.

4.2.7. Glycan Analysis of the reference product

De-glycosylation and labelling were performed as above. LC-MS analysis was performed on a Waters Xevo G2 QToF coupled to an Aquity UPLC using a Waters BEH Amide column of dimensions 1.0 x 150mm, 1.7 µm particle size. Solvents and gradients were the same as above. MS setting were as follows:

The outlet of the chromatographic system was coupled directly to a Waters Xevo G2 QToF mass spectrometer (Milford, MA, USA) equipped with an electrospray ionization interface. The instrument was operated in negative ion mode with a capillary voltage of 1.80 kV. The ion source and nitrogen desolvation gas temperatures were set at 120 °C and 400 °C, respectively. The desolvation gas flow rate was 600 L/h. The cone voltage was maintained at 50 V. Full-scan MS data was acquired over the range of 450 to 2500 m/z [11].

Chapter 4.

4.2.8. Peptide Mapping

Peptide mapping analysis was performed on a ThermoScientific Vanquish Flex UHPLC couple to a Q Exactive Plus Mass Spectrometer. Chromatographic separation was performed on an Acclaim C18 column of dimensions 2.1 × 250mm, particle size 2.2 µm. Protein collected from the charge variant fractionation deglycosylation and the intact protein de-glycosylation were trypsin digested using Prozyme sequencing grade Trypsin. The enzyme to protein ratio was 1:50 with a final digestion volume of 100 µL. The protein was digested overnight at 37°C. The following day the enzyme was denatured by adding 5 µL formic acid. The peptides were then completely dried *via* vacuum centrifugation and subsequently resuspended in 20 µL LC-MS grade 0.1% formic acid in water. Gradient conditions were:

Table 4.2.7: Peptide mapping gradient conditions

Time (minutes)	**Flow rate ($mL.min^{-1}$)**	**%B**
0	0.3	2
45	0.3	40
46	0.3	80
50	0.3	80
50.5	0.3	2
65	0.3	2

Table 4.2.8: Full MS, dd-MS2 & dd conditions

Full MS	
Polarity	Positive
Capillary Voltage (kV)	3.8
Capillary Temperature (°C)	320
Probe Heater Temperature (°C)	400
Sheath Gas (AU)	40
HESI source	
Collision Induced Dissociation (eV)	0.0
Resolution	70,000
Automatic Gain Control (AGC)	3e6
Maximum Inject Time (ms)	100
Mass Range (m/z)	200 – 2000
dd-MS^2	
Resolution	17,500
AGC	1e5
Max IT	200
Loop	5
MSX	1
TopN	5
Iso window (m/z)	2.0
(N)ce	28
dd	
Min AGC	2e3

Chapter 4.

Peptide mapping data was processed using ThermoScientific Biopharma Finder 3.0 and Protein Metrics Byonic & Byologic. The following processing parameters were used: precursor mass accuracy - 5 ppm, maximum number of modifications per peptide – 1, protease – trypsin. The variable modifications searched for were; N-terminal pyroglutamate, C-terminal lysine clipping, side chain carbamidomethylation and carboxymethylation, asparagine deamidation, glutamine deamidation and oxidation of methionine and tryptophan. The following amino acid sequences were used to process the HEK and CHO produced biosimilars. The reference product sequence was obtained from Ayoub *et al.*(2013)[12].

Heavy Chain:

QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWSGGN
TDYNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTL
VTVSAASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHT
FPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCP
PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV
HNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ
PREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD
SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

Light Chain:

DILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPS
RFSGSGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKRTVAAPSVFIF
PPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYS
LSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

Chapter 4.

4.3. Results & Discussion

4.3.1. Strong Cation Exchange Chromatography

Following, a single step, protein A purification, the transiently expressed in-house cetuximab variants were analysed by SCX chromatography using the scouting methods developed in Chapter 2. The charge variant profiles, of each HEK and CHO produced cetuximab, were expected to be less populated than the reference drug product. As they were produced by cell lines which are known to produce human glycosylation patterns, the contribution to charge of these structures was expected to be less[13,14]. Figure 4.3.1 shows the SCX chromatograms obtained. The charge variant profile of the reference drug product (RDP) consists of three acidic variants, one main and three basic. The main variant is clearly distinguishable in 4.3.1 (C). The HEK variants exhibit the same profile with the same pI, with seven acidic variants and four basic. The abundance of basic variant 1 & 2 (B1 & B2) are almost equivalent to the main isoform. Comparison to the RDP shows that the HEK variants have a more acidic pI as they elute earlier and have a higher quantity of acidic species. However, the distribution of the HEK CV is similar to the RDP. The CHO variants 4.3.1 (B)) have a considerably different distribution. The CHO variants consist of four acidic peaks and four basic peaks, less than HEK but with varying abundance compared to the RDP. CHO variant two, red line 4.3.1 (B), exhibits a pI which is fractionally more acidic than the other variants.

The complexity unveiled by charge variant analysis was surprising. To investigate this observation further a number of additional experiments were performed. These included the use of size exclusion chromatography to determine whether aggregation played any role in the observed complexity and in depth analysis of posttranslational modifications, both N-glycosylation and PTMs present in the primary sequence, such as deamidation and C-terminal lysine clipping to elucidate their contribution to the overall charge variant pattern.

Chapter 4.

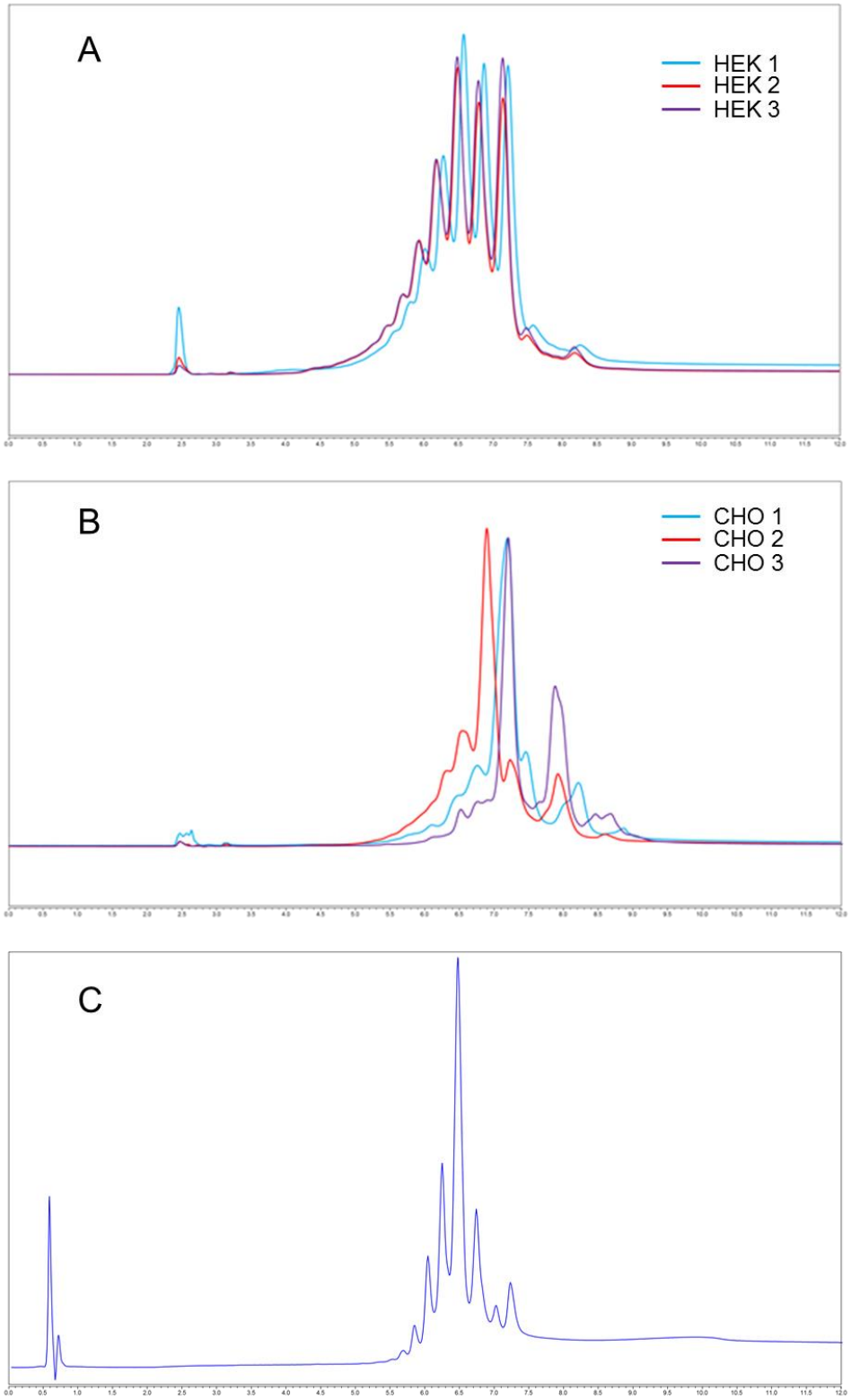


Figure 4.3.1: SCX chromatograms for; (A) HEK produced variant, (B) CHO produced variant & (C) reference drug product. The in-house variants have complex CV profiles which exhibit biological heterogeneity.

4.3.2. Size Exclusion Chromatography

Size exclusion chromatography indicated that were no dimers in the cetuximab variants. Figure 4.3.2 shows the chromatograms obtained. There appears to be low molecular weight fragments occurring in the in-house variants, after the monomer peaks in (A) and (B). The CHO variants appear to have some higher molecular weight fragments, peaks at ~4.5 mins. The RDP appears to have minor low molecular weight fragments also, but overall the SEC is as expected for a highly polished product. Fragmentation of the mAb, by both enzymatic and non-enzymatic processes, was shown to contribute to the CV profile in Chapter 3. However, conclusive identification of the contributing PTM's, required further analysis. The glycan profile of each variant was subsequently determined.

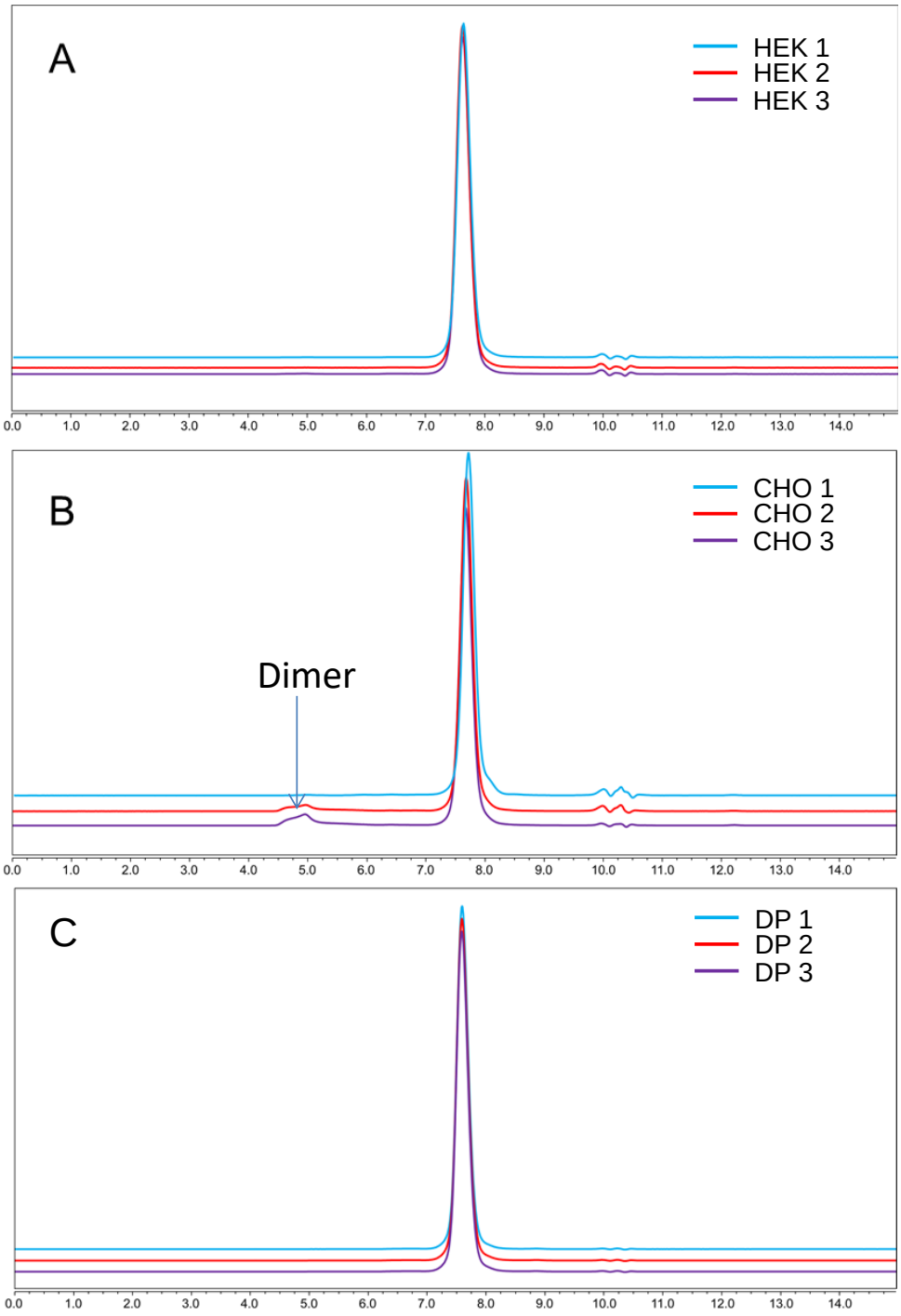


Figure 4.3.2: Size exclusion chromatograms for; (A) HEK produced variant, (B) CHO produced variant & (C) reference drug product

Chapter 4.

4.3.3. Glycan Profiling

The glycan profile (GP) of cetuximab is complex and populated by many non-human glycoforms. Allergic reactions to the product were reported upon mass market release[8]. In this section, the glycan profiles in-house variants of cetuximab are compared to that of the drug product. The GP of the variants provide an interesting insight into the effects cell line and culture conditions can have on the glycoform.

4.3.3.1. Expi293 glyco profiling

The glycan profile of the HEK produced cetuximab variant was expected to contain predominantly complex type N-glycans[15]. As HEK is a human line, there were no α-Gal detected, sialylated forms were present but only of NANA, the human form of sialic acid. Figure 4.3.3 shows the HILIC chromatogram obtained from the MS analysis of the released glycans.

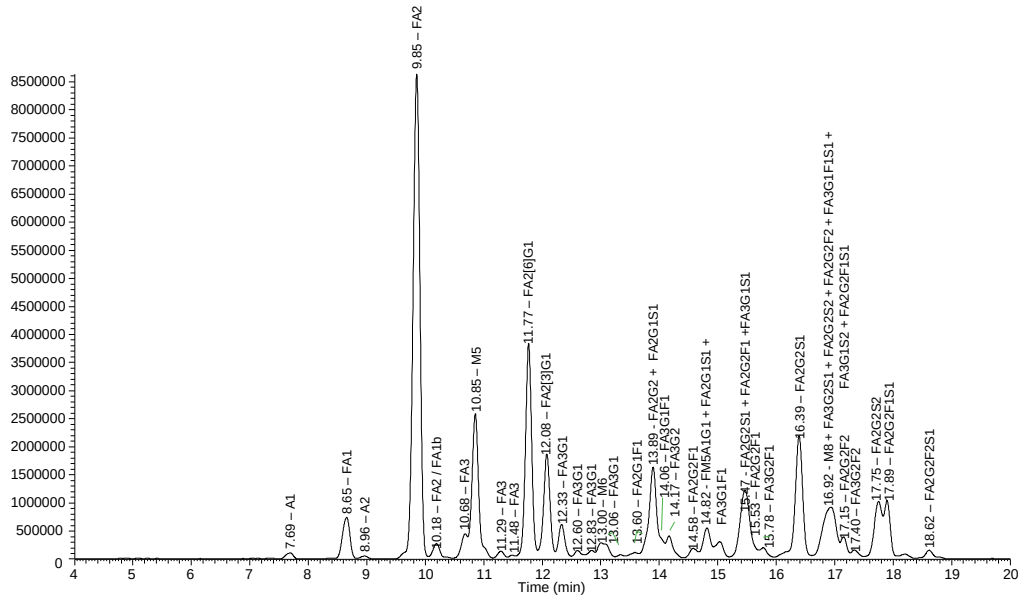


Figure 4.3.3: HILIC-fluorescence chromatogram of HEK variant.

Table A in the appendix summarises the glycan structures found. Forty-seven glycan structures were identified for the HEK variant. The most abundant N-glycans of the HEK variant are FA2, FA2[6]G1 and FA2G2S1. There are numerous sialylated glycoforms, which could account for the highly acidic CV profile. Interestingly, there are a number of bisected forms which are not present on the RDP. Figure 4.3.3.2 shows, taking

Chapter 4.

FA2G2S1F1 as an example, how exoglycosidase panels were used to determine the identity of glycoforms. Taking into account retention time and enzyme specificity the site and linkage of glycan residues can be determined. This is particularly useful in the absence of good quality MS/MS data, where only MS^1 values of m/z are available, as many glycan residues are isobaric.

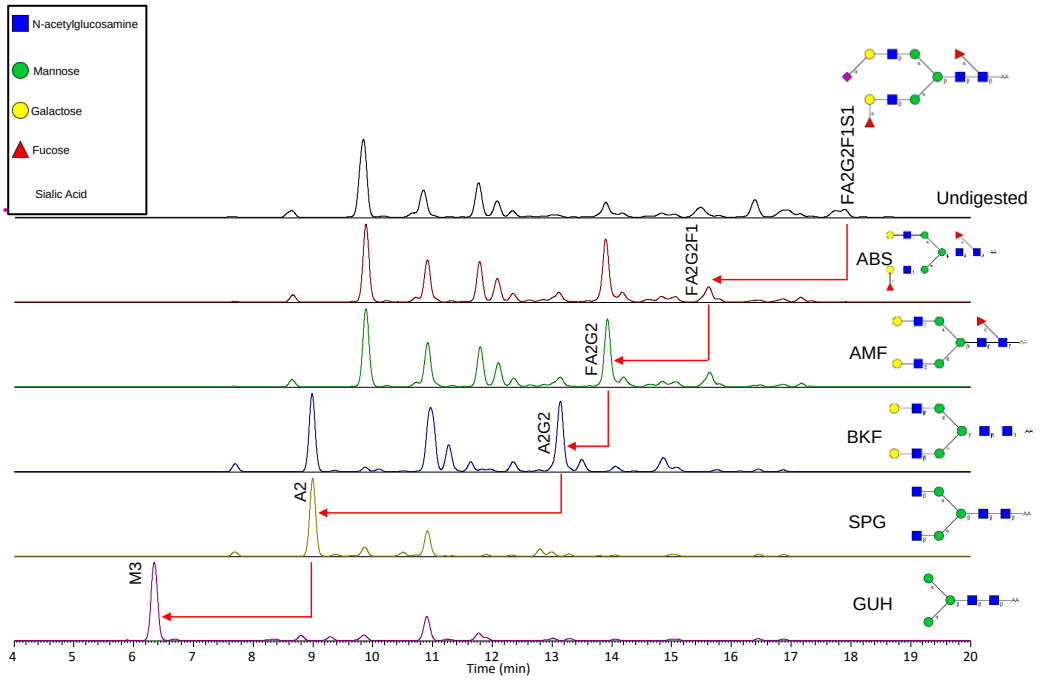


Figure 4.3.4: Exoglycosidase panel showing the deduction of residue identity and linkage by reverse addition**.**

Figure 4.3.5 shows the corresponding MS signals where there is an evident decrease in MS peaks as the number of exoglycosidases increases. Taking FA2G2A1F1 as an example again, Figure 4.3.5 shows the mass shift associated with each enzyme cleavage

Chapter 4.

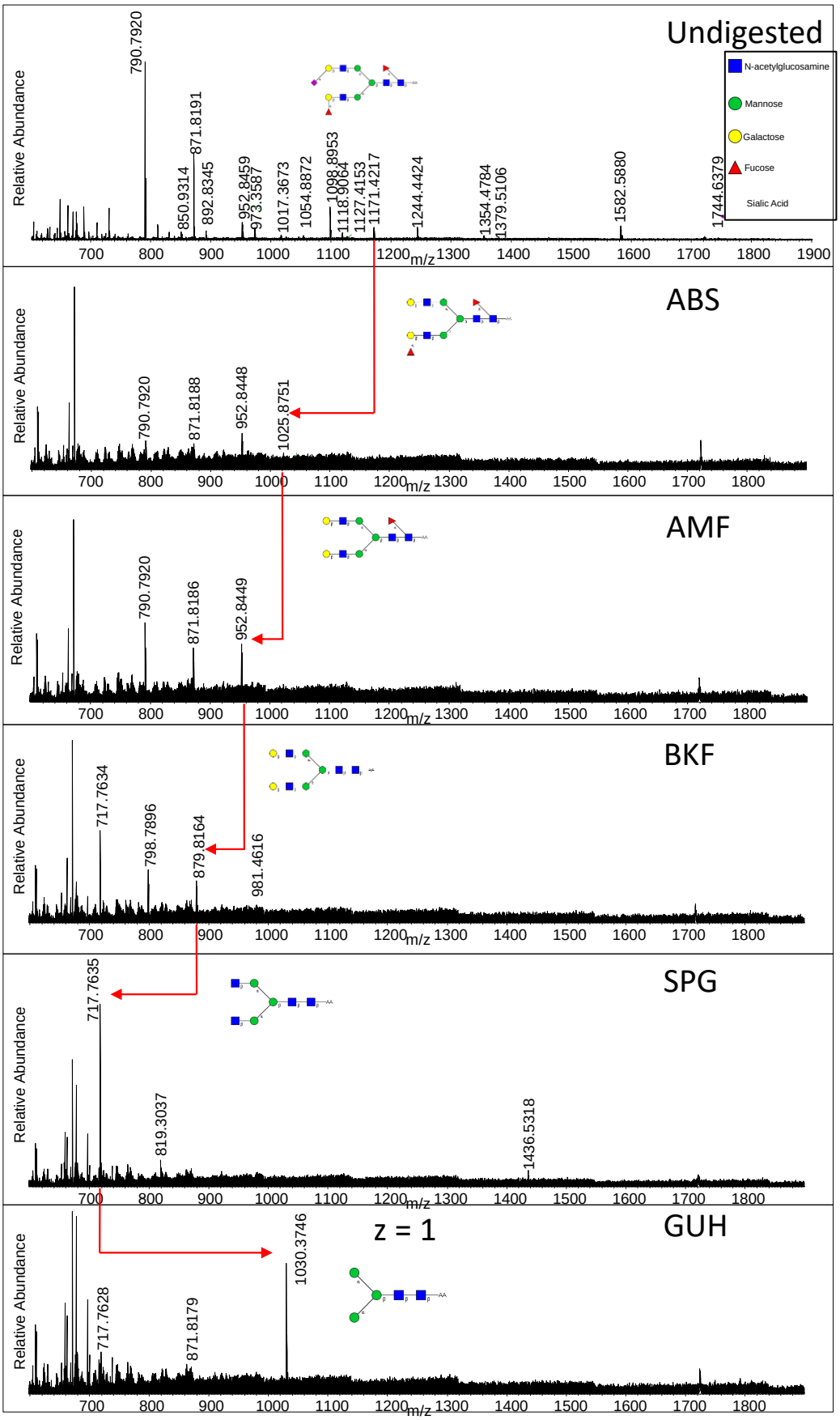


Figure 4.3.5: Averaged MS Spectra of exoglycosidase digests for HEK variant

Chapter 4.

The mass matching error (ppm) for the HILIC-MS analysis of the HEK variant, intact glyco profile were excellent in comparison with the RDP MS runs, Table 1 (Appendix). The RDP intact glycoform analysis was performed on a Waters Xevo instrument and the considerable improvement in mass accuracy on the QE is evident. To ensure consistency across all glyco profile annotation, the glycan samples from the RDP site specific digestions were performed in the Q-Exactive instrument using the same parameters as the in-house biosimilars, and gave the same mass matching error.

4.3.3.2. ExpiCHO glyco profiling

The CHO glycan profile was expected to be less complex than HEK or RDP with mostly high mannose forms detected[16]. Similarly to HEK contrasting to RDP, there was no α-Gal detected or NGNA. Figure 4.3.6, below, shows the HILIC chromatogram for CHO variant cetuximab. There were twenty-eight glycoforms identified, of which FA2 and M5 were the most abundant. Interestingly the FA2[3]G1 isoform is more abundant compared to the FA2[6]G1 in CHO than in HEK. Some sialylated N-glycans were observed but were not considered to be significantly contributing to the acidic nature of CV profile.

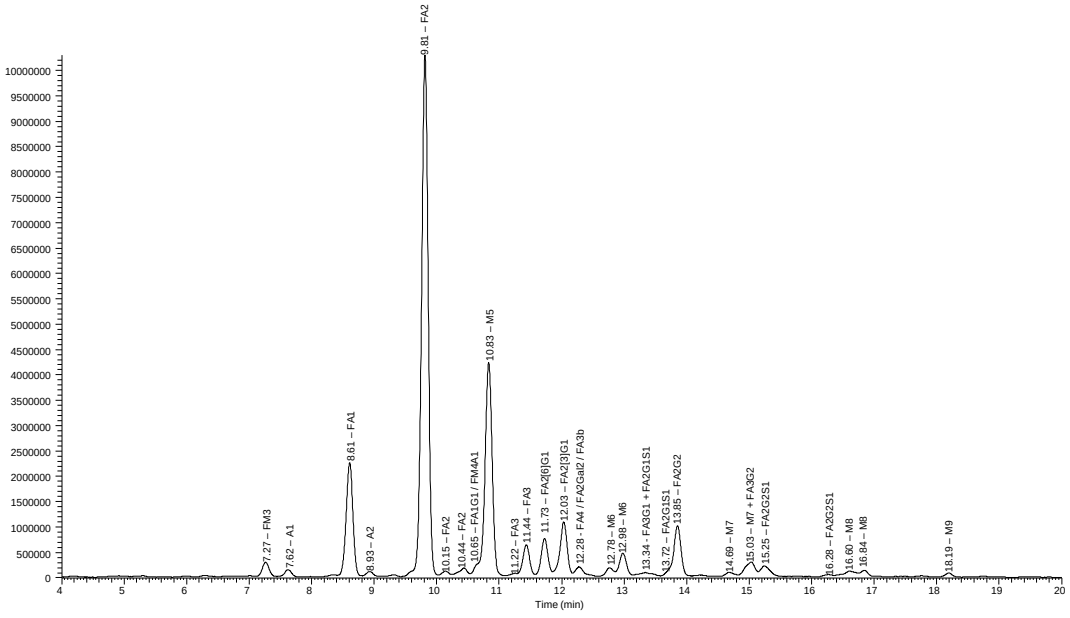


Figure 4.3.6: HILIC-fluorescence chromatogram of CHO variant.

Figure 4.3.7 illustrates how exoglycosidase panels can be used to rule out glycoforms and aid identification. Taking FA3b and FA4 as two potential glycan candidates, the lack of an A1 peak and the abundance of only high mannose in GUH chromatogram indicates that there are no bisected forms in the profile and that the assignment of FA4 is correct.

Chapter 4.

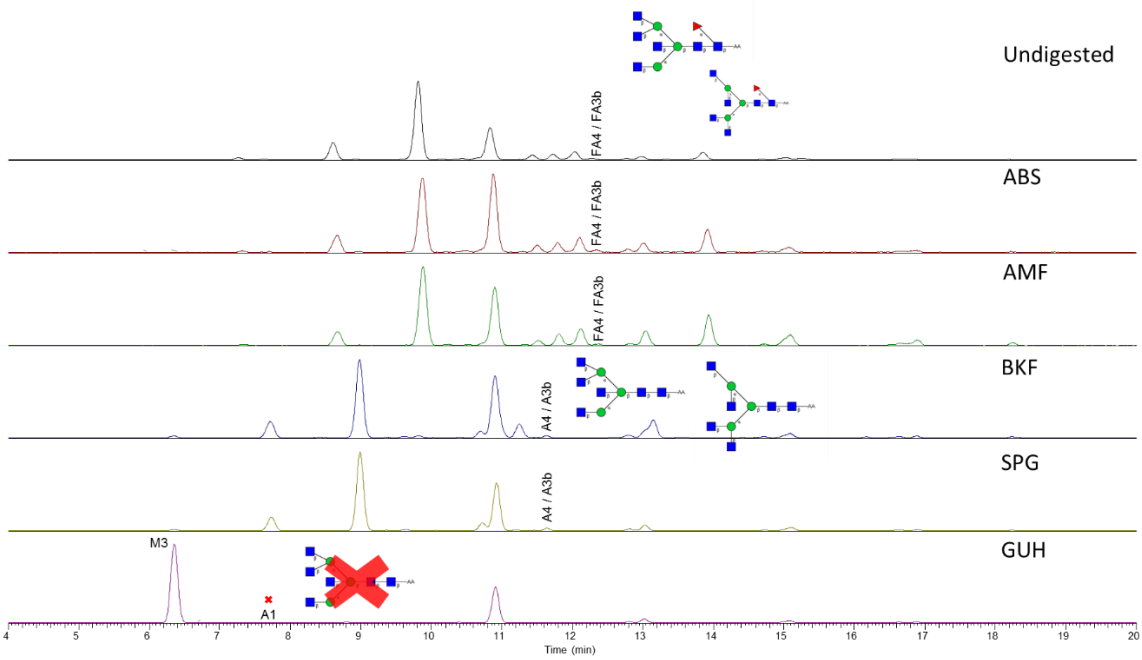


Figure 4.3.7: Exoglycosidase panel of the CHO variant illustrating structural elucidation of a potentially bisected glycan**.**

The corresponding MS traces are shown in Figure 4.3.8, below. The simplicity of the CHO profile relative to HEK is apparent. FA2 is given as an example of the effect enzyme digestion has on the complexity of the averaged MS spectra. The full annotated glycan structures are provided in Table B, Appendix 1.

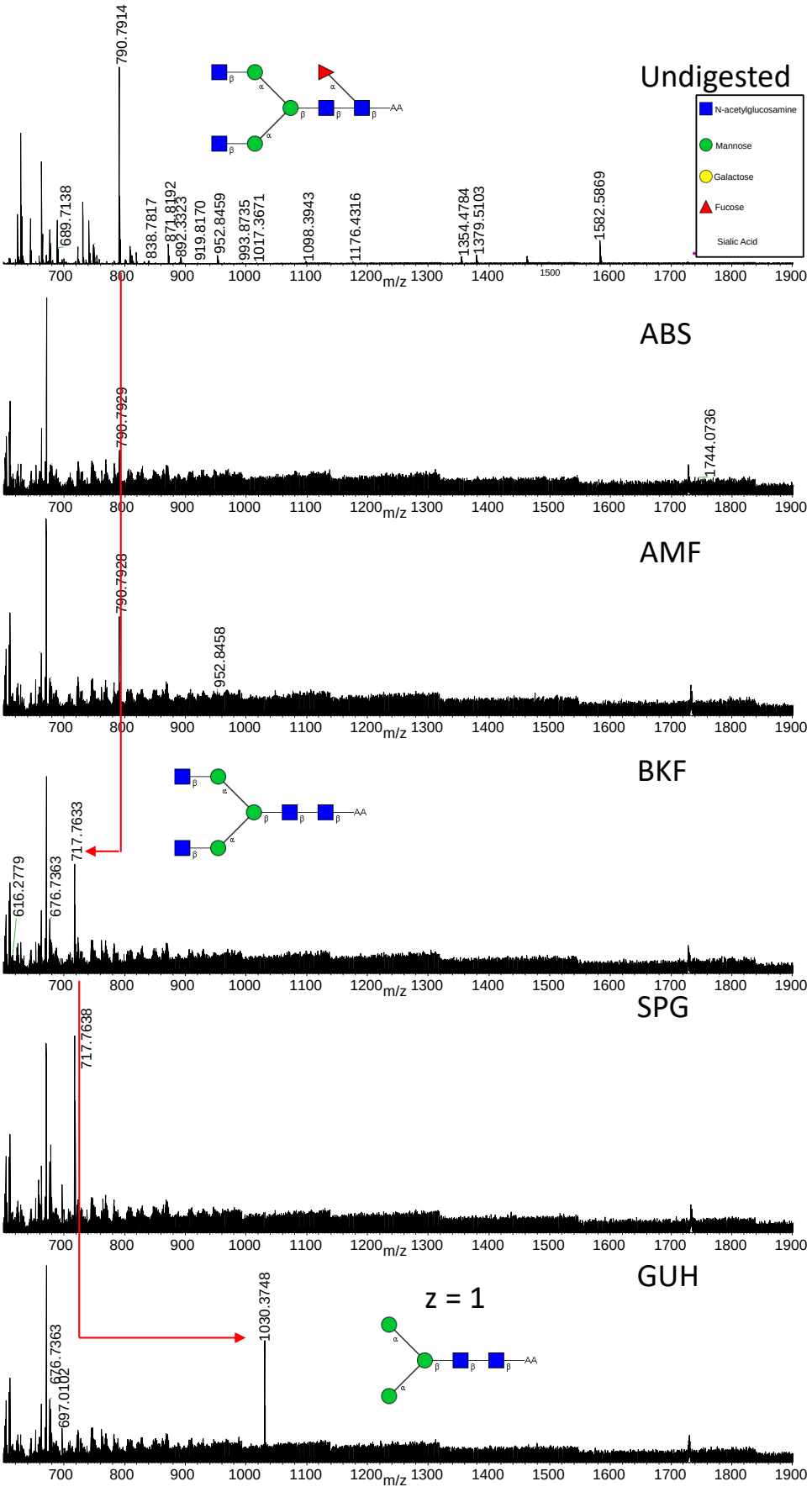


Figure 4.3.8: Averaged MS spectra of exoglycosidase panel for the CHO variant

4.3.3.3. Reference Drug Product Glycan Profile

Glycan profiling of the reference drug product was performed on a Waters Aquity UPLC couple to a Xevo G2 QToF. The complexity of cetuximab's glycan profile is well studied [12,17,8,6,7]. The adverse effects seen patients upon administration highlighted the importance of in-depth glyco-profiling of mAb therapies. The α-Gal and NGNA residues are not human and so result in IgE antibodies being produced against them.

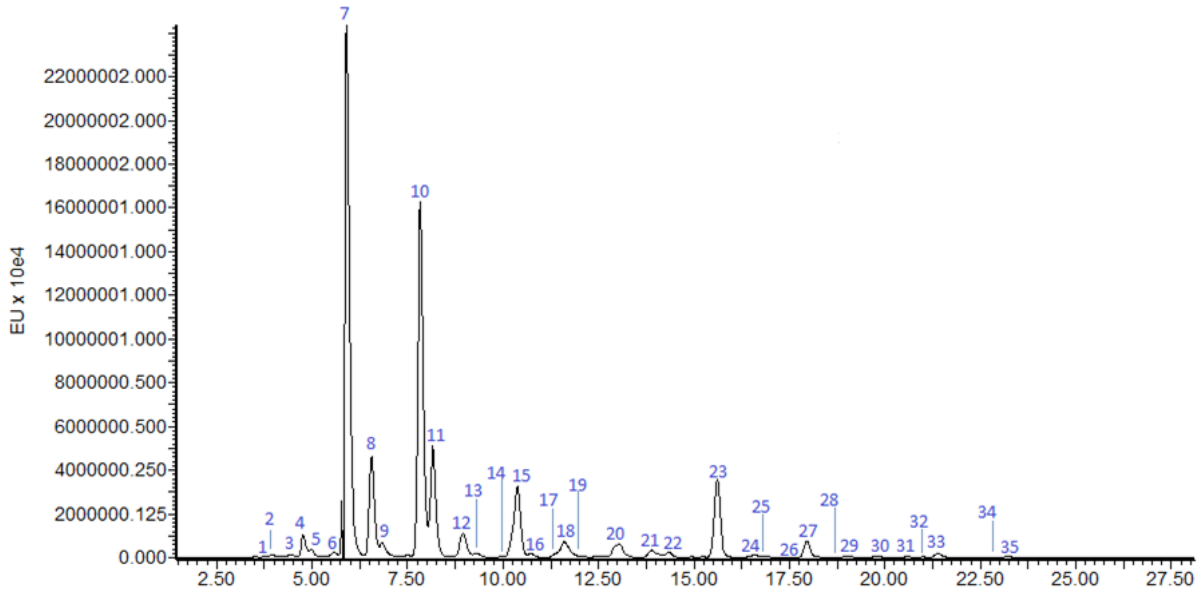


Figure 4.3.9: HILIC-fluorescence chromatogram of reference drug product performed on an Aquity UPLC.

The problematic cetuximab glycan structures widely reported in the literature were all identified in this study. The RDP profile is dominated by complex and hybrid type structures. Forty-eight glycoforms were detected. Table C, appendix, summarises the glyco profile of the RDP. Mass error for these identities is considerably higher than those observed on samples run on the Q Exactive system. For confident assignment the glycan structures annotated from data derived from the Xevo system were cross referenced with the data derived by the Q-Exactive for the RDP site localisation study.

Figure 4.3.10 illustrates the differences and similarities across the cetuximab variants glycan profiles. HEK ∩ CHO ∩ RDP shows all glycans common amongst all samples, this set consists of complex and high mannose forms, which were some of the most

abundant found including FA2, M5 and FA2G1. HEK ∩ RDP is dominantly hybrid type and complex, characteristic of the enzymatic phenotype of the cell lines used. HEK ∩ CHO are predominantly high mannose, illustrative of the relative simplicity of the CHO cells glyco enzyme production as well as the HEK cells. Only one glycoform inhabits the CHO ∩ RDP set, FM3.

The glycoforms unique to the HEK variant are mostly terminally sialylated and/or fucosylated. *In vivo,* core fucosylation has been shown to reduce binding to FcγRIIIa and negatively impacting ADCC activity[18]. Higher levels of terminal sialic acids have been correlated to longer serum half-life of biotherapeutics[19]. The glycan profile for HEK is consistent with hybrid/complex type profile of the RDP. *In vivo,* the HEK variant, in terms of glycan mediated behaviour could potentially exhibit a similar impact as the RDP, without the anaphylactic reaction as it does not have non-human glycoforms.

Overall, the CHO variant has the simplest profile of the sample set. The CHO cell line is a staple in biopharmaceutical manufacturing as it produces large quantities of protein with human-like features.

Once, the glycan profiles of the variants had been determined, site specific annotation was undertaken. Heterogeneity of the glycan profile across the two glycosylation sites of cetuximab were considered to be affecting the charge variant patter of the mAb, particularly in the case where fragmentation was occurring.

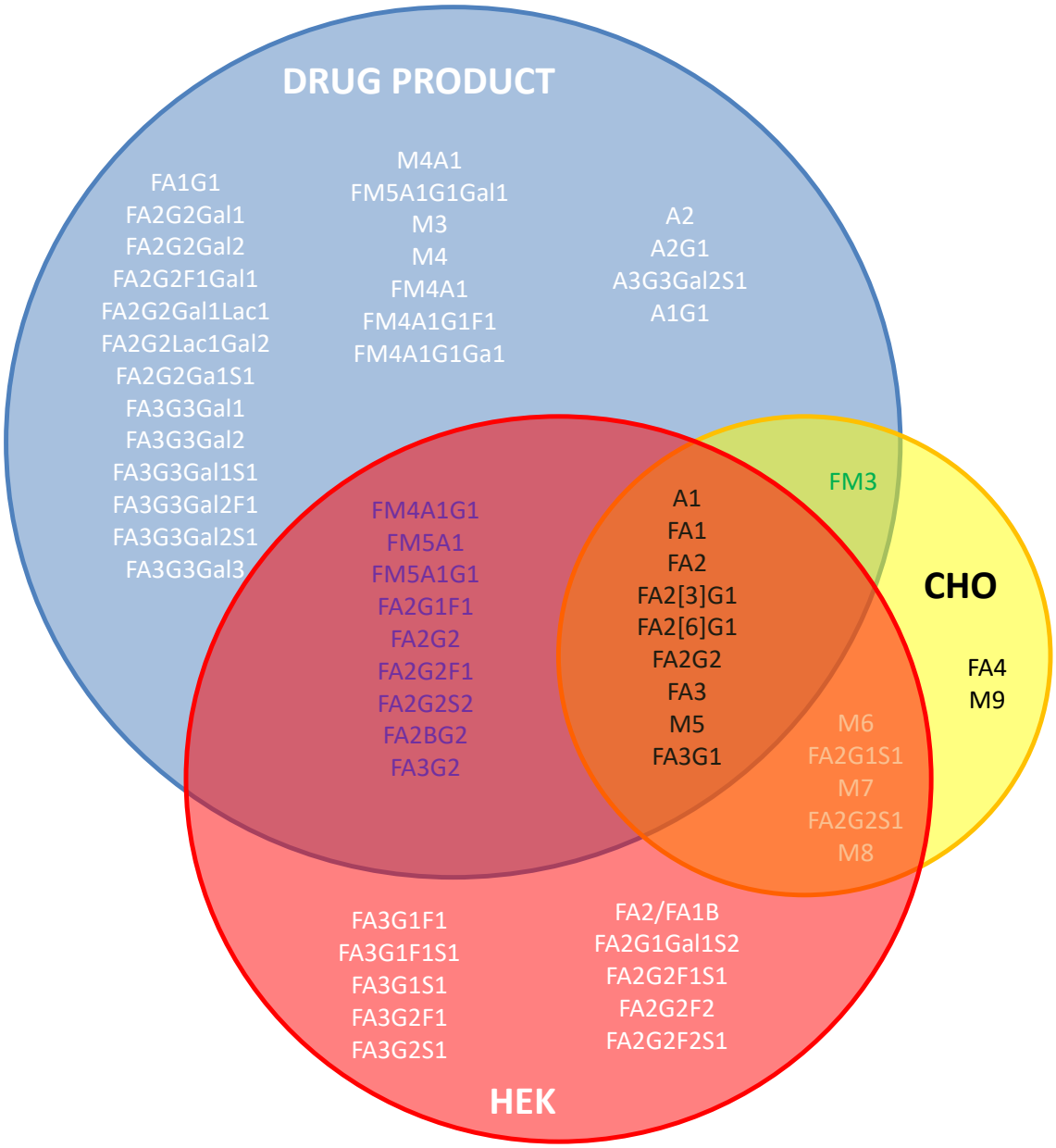


Figure 4.3.10: Venn diagram of the glycan profiles for the HEK & CHO variants and the reference drug product.

Chapter 4.

4.3.4. Site-Specific glycan profiling

The two glycosylation site of cetuximab are known to be located on Asn88 of the light chain (LC) and Asn299 of the heavy chain (HC). Asn299 on the HC is located within the structure (Figure 4.3.14). This results in less complex glycan structures on the HC as the enzymes required for mannose cleavage/complex addition cannot physically reach the residue. For site specific annotation the HEK, CHO and RDP mAbs were digested with IdeS enzyme and subsequently separated by SEC and de-glycosylated. Using retention time and peak shape to inform annotation the resulting HILIC chromatograms provide an interesting insight into the machinery at work in glycosylation. The HILIC trace for the Fab and Fc domain glycans shows the complexity and diversity of N-glycans on the structure, and the effect 3D orientation of the protein plays.

4.3.4.1. IdeS digestions

The HEK and CHO variants sub-unit separation shows similarly distributed glycans. Figure 4.3.11 shows the HILIC chromatogram obtained for the HEK variant. The Fab fragment profile is dominated by highly abundant complex sialylated forms. However, there is also a high abundance of mannose type glycans with the peak for M5 being most prominent in the chromatogram. The Fc region consists of mostly bi-antennary glycans with varying degrees of galactosylation. FA2[6]G1 is disproportionally more abundant than the FA2[3]G1 isoform. This could infer that the orientation of FA2[6]G1 is more favourable in the confined space of Asn299. Fc glycosylation has been shown to have an effect of conformational stability [18]. The shorter chain glycans found on the Fc contribute to this by preventing the molecule being pushed apart by highly branched structures. Figure 4.3.4.2 illustrates the physical space occupied by the FA2 glycan structure on a human IgG Fc region.

Chapter 4.

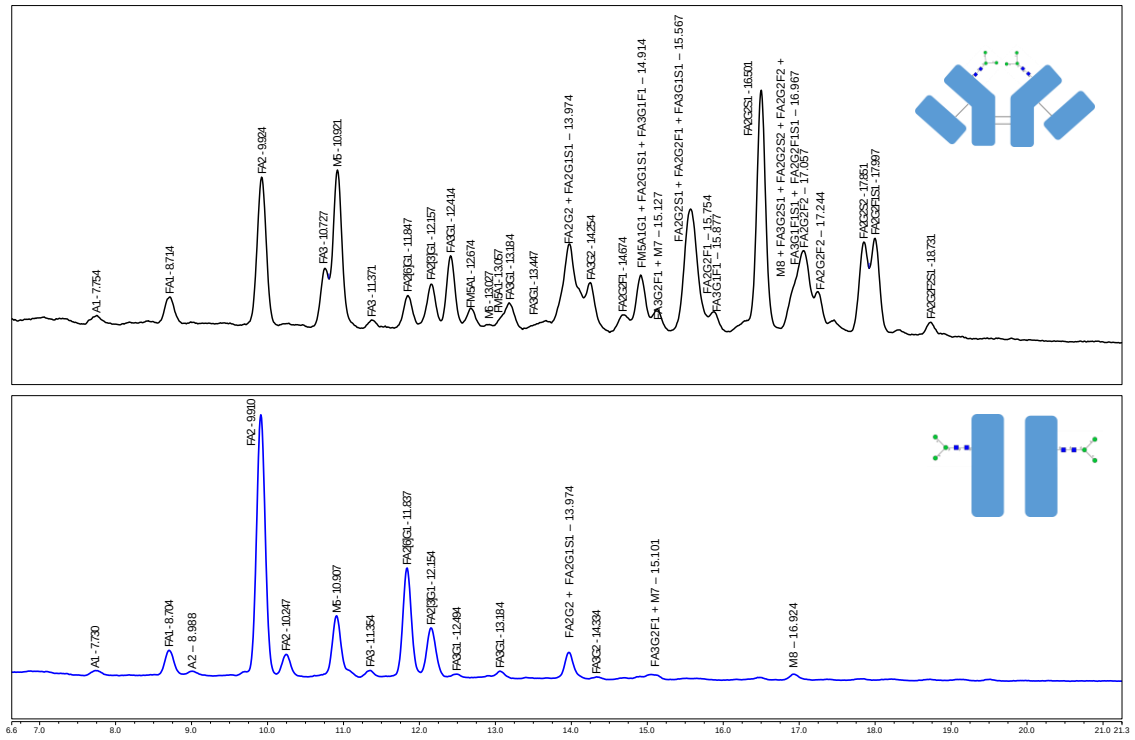


Figure 4.3.11: HILIC chromatogram for HEK variant. Top: Fab glycans, bottom: Fc glycans.

Figure 4.3.12 shows the HILIC chromatogram for the CHO variant. The glycan profile of CHO variants Fc region is predominantly high mannose forms. This is further evidence of effect that the physical orientation of the Asn site has on the complexity of glycan residues. More complex types are found on the Fab as seen previously.

Chapter 4.

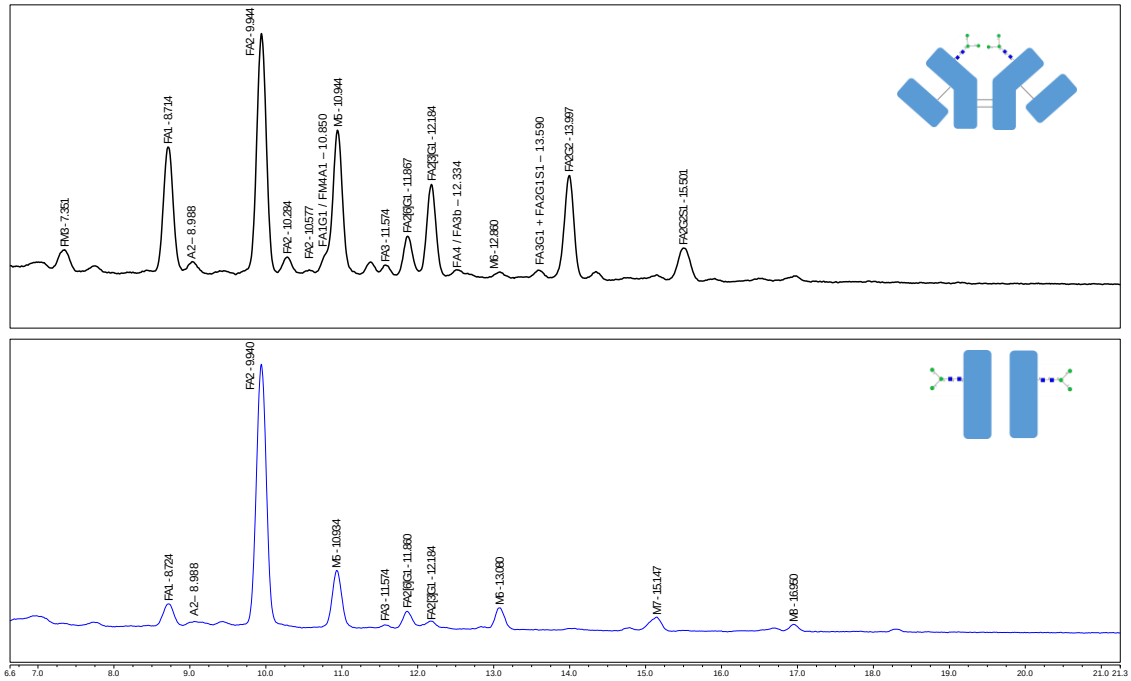


Figure 4.3.12: HILIC chromatogram for CHO variant. Top: Fab glycans, bottom: Fc glycans.

The most dramatic contrast in domain specific glycosylation is seen in the RDP, Figure 4.3.13, has a defined 50:50 split in abundance. The Fab features more hydrophilic larger, later eluting glycans compared to the Fc which is predominantly early eluting peaks that correspond to smaller bi-antennary structures. The majority of glycoforms on the Fab are highly sialylated, α-Gal containing structures. Whereas, the Fc is predominantly simpler bi-antennary type. What is also striking is the disparity between the in-house variants and the RDP. The in-house variants had a more even distribution of FA2 based glycans, with high abundances on each domain; The RDP on the other hand has very few FA2 type glycans on the Fab. Du *et al.* showed that though high mannose and core fucose did not contribute to charge, the effects these structures have on conformation could provide a small contribution to acidic variants[20]. This could be one part of a complicated amalgamation of reasons why the in-house produced variants exhibit significantly different charge profile to the RDP.

Chapter 4.

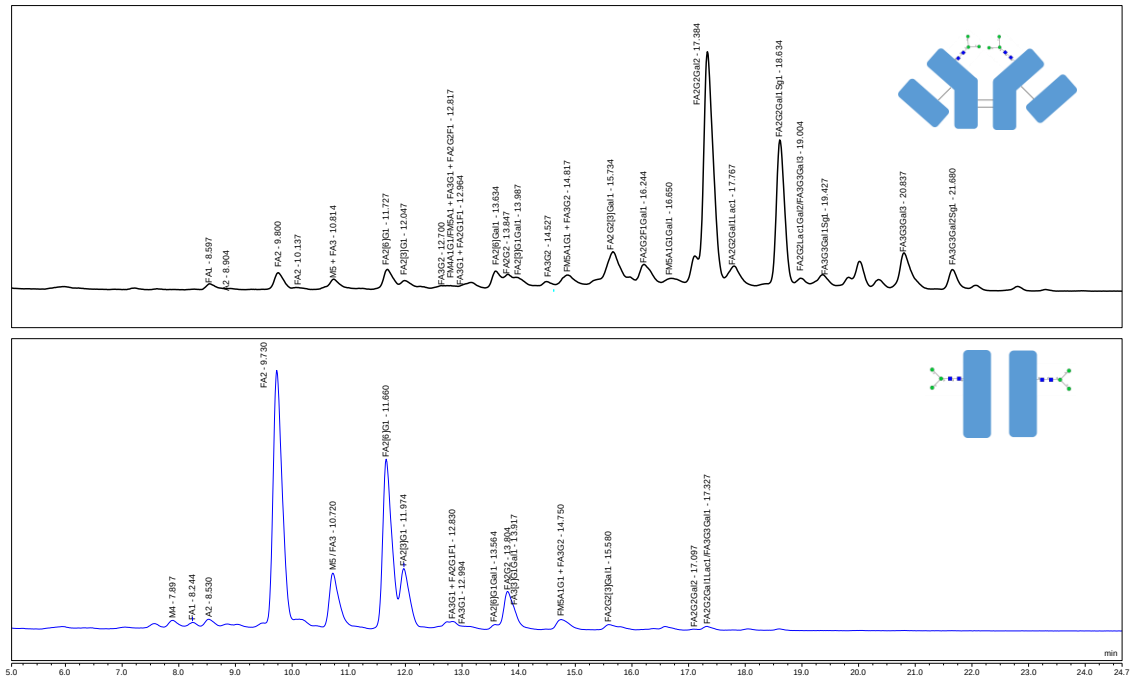


Figure 4.3.13: HILIC chromatogram for RDP. Top: Fab glycans, bottom: Fc glycans.

4.3.4.2. Charge Variant Glycans

The glycan profiles determined in the previous section provided insight into the potential contribution of those forms have on the charge variant profiles of the in-house variants. Site specific localisation determined that the majority of charged glycans were found on the Fab region, with the Fc region containing mostly high mannose, less complex glycoforms. The charge variant fractions of each mAb were subsequently collected, de-glycosylated and their specific glycoforms analysed by HILIC.

Figure 4.3.16 shows the SCX chromatogram obtained during fraction collection peak A7 has a high UV absorbance, but less than the first basic peaks. This infers that though there was less protein contained within this peak the number of glycan residues contained was considerably higher. This supports the observations of Du *et al.* that glycoforms can contribute to charge heterogeneity not necessarily by their individual charge but as an effector of conformational charge variation[20].

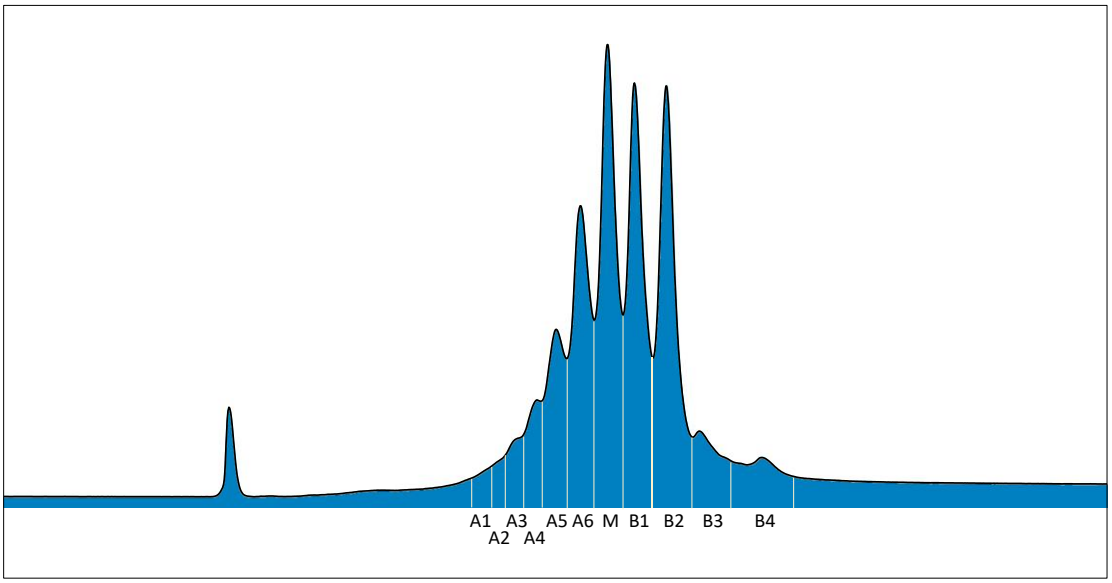


Figure 4.3.14: SCX chromatogram showing fractionation regions of HEK variant cetuximab

The majority of glycoforms for the HEK variant were detected across all charge variants but with considerably different abundances. Figure 4.3.17 shows the HILIC chromatogram. For instance, in the earlier acidic region peaks the abundance of sialylated forms is greater than the basic region peaks. Overall glycan distribution appears to favour the acidic region peaks. Peak A1 shows equivalently high absorbance of the earlier FA2 series of glycans as the main variant itself. One interesting feature is the high quantity of M% variant in this peak, larger than any other charge variant. Also worth noting is the abundance of glycoforms in the A7 peak.

Chapter 4.

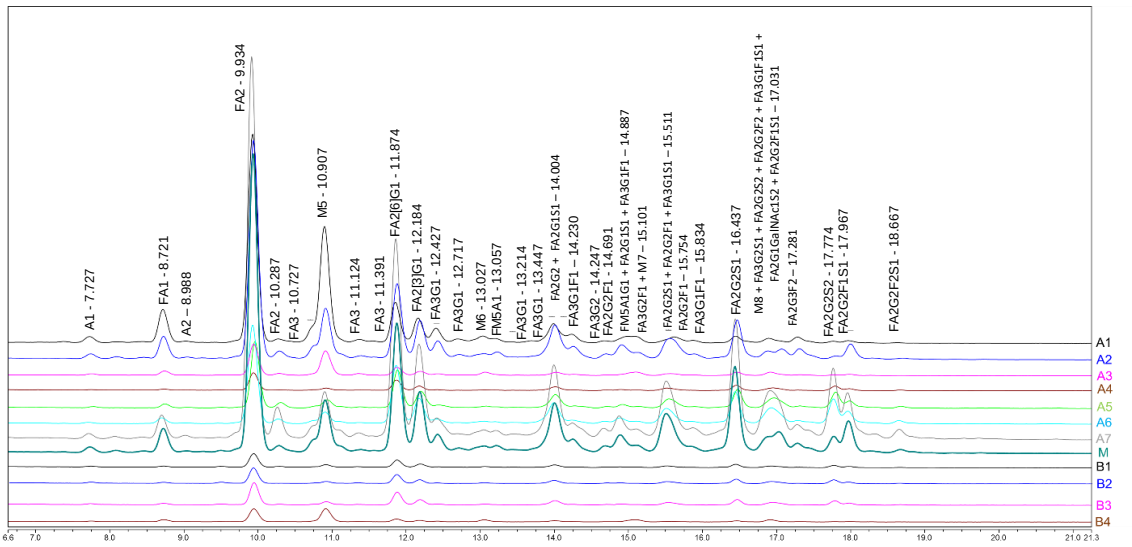


Figure 4.3.15: The distribution and relative abundance of glycoforms across twelve charge variants of the HEK cetuximab variant.

The SCX chromatogram of CHO variant, Figure 4.3.18, below. Abundance of charge variants considerably different to HEK variants has an observable impact on glycoform distribution.

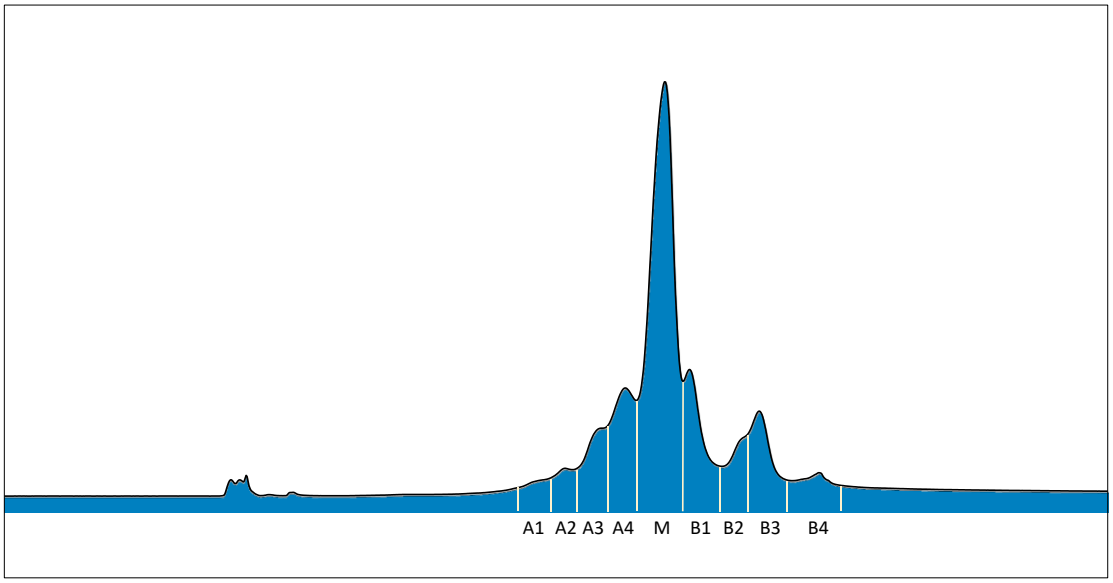


Figure 4.3.16: SCX chromatogram showing fractionation regions of CHO variant cetuximab

Figure 4.3.19 shows the distribution of the N-Glycans across the charge variant peaks. The glycoforms present on the CHO cetuximab variant are largely limited to the main charge variant. As this is the most abundant form, this was expected. However, this raises the question of what is contributing to the acidic nature of the CHO variant. Section

Chapter 4.

4.3.5 discusses the other PTMs which appear to contribute to the CHO variants acidic profile relative to the RDP.

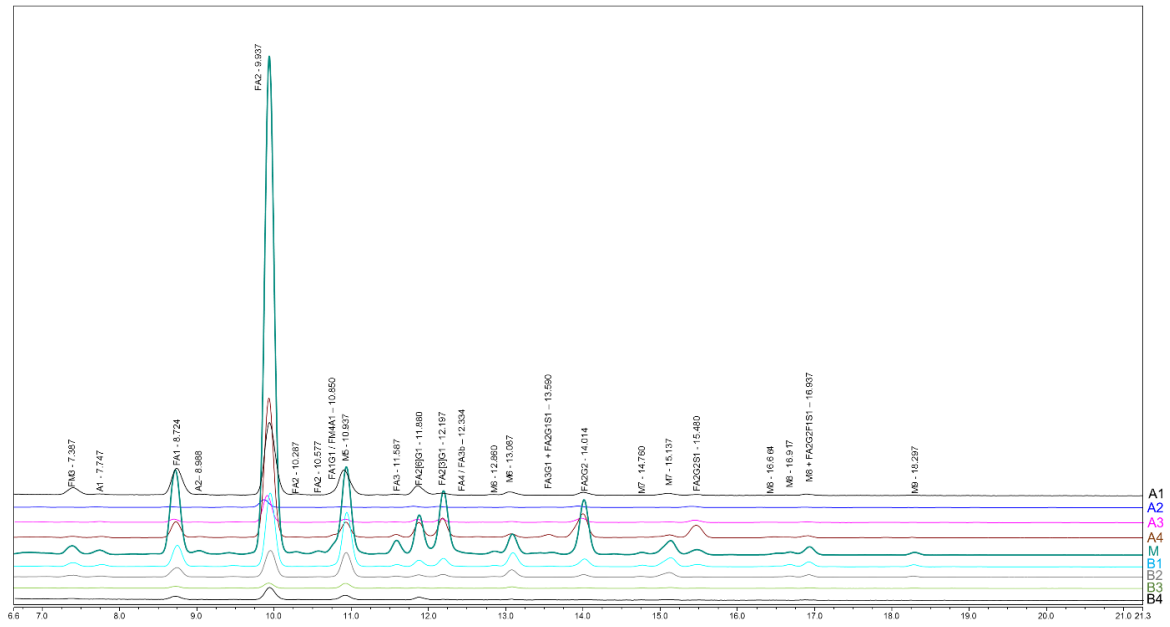


Figure 4.3.17: The distribution and relative abundance of glycoforms across twelve charge variants of the CHO cetuximab variant.

Figure 4.3.20 shows the SCX chromatogram of the RDP.

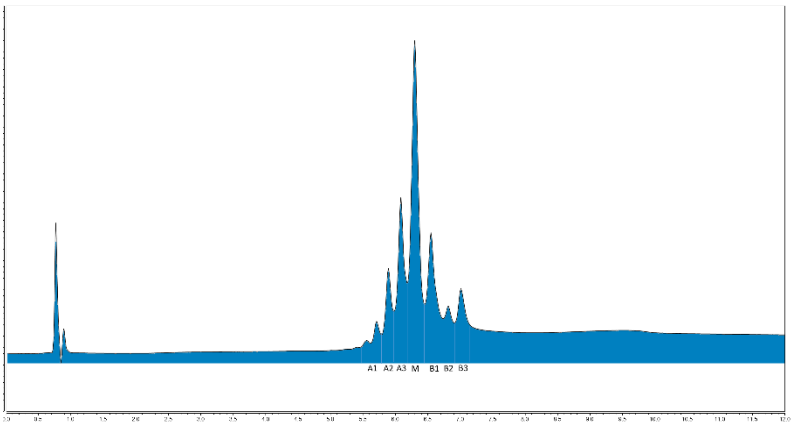


Figure 4.3.18: SCX chromatogram showing fractionation regions of reference drug product

The reference drug product has arguably the most striking distribution of glycoforms across the charge variant fractions. Figure 4.3.21 shows the HILIC chromatogram obtained. There is an observable decrease in highly sialylated forms as the peaks shift into the basic region. The relative abundance of α- Gal containing glycoforms increases compared to other glycan residues. As observed in the HEK and CHO variants there is

an overall decrease in glycan content in the basic peaks. As they are relatively lower in abundance than the acidic variants the contribution of less glycoprotein is also a factor.

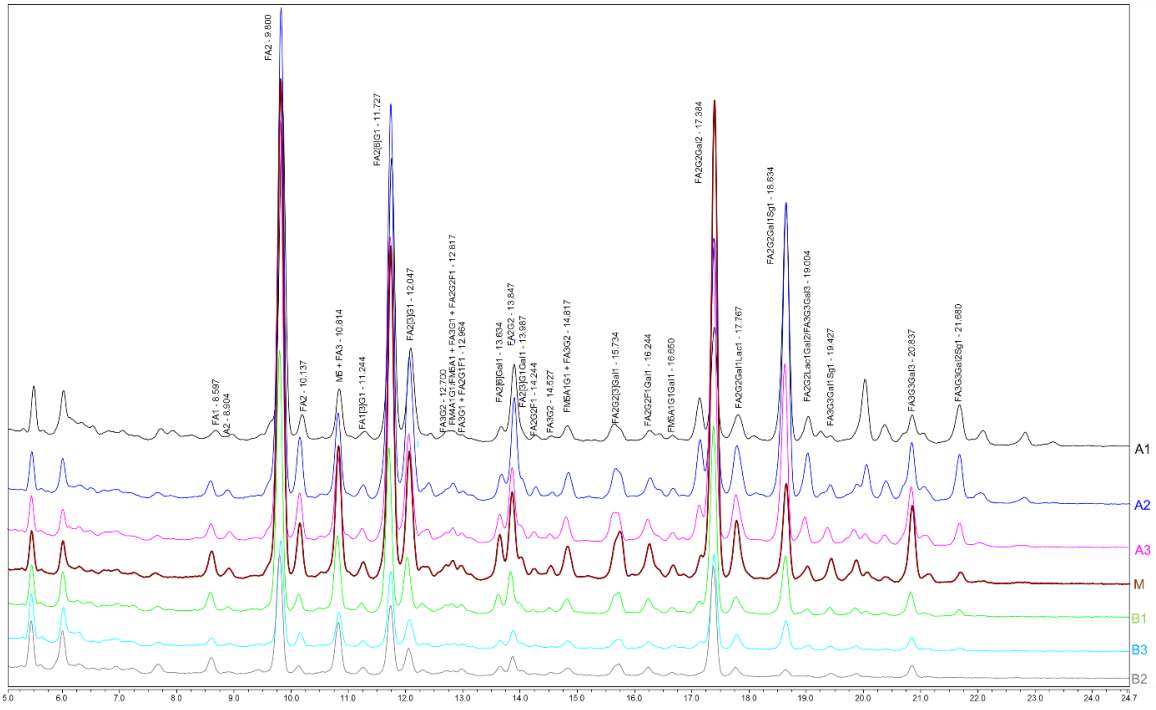


Figure 4.3.19: The distribution and relative abundance of glycoforms across twelve charge variants of the reference drug product.

Chapter 4.

4.3.5. Peptide Mapping

Peptide mapping analysis is one of the primary routes for establishing charge contribution of PTM's. The levels of modification across the charge variant fractions de-glycosylated in the previous section were analysed and compared. As with the glycan profile unique modifications were found for each sample. Biopharma Finder 3.0 was used to identify modifications; Byonic was used as a secondary resource to confirm the accuracy of the Biopharma Finder identifications. The modifications presented here were those which were identified by both software platforms.

4.3.5.1. Expi293 peptide mapping

Sequence coverage is the most fundamentally important features of a peptide mapping experiment. It is an indicator of the quality of the data and infers confidence in the modifications identified. Sequence coverage of ≥90% is required for accurate determination of PTMs. Table 4.3.4, below; shows the averaged % sequence coverage across the three HEK variants fractionated charge variants. Overall, sequence coverage was acceptable, with 86% being the lowest value. Accurate identification of PTM's was still achievable. Tables 4.3.5 and 4.3.6 show the modification identified for the heavy and light chain respectively.

Table 4.3.1: Average sequence coverage for the HEK variants charge variant fractions for peptide mapping analysis

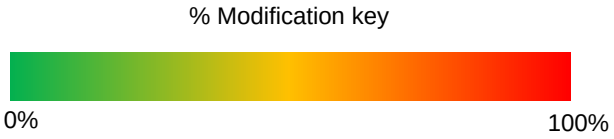
CV Fraction	% Light Chain Sequence Coverage (n = 3)	% Heavy Chain Sequence Coverage (n = 3)
A1	86.13	87.73
A2	88.47	87.50
A3	97.97	98.37
A4	97.33	98.67
A5	92.10	89.63
A6	100.00	99.43
A7	100.00	99.65
M	100.00	99.87
B1	100.00	99.87
B2	100.00	99.87
B3	96.73	95.70
B4	95.97	94.73

C-terminal lysine truncation has been shown to be one of the contributors to acidic variants by the removal of one positively charge molecule [21] [22]. *In vivo*, lysine clipping readily occurs, as discussed in Chapter 3, and due to the residues low impact on functionality it is not considered a CQA. However, in an effort to fully understand the heterogeneity of the in-house produced cetuximab variants, the levels of Lys are considered critical. Table 4.3.5, below, shows the abundance of Lys loss across the variant profiles. Taking into account the sub-optimum sequence coverage for the early acidic peaks no modification was detected. However, A4 has a Lys loss abundance of almost 97%, indicating that the majority of mAbs in this peak have no Lys at the C-terminal. Moving to the main peak, where 75% of peptides had no C-terminal K. As the molecules become more basic, more peptide fragments have C-terminal K.

Chapter 4.

Table 4.3.2: Post translational modifications for the HEK variant Heavy chain with relative abundance across the charge variant fractions

			A1	A2	A3	A4	A5	A6	A7	M	B1	B2	B3	B4
Modification	**Δ Mass**	**Domain**												
K449 Lys-loss	-128.095	Heavy Chain				96.7	78.1	76.3	84.6	75	71.6	61.3	55	
M254 Oxidation	+15.9949	Heavy Chain	49.9	23.7	20	13.2	6.02	6.35	15.3	8.27	6.14	8.39	31.5	
M360 Oxidation	+15.9949	Heavy Chain	42.6	15.5	13.1	8.08	5.05	4.98	10.9	5.05	5.53	7.5	25.3	
M430 Oxidation	+15.9949	Heavy Chain		13.6	10.8	10.1	7.91	9.98	10.7	5.45	15	16.6	17.8	
N205 Deamidated	+0.9840	Heavy Chain		28.2	32.6	29.1	32.6	30.1	31	30	32.7	37	33	
N210 Deamidated	+0.9840	Heavy Chain			23	38.3	35.5	23.8	39	27.2	28.6	46.4		
N288 Deamidated	+0.9840	Heavy Chain	3.4	11.9	13.3	11	8.1	12.1	8.87	9.64	9.75	11.6	13	
N299 Deamidated	+0.9840	Heavy Chain			99.4	99.4	99.1	99.1	99.2	99.1	99.2	99.3		
N317 Deamidated	+0.9840	Heavy Chain	82.1	83.3	85	78.5	76.6	78.1	78.8	72.1	84	85.2	83	89.8
N327 Deamidated	+0.9840	Heavy Chain			9.18	9.1	10.4	13.2		13	7.74	9.08		
N363 Deamidated	+0.9840	Heavy Chain	6.58	9.49	8.27	8.08	8.36	23.1	6.84	26.9	19.6	6.96	2.69	
N386 Deamidated	+0.9840	Heavy Chain	61.6			76.9		87.9	77.2	87.2				
N436 Deamidated	+0.9840	Heavy Chain		8.89	11.4	10.9	10.3	14.5	8.59	13.3	16	15.2	10.5	
N56 Deamidated	+0.9840	Heavy Chain			6.8	6.8	7.17	4.74	3.48	4.57	4.59	5.54		
N88 Deamidated	+0.9840	Heavy Chain						99.9		99.9	99.8			
Q3 Gln->pyro-Glu	-17.0265	Heavy Chain				58.9	93.6			94.4				
Q364 Deamidated	+0.9840	Heavy Chain			2.76	9.59	9.91	9.33	8.1	9.11	10.1	9.86	8.07	
Q364 Gln->pyro-Glu	-17.0265	Heavy Chain					13.4	12.8		13.8	14.9			
Q39 Gln->pyro-Glu	-17.0265	Heavy Chain	53.5	60.7	59.4	56.1	42.7	28.2	49.2	26.1	42.5	42.2	57.9	
Q421 Deamidated	+0.9840	Heavy Chain					0.359	0.483			60	0.833		
Q440 Deamidated	+0.9840	Heavy Chain					0.153	0.12		0.115	0.112	0.123		
Q6 Gln->pyro-Glu	-17.0265	Heavy Chain			20.2	18.8	14	21.5	16	24.8	22.4	26.7	20.9	18.6
Q77 Deamidated	+0.9840	Heavy Chain					0.113	0.111		0.105	0.104	0.111		
Q77 Gln->pyro-Glu	-17.0265	Heavy Chain					52.9	52.4		42.9	52.3	53.2		
W279 Oxidation	+15.9949	Heavy Chain			0.174		0.158	0.0926	0.27	0.165	0.138	0.2		
W36 Oxidation	+15.9949	Heavy Chain					0.338	0.457	0.186	1.04	0.354	0.54		
W52 Oxidation	+15.9949	Heavy Chain					2.7	0.877		0.783	2.09			



Taking the data from a statistical perspective, it is likely that 50% of, pre-trypsin, intact molecules of the main peak have no Lys, at both C-terminals, while the other 50% have 1:0 Lys, contributing to 75% of detected peptides having none, while 25% have one Lys. Going from acidic to basic pI, there is an increase in Lys on molecule, continuing into the basic peaks, up to B3, where 45% of peptides have Lys. This is one of the key findings which help to explain the acidic nature of the HEK variants. Even at the most basic pI, the levels of Lysine clipping are contributing to the overall shift of the profile toward the acidic region.

Interestingly, N-terminal pyro-Glu formation was not detected in the HEK variant; however Q3 at the N-terminal does show modification of almost 100%, but erratically detected across fractions. Peptide mapping coverage, below, shows that the peptide sequence containing Q1 is not an abundant fragment. Though sequence coverage for the one peptide identified is in the highest region, it is a curious occurrence. The hypothesis for this is potential cleavage of the AA sequence or degradation.

There are a large number of Asparagine residues which have been deamidated, the majority of which are low abundance, 5 – 20 %; however some residues have upwards of 80% modification. The Asn residues which are part of glycosylation consensus sequence, Asn 88 and Asn 299, are two such molecules. This is due to the de-glycosylation process wherein asparagine is deamidated to form aspartic acid based on the action of PNGase F. Asn 317 is another highly deamidated residue, this residue is located in the Fc portion of the molecule and is unlikely to affect function, however deamidation is widely accepted as one of the main contributors to acidic charge variants [22].

Oxidation is also a known contributor to local charge heterogeneity on mAbs. Three of four methionine residues of the HEK variant have some level of oxidation present. Dissolved oxygen (DO) content during cell culturing is one of the main pathways of M oxidation[23]. As these mAbs were produced in-house on pilot scale, the DO was not

monitored; this has an observable impact on the quality of the product produced and illustrates the importance of DO in CPP.

Table 4.3.6 shows the modification detected for the light chain of the HEK variant. As expected there were fewer modifications than the HC but the same type including high levels of deamidation, oxidation and pyro-Glu formation. All known contributors to acidic variants.

The most interesting modification detected was light chain N41 glycosylation, consisting of high mannose type N-glycans. It is widely known that cetuximab has a third consensus sequence for glycosylation on the light chain variable region (Figure 4.3.14 and Figure 4.3.15).

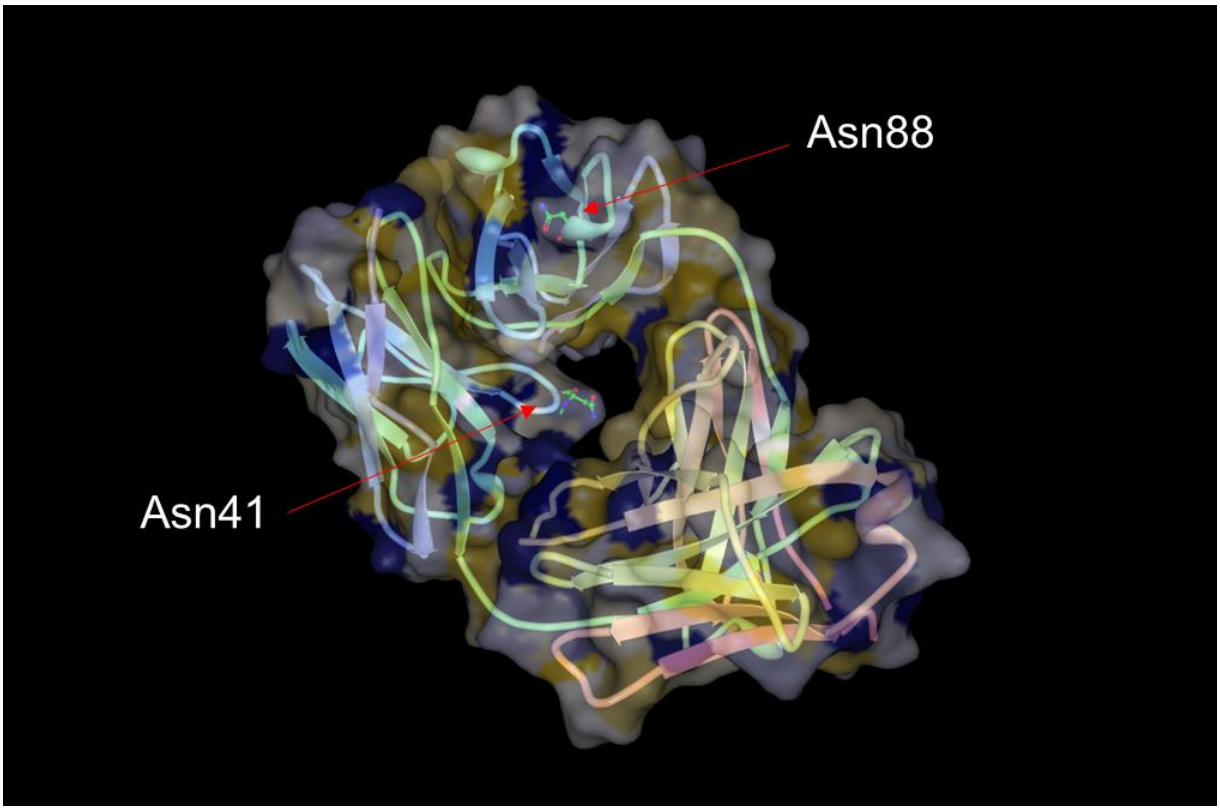


Figure 4.3.20: Crystal structure, with hydrophobicity view, of cetuximab Fab region, indicating the consensus sequences for glycosylation. Crystalisation obtained from Protein Data Bank

Chapter 4.

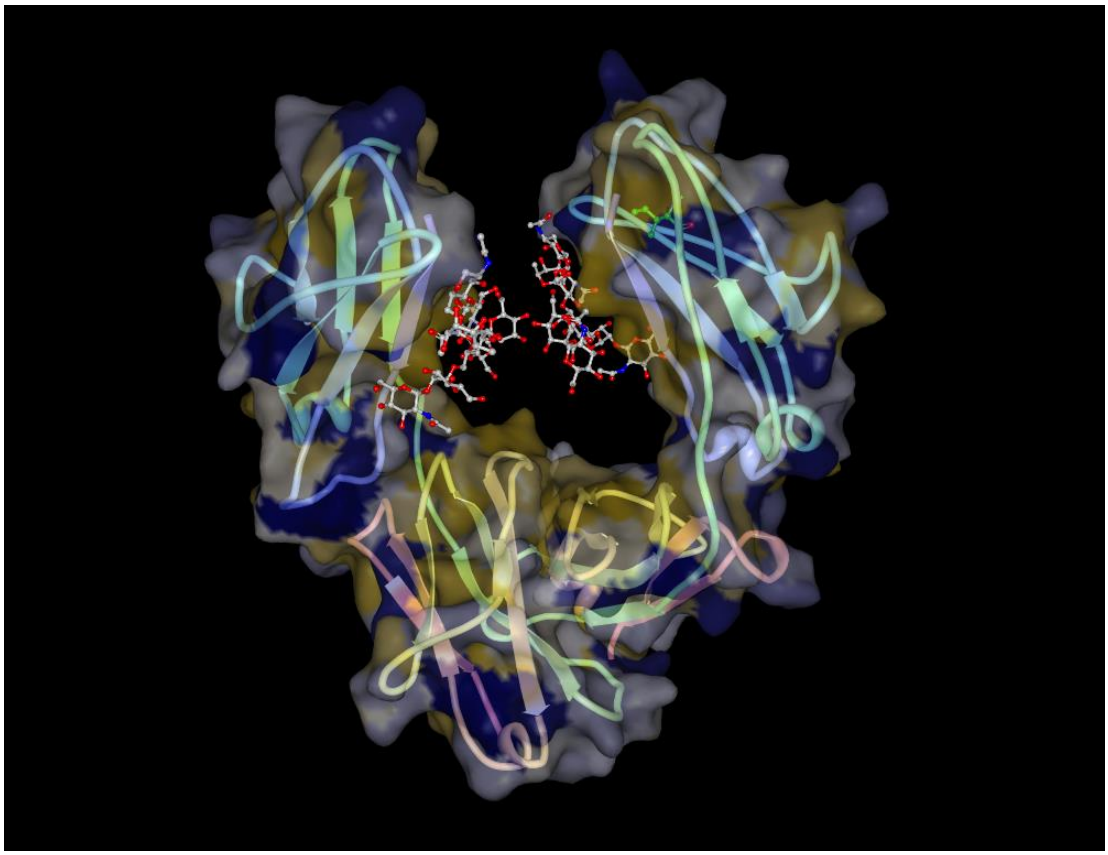


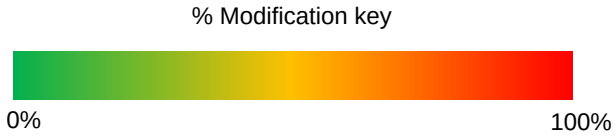
Figure 4.3.21: Crystal structure of the human IgG1 Fc region with FA2 glycosylation. Crystallisation obtained from Protein Data Bank

Though this is site not normally occupied on the RDP, in the in-house variants it was. The effects of cell culturing process and cell line choice have an obvious impact here. Expi293 cells have been specially engineered for high titre protein production. This could have a downstream effect of high sensitivity to glycosylation sequences. Considering the relatively simple glycan profile of CHO, these cells have been clonally selected to produce high titres of protein but potentially at the expense of correct protein processing and conformation. Glycosylation is known to act as a protein folding checkpoint, the occurrence of high mannose structures is indicative of either a lack of complex glycan building enzymes or a rapid expulsion of the protein into the cytosol, before these enzymes have had time to work. This is the hypothesis for the occurrence of N41 glycosylation and is an example of the importance of full product characterisation. These site specific glycoforms were not detected by traditional means, however the presence of these glycans explain the occurrence of high mannose forms on the Fab region.

Chapter 4.

Table 4.3.3: Post translational modifications for the HEK variant Light chain with relative abundance across the charge variant fractions

			A1	A2	A3	A4	A5	A6	A7	M	B1	B2	B3	B4
Modification	**Δ Mass**	**Domain**												
N137 Deamidated	+0.9840	Light Chain		14.9	15.2	14.1	15.6	16.9	20.3	15.7	17.7	18.5	18.7	
N158 Deamidated	+0.9840	Light Chain		13.1	14	11.1	14.8	14.6	10	13.8	15.3	14.9	14.4	
N210 Deamidated	+0.9840	Light Chain		1.51	1.69	1.75	29.5	29	1.36	18.1	28	26.4	1.42	
N32 Deamidated	+0.9840	Light Chain			1.31	1.23	1.02	0.967	0.813	1	1.01	1.06		
N41 Deamidated	+0.9840	Light Chain		92.2	93.5	92.3	93.1	71	86.8	70.6	94	94.3	91.5	
N41 N-glycan	+1216.4229 (M5)	Light Chain					13.1	9.43	5.77	6.27	3.71			
N41 N-glycan	+1378.4757 (M6)	Light Chain					5.05	3.54		2.82	2.16			
N41 N-glycan	+1540.5285 (M7)	Light Chain					3.4	2.56		2.58	1.79			
N41 N-glycan	+1702.5813 (M8)	Light Chain						1.27		1.76	1.67			
N76 Deamidated	+0.9840	Light Chain					6.57	41.3		76.5	35.6	22.8		
N92 Deamidated	+0.9840	Light Chain		19.7	24.1	27.4	28.1	33.3	21.3	33.3	32.9	28.2		29.4
Q147 Deamidated	0.9840	Light Chain					0.0801	0.0956		0.0749	0.0885	0.114		
Q147 Gln->pyro-Glu	-17.0265	Light Chain					15.5	14.5		18.4				
Q155 Gln->pyro-Glu	-17.0265	Light Chain					12.7	12.2		11.1	13	12.6		
Q199 Deamidated	+0.9840	Light Chain			0.741	0.749	0.143	0.118	0.486	0.186	0.101	0.503		
Q199 Gln->pyro-Glu	-17.0265	Light Chain					23.1	21.4		19.9	22.1	3.64		
Q37 Gln->pyro-Glu	-17.0265	Light Chain					94.3	93.2		93.2	97.8	97.4		
Q6 Gln->pyro-Glu	-17.0265	Light Chain				24	26.9	27.9		25.3	23.1	21.7		
W148 Oxidation	+15.9949	Light Chain					1.15	1.13		1.33	1.15	2.21		
W35 Oxidation	+15.9949	Light Chain					0.119	0.087		0.104	0.0894	0.25		
W94 Oxidation	+15.9949	Light Chain					7.24	12.9	9.01	13.3	6.4	4.99		



4.3.5.2. ExpiCHO peptide mapping

The CHO variant had low levels of sequence coverage across both peptide matching platforms. The relative standard deviation was again high and is attributable to one biological replicates exceptionally low sequence coverage. However using both Byonic and Biopharma Finder as orthogonal identification methods, resulted in some insight into the charge contribution of various modifications.

Table 4.3.4: Average sequence coverage for the CHO variants charge variant fractions for peptide mapping analysis

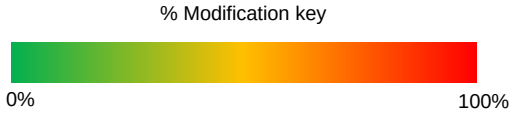
CV Fraction	% Light Chain Sequence Coverage (n = 3)	% Heavy Chain Sequence Coverage (n = 3)
A1	47.37	23.50
A2	64.70	43.10
A3	75.23	68.53
A4	74.30	68.90
M	86.43	76.73
B1	70.10	67.97
B2	75.40	68.37
B3	57.80	54.47
B4	57.00	47.90

Table 4.3.8 shows the modification summary and relative abundance across the charge variants of the CHO cetuximab. C-terminal Lys loss was detected in the main, B1, B2 and B4 peaks of the CHO variant. The relative abundances of the modification are contrasting to what was seen in the HEK variants. The main peak consists of 21.5% modified variants; almost 80% of C-terminal peptides have intact Lys residues. B2 has a high abundance of Lys loss. Considering the overall shift toward the acidic region the possible explanation for this is that the majority of isoforms have lost Lys and very few have intact C-terminal K to contribute to basic charge. However, this does not fully explain the inconsistency as it is widely known that basic charge variants are predominantly due to Lys residues. Further investigation into this occurrence by CVA-MS is required to fully understand the reasons behind this.

Chapter 4.

Table 4.3.5: Post translational modifications for the CHO variant Heavy chain with relative abundance across the charge variant fractions

			A1	A2	A3	A4	M	B1	B2	B3	B4
Modification	**Δ Mass**	**Domain**									
K449 Lys-loss	-128.0950	Heavy Chain					21.5	59.5	94.8		32.4
M254 Oxidation	+15.9949	Heavy Chain	67.4	53.6	39	29.4	10.8	67.5	15.4	43.5	59.1
M360 Oxidation	+15.9949	Heavy Chain		42.6	16.2	12.2	3.01	29.3	4.87	25.6	40.3
M430 Oxidation	+15.9949	Heavy Chain			15.8	12	5.44	5.19	8.29		
N161 Deamidated	+0.9840	Heavy Chain					21				
N205 Deamidated	+0.9840	Heavy Chain			19.1	25.5	21.5	29.6	24.2	51	
N210 Deamidated	+0.9840	Heavy Chain			9.98		13.5		12.1		
N288 Deamidated	+0.9840	Heavy Chain			7.93	9.68	8.16	11.3	9.9	10.2	8.98
N299 Deamidated	+0.9840	Heavy Chain			98.8	99	99.6		99.5		
N31 Deamidated	+0.9840	Heavy Chain			1.74	1.55	4.07		1.86		
N317 Deamidated	+0.9840	Heavy Chain	76.6	77.6	79.1	77.5	75.3	82.1	78.8	75.5	77.5
N327 Deamidated	+0.9840	Heavy Chain			1.86		5.07	4.85	1.24		
N363 Deamidated	+0.9840	Heavy Chain			2.04	2.11	8.24	11.5	6.79		
N386 Deamidated	+0.9840	Heavy Chain		67.9	77.5	65.4	72.8		72.9		64.4
N423 Deamidated	+0.9840	Heavy Chain			12.5		2.35		2.35		
N436 Deamidated	+0.9840	Heavy Chain			11.2	9.87	11.5	17.2	9.78		
N56 Deamidated	+0.9840	Heavy Chain			3.67	3.68	4.38	4.41	4.27		
N70 Deamidated	+0.9840	Heavy Chain			20.6	28.4	40.5		41.7		
N88 Deamidated	+0.9840	Heavy Chain					100	99.5			
Q1 Deamidated	+0.9840	Heavy Chain					24.3				
Q1 Gln->pyro-Glu	-17.0265	Heavy Chain					99.9		99		98.3
Q364 Deamidated	+0.9840	Heavy Chain			5.91	6.42	8.92		7.99		
Q39 Deamidated	+0.9840	Heavy Chain					0.159				
Q39 Gln->pyro-Glu	-17.0265	Heavy Chain		65.2	64.8	59.3	28.6	59.1	49.4	54.7	64
Q6 Gln->pyro-Glu	-17.0265	Heavy Chain	40.5		23.5	21.1	25.2	27	24.6	33.3	32.7
Q77 Deamidated	+0.9840	Heavy Chain					0.0815		0.0981		
W279 Oxidation	+15.9949	Heavy Chain					0.193	3.21	0.41		
W315 Oxidation	+15.9949	Heavy Chain							0.156		
W36 Oxidation	+15.9949	Heavy Chain					0.422		0.521		
W419 Oxidation	+15.9949	Heavy Chain					0.216				
W47 Oxidation	+15.9949	Heavy Chain					0.379				



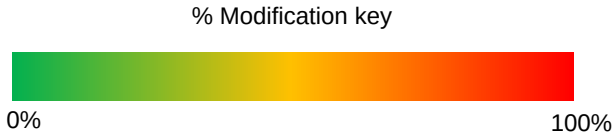
Other modification contributing to the acidic profile of the CHO variants are extensive deamidation across a high number of asparagine residues. Similar to the HEK variants Asn 317 is one of the more heavily modified residues. This residue is most likely located in a solvent accessible area and is prone to deamidation. N-terminal pyro-Glu formation is a modification monitored during characterisation studies [24]. However it has no impact on structure or function but is an indicator of molecule stability. In the CHO variants high levels of pyro-Glu formation were identified in the main, B2 and B4 peaks. Pyro-Glu is a PTM which contributes to acidic variation, this is another instance on this variant, where a known acidic modification is occurring in the basic region and strengthens the evidence that the modifications at the amino acid level are the main reasons for the mAbs pI shift.

Glycosylation at N41 was also found on the CHO variant, though to a lesser extent than HEK. M5 and M7 glycans were detected on the main variant, Table 4.3.9, below. Other modification on the light chain were relatively low abundant with few being higher than 20%. N41 deamidation was the highest, further implicating this site as a glycosylation site. As this residue is buried within the protein structure it is not readily accessible and so natural deamidation would be at similar levels to other Asn residues; however the activity of PNGase F would readily induce such levels.

Chapter 4.

Table 4.3.6: Post translational modifications for the CHO variant light chain with relative abundance across the charge variant fractions

			A1	A2	A3	A4	M	B1	B2	B3	B4
Modification	**Δ Mass**	**Domain**									
N137 Deamidated	+0.9840	Light Chain	22.2		19.6	17.3	15.4	20.7	20.5		
N152 Deamidated	+0.9840	Light Chain			4.58		5.65		6.39		
N158 Deamidated	+0.9840	Light Chain	9.18		13.2	15.4	11.9	13.9	11.1	12	13.7
N210 Deamidated	+0.9840	Light Chain			0.929	9.94	28.7	20.8	10.7		
N32 Deamidated	+0.9840	Light Chain			0.822	0.626	1.18	0.377	0.937		
N41 Deamidated	+0.9840	Light Chain			89.2	87.5	70.5	76.8	85.8	89	
N41 N-Glycan	+1216.4229 (M5)	Light Chain					2.92				
N41 N-Glycan	+1540.5285 (M7)	Light Chain					0.673				
N76 Deamidated	+0.9840	Light Chain					13.9	2.74	6.9		
N91 Deamidated	+0.9840	Light Chain			6.8	31.2	7.24	15.3		0.649	
Q147 Deamidated	+0.9840	Light Chain					0.0906				
Q199 Deamidated	+0.9840	Light Chain			0.442	0.287	0.322		0.6		
Q6 Gln->pyro-Glu	-17.0265	Light Chain					32.5				
W148 Oxidation	+15.9949	Light Chain					0.629		0.686		
W35 Oxidation	+15.9949	Light Chain					0.101				
W94 Oxidation	+15.9949	Light Chain					0.941	0.944			



4.3.5.3. Reference Drug Product peptide mapping

Sequence coverage for the RDP was higher quality compared to the in-house variants.

Table 4.3.10, below shows the coverage obtained.

Table 4.3.7: Sequence coverage for the RDP charge variant fractions for peptide mapping analysis

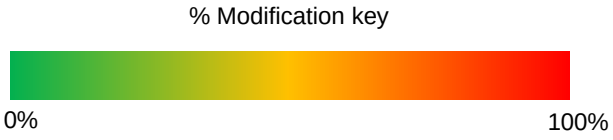
CV Fraction	Light Chain Sequence Coverage (%)	Heavy Chain Sequence Coverage (%)
A1	100.00	100.00
A2	100.00	99.60
A3	100.00	100.00
M	100.00	100.00
B1	100.00	100.00
B2	99.10	100.00
B3	90.70	88.20

Post translational modification of the RDP was significantly less than the in-house variants. The PTMs and their relative abundance on the heavy and light chains are sown together in Table 4.3.11. C-terminal Lys loss is more uniform in the RDP compared to the variants, the increase in Lys residue identification as the molecule becomes more basic is quite evident. This supports the previous observations that extensive Lysine clipping is contributing to the overall acidic nature of the in-house variants. Heavy deamidation occurs at the sites previously identified, Asn317, Asn88 and Asn299. Interestingly no deamidation is detected at the N41 site, indicating that there are no glycan presents in this case. Pyro-Glu formation was in all isoforms, at an abundance of 99%. Ayoub *et al.* showed that pyro-Glu formation at the N terminus of cetuximab was a common occurrence. Oxidation of M and W residues was also reduced in the RDP compared to the in-house variants.

Chapter 4.

Table 4.3.8: Post translational modifications for the RDP heavy chain and light chain with relative abundance across the charge variant fractions

			A1	A2	A3	M	B1	B2	B3
Modification	**Δ Mass**	**Domain**							
K449 Lys-loss	-128.095	Heavy Chain	89.8	87.5	79.2	67.5	26.3	8.2	15
M254 Oxidation	+15.9949	Heavy Chain	11.4	9.22	6.52	5.18	9.21	16.9	
M430 Oxidation	+15.9949	Heavy Chain	6.11		23.1	5.74	6.43	9.48	
N205 Deamidated	+0.984	Heavy Chain	17.4		24.9	28.4	22.6	20.5	
N210 Deamidated	+0.984	Heavy Chain			27.5	32.8			
N288 Deamidated	+0.984	Heavy Chain	11.1	11.5	8.43	10.3	8.84	9.98	
N299 Deamidated	+0.9840	Heavy Chain	99.5		99.4	99.3	99.5		
N31 Deamidated	+0.9840	Heavy Chain	2.37	2.21	2.77	1.8	2.24	2.04	
N317 Deamidated	+0.9840	Heavy Chain	80.9	82.8	81.5	81	81.7	82.4	
N327 Deamidated	+0.9840	Heavy Chain			74.2				
N363 Deamidated	+0.9840	Heavy Chain	1.98	1.84	6.69	6.32	7.6	1.96	
N386 Deamidated	+0.9840	Heavy Chain	71.3	76	81.9	85			
N436 Deamidated	+0.9840	Heavy Chain	7.8	12.1	21.8	17.7	11.4		
N56 Deamidated	+0.9840	Heavy Chain			3.59	3.66	3.8		
N60 Deamidated	+0.9840	Heavy Chain	5.34			4.4			
N88 Deamidated	+0.9840	Heavy Chain				99.9			
Q1 Gln->pyro-Glu	-17.0265	Heavy Chain	99.9	99.9	99.9	99.9	99.9	99.9	99.9
Q421 Deamidated	+0.9840	Heavy Chain			43				
Q6 Gln->pyro-Glu	-17.0265	Heavy Chain			15.2	13.8	20.5	36.5	36.3
Q77 Deamidated	+0.9840	Heavy Chain			0.0816	0.0787			
W279 Oxidation	+15.9949	Heavy Chain			0.1	0.0747			
N137 Deamidated	+0.9840	Light Chain	18.8	18.8	14.1	17.5	15.7	16.9	
N158 Deamidated	+0.9840	Light Chain	7.25	10.9	11	20.1	8.64	8.25	
N210 Deamidated	+0.9840	Light Chain	0.972		0.917	1.4	0.928	0.985	
N32 Deamidated	+0.9840	Light Chain	0.954		0.954	0.96	0.824		
Q155 Deamidated	+0.9840	Light Chain	5.97	4.12			4.26		
Q199 Deamidated	+0.9840	Light Chain	0.562		0.476	0.216	0.579		
Q6 Deamidated	+0.984	Light Chain			0.063	0.0579			
Q6 Gln->pyro-Glu	-17.0265	Light Chain				14			
Q90 Deamidated	+0.984	Light Chain				5.83	12.3	28.7	
W35 Oxidation	+15.9949	Light Chain				0.106			



Across the sample set (Figure 4.3.22), deamidation of Asp residues are the most common modification. Due to the extensive sample processing procedures required to analyse these molecules deamidation readily occurs. However, deamidation in the HEK ∩ CHO is most likely attributable to cell culturing and purification. The effects on un-optimised formulation buffer can also be contributing to the charge profile and modification. What is most interesting about his set is the high abundance of oxidation. This is almost certainly due to the cell culture process. In future DO content will have to be closely monitored to ensure modifications of the mAbs are kept to a minimum.

Chapter 4.

The root cause of Asn41 glycosylation will be investigated further, though this is a very interesting observation from a bio-manufacturing perspective, as this site was thought to be a defunct consensus sequence.

The HEK set is mostly composed of pyro-Glu formation; this is quite interesting, though pyro-Glu formation is not a CQA the distinctive levels solely on HEK are worth noting. The CHO set has additional deamidation and oxidation; this incriminates the cell culture process further. The relatively longer transfection time could be exposing mAbs to the harsh environment outside the cell for longer and resulting in damaged product. With only a one-step purification process being employed, it is most likely that these unfavourable isoforms would be removed by further polishing.

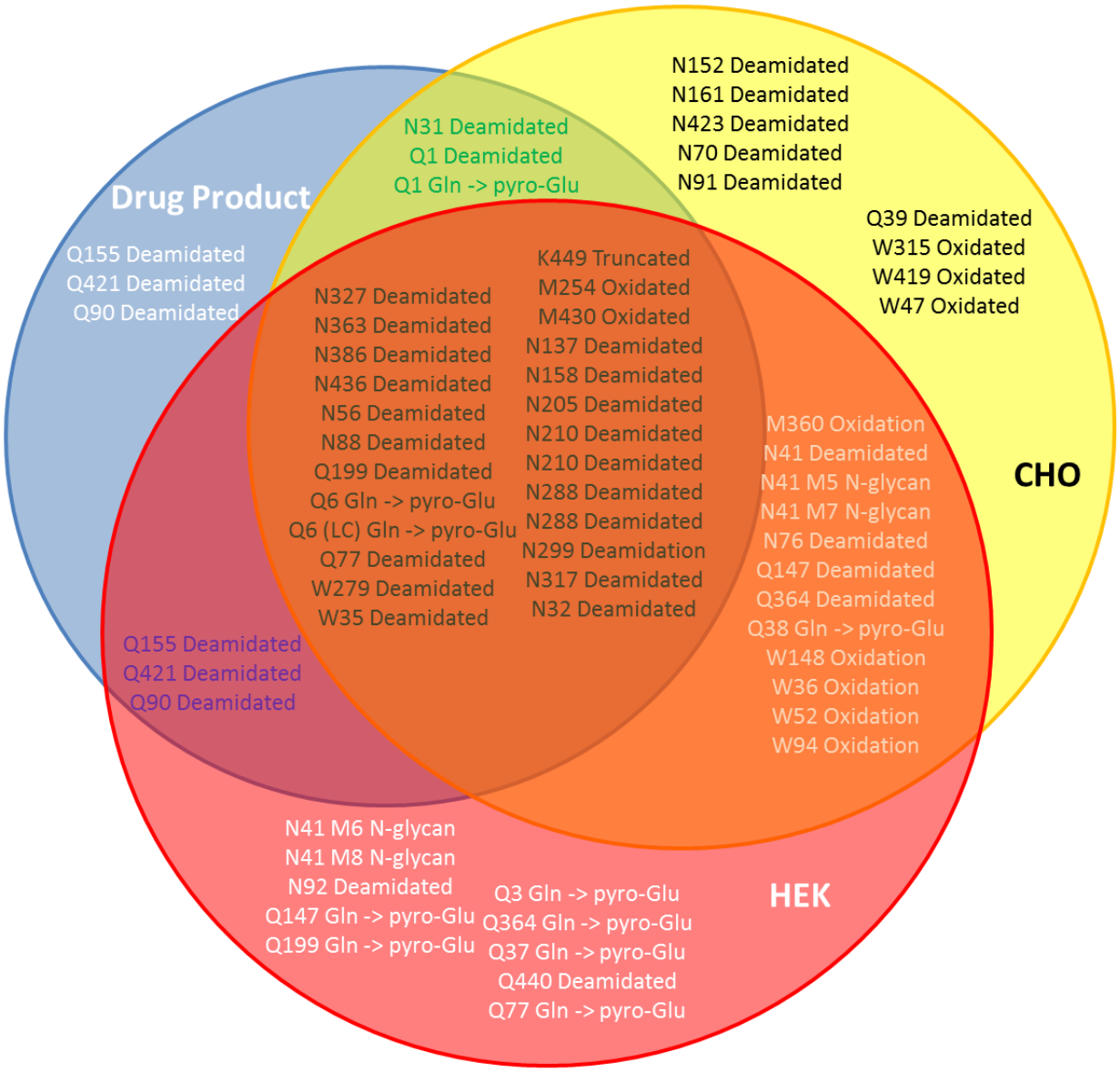


Figure 4.3.22: Venn diagram of the PTMs for the HEK & CHO variants and the reference drug product.

Chapter 4.

4.4. Conclusion & Future directions

In this Chapter, the charge variant profile of in-house produced cetuximab variants were investigated and compared to the reference drug product. N-glycan analysis showed varying degrees of sialylated glycoforms known to contribute to charge variation. No non-human glycan residues were found on the in-house variants. The glycoforms distribution across mAb domains was determined and shown to be consistent with physiological theories of glyco-enzyme access. This is evident by the fact that more complex glycans were found on Asn88, which is easily accessed by enzymes. Simpler high mannose type glycans were predominantly found on the Fc portion. Analysis of the glycans released form charge variant fractions showed a decrease of sialylated form going from the acidic variants to the basic variants. Levels of glycosylation also decreased indicating that glycan structures as a whole could contribute to negative charge on the molecule.

Peptide mapping provided a more in-depth insight into the modifications contributing to charge heterogeneity. The in-house variants, compared to the RDP showed higher levels of deamidation and oxidation. The HEK variant showed a higher propensity for pyro-Glu formation while the CHO variants had extensive deamidation and oxidation. The RDP had fewer modifications than either but still showed considerable levels of deamidation. These deamidation sites are most likely the result of sample processing as the de-glycosylation procedure exposes the molecule to adverse conditions which induce deamidation. However, cell culturing conditions are also implicated in the extended deamidation observed on the in-house variants. Dissolved oxygen content of the cell culture media is expected to be the reason behind the high levels of oxidation.

Further work on this subject will involve, close monitoring of cell culture conditions with experimentation on the levels of DO. The effects purification process has on molecules will also be investigated, to determine the loss of highly acidic variants through additional polishing steps such as IEC to remove unfavourable charge variants. As an immediate future direction, analysis of cetuximab by CVA-MS to identify intact C-terminal Lys

abundance is needed as well as determining the true effects sample processing has on deamidation levels.

4.5. Author Contributions

Research study devised by Anne Trappe, Stefan Mittermayr, Silvia Millan Martin and Jonathan Bones; Cell culturing performed by Anne Trappe; Sample preparation performed by Anne Trappe, N-glycan analysis of drug product performed by Silvia Millan Martin; LC-MS performed by Anne Trappe; Data analysis performed by Anne Trappe; Chapter written by Anne Trappe and reviewed by Jonathan Bones.

Chapter 4.

4.6. References

1.Qian, J.; Liu, T.; Yang, L.; Daus, A.; Crowley, R.; Zhou, Q., Structural characterization of N-linked oligosaccharides on monoclonal antibody cetuximab by the combination of orthogonal matrix-assisted laser desorption/ionization hybrid quadrupole–quadrupole time-of-flight tandem mass spectrometry and sequential enzymatic digestion. *Anal.Biochem.* **2007,** *364* (1), 8-18.

2.Zhang, H.; Berezov, A.; Wang, Q.; Zhang, G.; Drebin, J.; Murali, R.; Greene, M. I., ErbB receptors: from oncogenes to targeted cancer therapies. *Journal of Clinical Investigation* **2007,** *117* (8), 2051-2058.

3.Agency, E. M., European Public Assessment Report for Erbitux. 2004.

4.Bauman, J. E.; Grandis, J. R., Targeting secondary immune responses to cetuximab: CD137 and the outside story. *The Journal of Clinical Investigation* **2014,** *124* (6), 2371-2375.

5.Neumann, J.; Zeindl-Eberhart, E.; Kirchner, T.; Jung, A., Frequency and type of KRAS mutations in routine diagnostic analysis of metastatic colorectal cancer. *Pathology - Research and Practice* **2009,** *205* (12), 858-862.

6.Gerdes, C. A.; Nicolini, V. G.; Herter, S.; van Puijenbroek, E.; Lang, S.; Roemmele, M.; Moessner, E.; Freytag, O.; Friess, T.; Ries, C. H.; Bossenmaier, B.; Mueller, H. J.; Umaña, P., GA201 (RG7160): A Novel, Humanized, Glycoengineered Anti-EGFR Antibody with Enhanced ADCC and Superior In Vivo Efficacy Compared with Cetuximab. *Clinical Cancer Research* **2013,** *19* (5), 1126-1138.

7.Patel, D.; Xuemei, G.; Ng, S.; Melchior, M.; Balderes, P.; Burtrum, D.; Persaud, K.; Luna, X.; Ludwig, D. L.; Xiaoqiang, K., IgG isotype, glycosylation, and EGFR expression determine the induction of antibody-dependent cellular cytotoxicity in vitro by cetuximab. *Human Antibodies* **2010,** *19* (4), 89-99.

8.Chung, C. H.; Mirakhur, B.; Chan, E.; Le, Q.-T.; Berlin, J.; Morse, M.; Murphy, B. A.; Satinover, S. M.; Hosen, J.; Mauro, D.; Slebos, R. J.; Zhou, Q.; Gold, D.; Hatley, T.;

Chapter 4.

Hicklin, D. J.; Platts-Mills, T. A. E., Cetuximab-Induced Anaphylaxis and IgE Specific for Galactose-α-1,3-Galactose. *New England Journal of Medicine* **2008,** *358* (11), 1109-1117.

9.Zhang, P.; Woen, S.; Wang, T.; Liau, B.; Zhao, S.; Chen, C.; Yang, Y.; Song, Z.; Wormald, M. R.; Yu, C.; Rudd, P. M., Challenges of glycosylation analysis and control: an integrated approach to producing optimal and consistent therapeutic drugs. *Drug Discovery Today* **2016,** *21* (5), 740-765.

10.Beck, A.; Sanglier-Cianférani, S.; Van Dorsselaer, A., Biosimilar, Biobetter, and Next Generation Antibody Characterization by Mass Spectrometry. *Analytical Chemistry* **2012,** *84* (11), 4637-4646.

11.Millan Martin, S.; Delporte, C.; Farrell, A.; Navas Iglesias, N.; McLoughlin, N.; Bones, J., Comparative analysis of monoclonal antibody N-glycosylation using stable isotope labelling and UPLC-fluorescence-MS. *Analyst 140* (5), 1442-1447.

12.Ayoub, D.; Jabs, W.; Resemann, A.; Evers, W.; Evans, C.; Main, L.; Baessmann, C.; Wagner-Rousset, E.; Suckau, D.; Beck, A., Correct primary structure assessment and extensive glyco-profiling of cetuximab by a combination of intact, middle-up, middle-down and bottom-up ESI and MALDI mass spectrometry techniques. *MAbs 5* (5), 699-710.

13.Böhm, E.; Seyfried, B. K.; Dockal, M.; Graninger, M.; Hasslacher, M.; Neurath, M.; Konetschny, C.; Matthiessen, P.; Mitterer, A.; Scheiflinger, F., Differences in N-glycosylation of recombinant human coagulation factor VII derived from BHK, CHO, and HEK293 cells. *BMC Biotechnology* **2015,** *15* (1), 87.

14.Croset, A.; Delafosse, L.; Gaudry, J.-P.; Arod, C.; Glez, L.; Losberger, C.; Begue, D.; Krstanovic, A.; Robert, F.; Vilbois, F.; Chevalet, L.; Antonsson, B., Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells. *Journal of Biotechnology* **2012,** *161* (3), 336-348.

15.Göritzer, K.; Maresch, D.; Altmann, F.; Obinger, C.; Strasser, R., Exploring Site-Specific N-Glycosylation of HEK293 and Plant-Produced Human IgA Isotypes. *Journal of Proteome Research* **2017,** *16* (7), 2560-2570.

16.Carillo, S.; Mittermayr, S.; Farrell, A.; Albrecht, S.; Bones, J., Glycosylation Analysis of Therapeutic Glycoproteins Produced in CHO Cells. *Methods in molecular biology (Clifton, N.J.)* **2017,** *1603*, 227-241.

17.Beck, A.; Debaene, F.; Diemer, H.; Wagner-Rousset, E.; Colas, O.; Dorsselaer, A. V.; Cianférani, S., Cutting-edge mass spectrometry characterization of originator, biosimilar and biobetter antibodies. *Journal of Mass Spectrometry 50* (2), 285-297.

18.Sakae, Y.; Satoh, T.; Yagi, H.; Yanaka, S.; Yamaguchi, T.; Isoda, Y.; Iida, S.; Okamoto, Y.; Kato, K., Conformational effects of N-glycan core fucosylation of immunoglobulin G Fc region on its interaction with Fcγ receptor IIIa. *Scientific Reports* **2017,** *7* (1), 13780.

19.Bones, J.; McLoughlin, N.; Hilliard, M.; Wynne, K.; Karger, B. L.; Rudd, P. M., 2D-LC Analysis of BRP 3 Erythropoietin N-Glycosylation using Anion Exchange Fractionation and Hydrophilic Interaction UPLC Reveals Long Poly-N-Acetyl Lactosamine Extensions. *Analytical Chemistry* **2011,** *83* (11), 4154-4162.

20.Du, Y.; Walsh, A.; Ehrick, R.; Xu, W.; May, K.; Liu, H., Chromatographic analysis of the acidic and basic species of recombinant monoclonal antibodies. *mAbs* **2012,** *4* (5), 578-585.

21.Hintersteiner, B.; Lingg, N.; Zhang, P.; Woen, S.; Hoi, K. M.; Stranner, S.; Wiederkum, S.; Mutschlechner, O.; Schuster, M.; Loibner, H.; Jungbauer, A., Charge heterogeneity: Basic antibody charge variants with increased binding to Fc receptors. *MAbs* **2016,** *8* (8), 1548-1560.

22.Leblanc, Y.; Ramon, C.; Bihoreau, N.; Chevreux, G., Charge variants characterization of a monoclonal antibody by ion exchange chromatography coupled on-line to native

mass spectrometry: Case study after a long-term storage at +5°C. *Journal of Chromatography B* **2017,** *1048* (Supplement C), 130-139.

23.Ghesquière, B.; Jonckheere, V.; Colaert, N.; Van Durme, J.; Timmerman, E.; Goethals, M.; Schymkowitz, J.; Rousseau, F.; Vandekerckhove, J.; Gevaert, K., Redox Proteomics of Protein-bound Methionine Oxidation. *Molecular & Cellular Proteomics : MCP* **2011,** *10* (5), M110.006866.

24.Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianférani, S., Characterization of Therapeutic Antibodies and Related Products. *Analytical Chemistry* **2013,** *85* (2), 715-736.

Chapter 5.

5. CVA-MS Analysis of Cetuximab Drug Product for orthogonal confirmation of glycan structures

5.1. Introduction

One of the major limitations of denaturing intact-MS is the limited complexity of samples which can be analysed, due to overlap of signals which have similar m/z values[1]. Native MS analysis allows more complex samples to be analysed, as the protein is still conformationally folded and so exhibits less charge, also generating greater spatial spectral resolution in the acquired data[2].

Chapter 3 illustrated the benefit of using CVA-MS for native analysis of charge variant isoforms. The wealth of information which could be gained from a single injection was immense. The samples used during that study were mAbs which had one glycosylation site, with a relatively simple N-glycan profile. The next challenge was to use this method to analyse a more complex sample. In this chapter, the cetuximab reference product was analysed by CVA-MS, using the information gained on the glycan profile from Chapter 4 for accurate annotation and localisation of complex glycoforms.

Intact and native MS identification of complex glycoforms is now achievable with state of the art instrumentation. There are many studies reported on the analysis of large, intact biomolecules using native MS to accurately quantify mass isoforms[3,4,5]. However, sample preparation for these studies involved de-glycosylation of the molecule prior to analysis. Subsequent annotation of the major glycoforms is performed as a separate analysis. Wohlschlager *et al.* (2018) demonstrated a native MS analysis of the fusion protein Etanercept [2]. This molecule has four N-glycosylation sites, which results in relatively high heterogeneity. Annotation of the glycan structures in native form was not possible due to the complexity of spectra obtained, thus, de-glycosylation and de-sialylation of the molecule enriched signal intensity for less abundant forms. This work was an excellent approach to tackling the complex signals of these highly glycosylated molecules. The native MS procedure involved buffer exchange of the molecule to

ammonium acetate followed by direct infusion[2]. In that work a statistical software program was used to annotate glycoforms up to the desialylated sample set. This software, termed Mofi, performs a two-stage search algorithm which initially assigns monosaccharide compositions to each peak and subsequently searches a list of glycan combinations which match with the identified monosaccharides[6]. This enabled Wohlschlager *et al.* to overcome the perceived limitations of the method and resulted in the accurate annotation of complex glycoprotein structures.

In this chapter, the CVA-MS approach provides an extra separation prior to MS analysis which enables glycoform annotation *in situ,* without the need to de-glycosylate.

Heck and co-workers have worked intensively to natively analyse glycopeptides, however prior deglycosylation has been required [7,8,9,10]. However, Yang *et al.* (2013) reported on the native-MS analysis of the chicken ovalbumin glycoprotein profile. Structural annotation was achieved by reference to previous profiling experiments[11]. In this Chapter, a similar approach is taken, with the exception of prior separation with SCX and performed on a considerably larger, therapeutically relevant protein.

More common methods for separation prior to MS analysis of intact glycopeptides includes CZE-MS and SEC-MS [12,13]. However, these methods are limited to proteins of smaller size, compared to mAbs.

Chapter 5.

5.2. Materials & methods

5.2.1. Reagents

Water (Optima™, LC-MS grade, Catalogue No 10505904) was provided by Fisher Scientific. Acetic acid (ACS reagent grade, ≥ 99.7%, Catalogue No A0808), ammonium bicarbonate (BioUltra, ≥ 99.5%, Catalogue No 09830) and ammonium hydroxide solution (BioUltra, 1M in H_2O, Catalogue No 09859) were purchased from Sigma-Aldrich.

The commercially available monoclonal antibody cetuximab was provided by the Hospital Pharmacy Unit of the University Hospital of San Cecilio in Granada, Spain.

Carboxypeptidase B (Catalogue No 10103233001) was obtained from Hoffmann-La Roche AG.

5.2.2. Sample Preparation for CVA-MS analysis

Carboxypeptidase digestion of the intact antibody was performed by the addition of the enzyme to obtain a substrate-enzyme ratio of 5:1 (w/w). Digestion was performed at 37 °C for 2 hours with agitation at 450 rpm.

5.2.3. CVA-MS Analysis

Separations were performed on a Thermo Scientific Vanquish™ Horizon UHPLC system equipped with a Fluorescence Detector F, a Split Sampler HT, a Binary Pump H and Column Compartment. Buffer A was 6.25 mM ammonium bicarbonate, 7.5 mM acetic acid in LC-MS grade water. Buffer B was 5 mM ammonium hydroxide in LC-MS grade water. A MAbPac™ SCX-10 RS column with dimensions of 2.1 × 50 mm and 5 µm particle size was used. The column oven was held at 25 °C, a flow rate of 0.4 mL/min^{-1} was applied.

Chapter 5.

Table 5.2.1: Optimised gradient for cetuximab

Time (min)	% buffer B
0	20
15	100
18	20
30	20

One hundred µg were injected per run and analysis was performed on a Q Exactive™ Plus mass spectrometer hybrid quadrupole Orbitrap mass spectrometer with extended mass Biopharma Option. Data acquisition was performed in triplicate. LC and MS systems were hyphenated *via* a Heated Electrospray Ionization-II (HESI-II) probe in a standard Ion Max ion source. The MS tune parameters used are shown in Table 5.2.2. The MS method parameters are shown in Table 5.2.3.

Table 5.2.2: Tune file parameters

Capillary Voltage (kV)	3.6
Capillary Temperature (°C)	275
Sheath Gas (AU)	20
Auxiliary Gas	5
Probe Heater Temp (°C)	275
S-Lens RF Level	200
High Mass Range mode	On
Trapping Gas Pressure Setting	1

Table 5.2.3: Method parameters

Polarity	Positive
Collision Induced Dissociation (eV)	150
Microscans	10
Resolution (@ m/z 200)	17,500; 35,000
Automatic Gain Control (AGC)	3e6
Maximum Inject Time (ms)	200
Mass Range (m/z)	2,500 – 8000

Deconvolution of spectra of main lysine variants was performed using the Sliding Window deconvolution feature. The average molecular masses obtained after deconvolution were compared to the theoretical cetuximab sequence mass considering major glycoforms. Fixed modifications were set at 16 disulphide bonds and 2 N-terminal pyroGlu.

Table 5.2.4: Parameters for sliding window deconvolution

Parameter	
m/z range	5,000 – 7,000
Deconvolution Algorithm	ReSpect
Model Mass Range	150,000 – 156,000
Mass Tolerance (ppm)	102
Charge State Range	20 – 30
Minimum Adjacent Charges	3 - 3
Retention Time Range	4 – 15
Target Avg. Spectrum Width (min)	0.3
Target Avg. Spectrum Offset (%)	25%
Merge Tolerance	12
Max RT gap (min)	0.250
Minimum number of detected intervals	3

Chapter 5.

5.3. Results & Discussion

CVA-MS analysis of the cetuximab reference product provided a more global view of the charge heterogeneity present. In Chapter 4, peptide mapping showed that the main peak consisted of a mix of isoforms which had lost C-terminal lysine. CVA-MS enabled more accurate annotation of the charge variant peaks and improved resolution enabled the identification of those which had not been detected using UHPLC alone. Nine unique charge variants were identified using this method, compared to seven in Chapter 4.

5.3.1. CVA-MS Analysis

CVA-MS analysis showed the distribution of C-terminal K moving across the charge profile. Figure 5.3.1 shows the charge variant distribution of the RDP with annotated peaks. The acidic variants, peaks 1 – 4, have a high abundance of sialylated glycans, some with up to 4 sialic acids across the molecule. The most acidic variants have no C-terminal lysine, consistent with the previous peptide mapping results. There is a steady decrease in sialylated forms an increase in C-terminal K to the main peak, where equilibrium between sialylation and lysine charge contribution is reached. Interestingly, sialylated forms are still present on some basic isoforms, peaks 6 and 8. This supports the observation in Chapter 4, where these types of glycans were still present in the charge variant fractions, Figure 4.3.21.

Chapter 5.

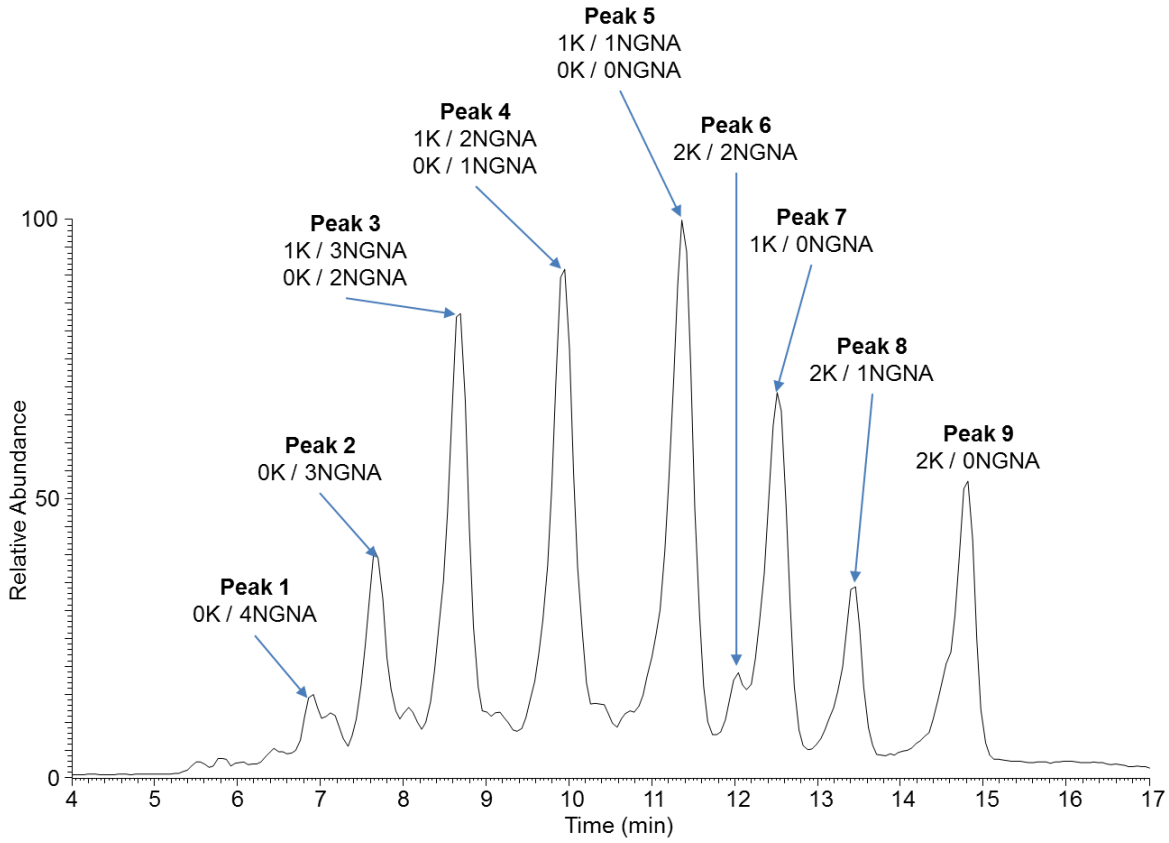


Figure 5.3.1: CVA-MS BPC of cetuximab drug product.

CpB digestion of the RDP resulted in a loss of peaks 6 – 9, confirming the identification of these variants, as having C-terminal Lysine. Figure 5.3.2, below shows the base peak chromatograms obtained.

Chapter 5.

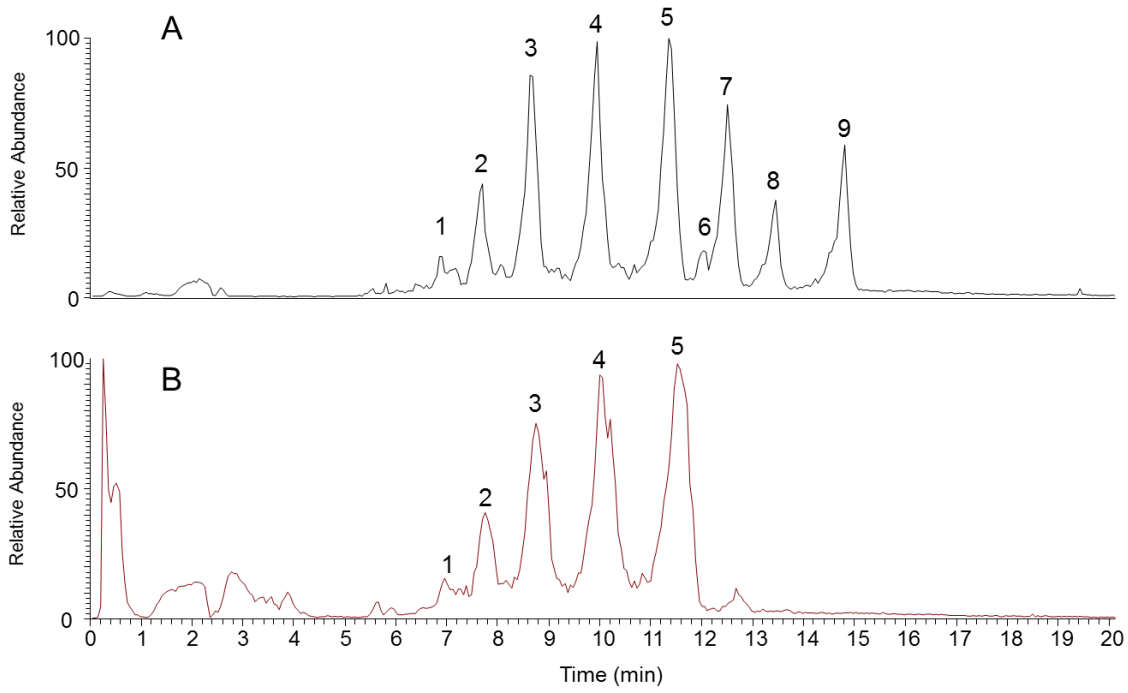


Figure 5.3.2: CVA-MS analysis of (A) Native RDP and (B) CpB digested RDP

Deconvolution was performed using the Sliding Windows algorithm in BioPharma Finder 3.0. Figure 5.3.3 shows the deconvoluted spectra obtained for peak 5, the main charge isoform.

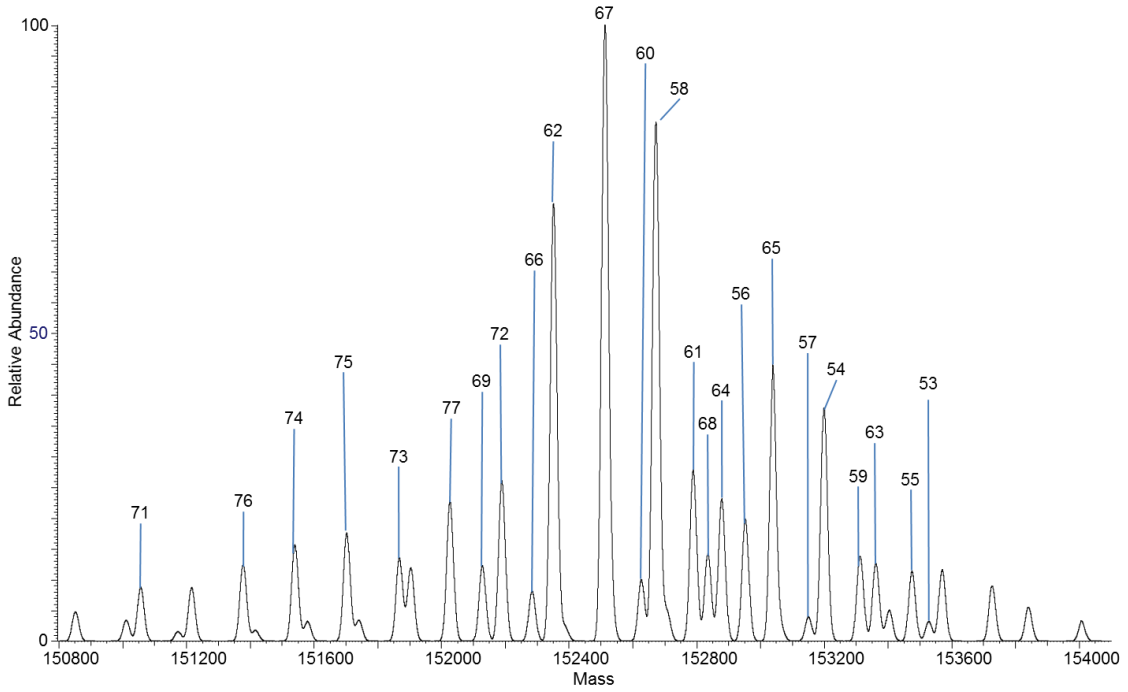


Figure 5.3.3: Deconvoluted Spectra of main charge variant, peak 5. Peak labels correspond to glycan forms in Table 5.3.1.

Chapter 5.

The main charge isoform peak was found to consist of 24, baseline separated, unique spectral features. Due to the complexity of the data derived from this experiment, structural confirmation could not be attained. The empirical formulas for the glycans, found, are given in Table 5.3.1 with the number of C-terminal K, the theoretical and average masses, relative abundance and Δ ppm. The mass difference for this experiment was quite high relative to previous studies, as much as 30 ppm. This is most likely due to low sample quantity and closely eluting isobaric substances. However, 104 isoforms were annotated; the most abundant forms for each peak are given in Table 5.3.2.

Chapter 5.

Table 5.3.1: Table of isoforms detected by the Sliding Windows Algorithm for CVA-MS analysis of cetuximab RDP.

Peak	Isoform number	Average Mass (mean in Da)	Relative Abundance	Variant (XK_FcGly1/FcGly2_FabGly1/FabGly2)	Theoretical mass (Da)	Δm (ppm)
1	1	153254.78	10.11	0K_H4N4F1/H4N4F1_H6N4F1NGNA2/H6N4F1NGNA2	153256.64	-12.09
1	2	153779.42	11.20	0K_H4N4F1/H4N4F1_H6N4F1NGNA2/H8N5F1NGNA2	153784.12	-30.62
1-2	3	152947.93	39.05	0K_H3N4F1/H4N4F1_H6N4F1NGNA2/H6N4F1NGNA1	152949.38	-9.46
1-2	4	154163.53	5.75	0K_H4N4F1/H4N4F1_H8N5F1NGNA2/H8N5F1NGNA1	154166.50	-19.28
2	5	153110.28	36.25	0K_H4N4F1/H4N4F1_H6N4F1NGNA2/H6N4F1NGNA1	153111.52	-8.12
2	6	153313.63	20.00	0K_H3N4F1/H3N4F1_H6N4F1NGNA2/H8N5F1NGNA1	153314.72	-7.12
2	8	153271.03	4.22	0K_H4N4F1_H5N4F1_H6N4F1NGNA2/H6N4F1NGNA1	153273.67	-17.19
2	9	152786.94	27.40	0K_H3N4F1/H3N4F1_H6N4F1NGNA2/H6N4F1NGNA1	152787.23	-1.94
2	10	153635.57	16.18	0K_H4N4F1/H4N4F1_H6N4F1NGNA2/H8N5F1NGNA1	153639.01	-22.39
2	11	153475.01	30.56	0K_H3N4F1/H4N4F1_H6N4F1NGNA2/H8N5F1NGNA1	153476.87	-12.08
2-3	12	152803.44	57.52	0K_H3N4F1/H4N4F1_H6N4F1NGNA1/H6N4F1NGNA1	152804.26	-5.40
2-3	13	153329.52	35.74	0K_H3N4F1/H4N4F1_H6N4F1NGNA1/H8N5F1NGNA1	153331.75	-14.58
2-3	14	153077.59	2.08	1K_H3N4F1/H4N4F1_H6N4F1NGNA1/H6N4F1NGNA2	153077.55	0.27
3	15	153242.72	3.07	1K_H4N4F1/H4N4F1_H6N4F1NGNA1/H6N4F1NGNA2	153239.70	19.75
3	16	153694.46	4.24	0K_H3N4F1/H3N4F1_H8N5F1NGNA1/H8N5F1NGNA1	153697.10	-17.16
3	17	154019.78	7.74	0K_H4N4F1/H4N4F1_H8N5F1NGNA1/H8N5F1NGNA1	154021.39	-10.45
3	18	153124.80	9.06	0K_H4N4F1_H5N4F1_H6N4F1NGNA1/H6N4F1NGNA1	153128.55	-24.55
3	19	152964.68	75.32	0K_H4N4F1/H4N4F1_H6N4F1NGNA1/H6N4F1NGNA1	152966.41	-11.31
3	20	152413.17	4.10	0K_H3N4F1/M5_H6N4F1NGNA1/H6N4F1NGNA1	152413.87	-4.61
3	22	152479.98	16.98	0K_H3N4F1/H3N4F1_H6N4F1NGNA1/H5N4F1NGNA1 or 0K_H3N4F1/H4N4F1_H5N4F1NGNA1/H5N4F1NGNA1	152479.97	0.08
3	23	153858.35	8.59	0K_H3N4F1/H4N4F1_H8N5F1NGNA1/H8N5F1NGNA1	153859.24	-5.80
3	24	153492.22	28.04	0K_H4N4F1/H4N4F1_H6N4F1NGNA1/H8N5F1NGNA1	153493.90	-10.94
3	25	153168.49	26.50	0K_H3N4F1/H3N4F1_H6N4F1NGNA1/H8N5F1NGNA1	153169.61	-7.32
3	26	152641.51	66.56	0K_H3N4F1/H3N4F1_H6N4F1NGNA1/H6N4F1NGNA1 or H4N4F1/H4N4F1_H5N4F1NGNA/H5N4F1NGNA1	152642.12	-3.99
3	27	152317.65	7.15	0K_H3N4F1/H3N4F1_H5N4F1NGNA1/H5N4F1NGNA1	152317.83	-1.15
3-4	31	153094.59	14.47	1K_H4N4F1/H4N4F1_H6N4F1NGNA1/H6N4F1NGNA1	153094.58	0.07
4	33	152772.49	4.29	1K_H3N4F1/H3N4F1_H6N4F1NGNA1/H6N4F1NGNA1	152770.29	14.40
4	34	153458.58	8.74	1K_H3N4F1/H4N4F1_H6N4F1NGNA1/H8N5F1NGNA1	153459.92	-8.76
4	35	152933.62	25.72	1K_H3N4F1/H4N4F1_H6N4F1NGNA1/H6N4F1NGNA1	152932.44	7.77
4	36	153874.72	6.55	0K_H4N4F1/H4N4F1_H8N5F1NGNA1/H9N5F1	153876.27	-10.11
4	37	153620.45	5.91	1K_H4N4F1/H4N4F1_H6N4F1NGNA1/H8N5F1NGNA1	153622.07	-10.54
4	39	152496.01	65.49	0K_H3N4F1/H3N4F1_H6N4F1NGNA1/H7N4F1	152497.00	-6.54
4	40	153713.78	10.36	0K_H3N4F1/H4N4F1_H8N5F1NGNA1/H9N5F1	153714.13	-2.23
4	41	152429.14	2.58	0K_H3N4F1/M5_H7N4F1/H6N4F1NGNA1	152430.90	-11.53
4	42	153185.71	51.36	0K_H3N4F1/H4N4F1_H6N4F1NGNA1/H9N5F1	153186.64	-6.06
4	43	152336.69	17.11	0K_G0F/G0F_H6N4F1NGNA1/H6N4F1	152334.86	12.05
4	44	153347.00	26.41	0K_H4N4F1/H4N4F1_H6N4F1NGNA1/H9N5F1	153348.78	-11.66
4	45	152268.35	5.09	0K_H3N4F1/M5_H7N4F1/H6N4F1NGNA1	152268.75	-2.67
4	46	153023.96	26.85	0K_H3N4F1/H3N4F1_H6N4F1NGNA1/H9N5F1	153024.49	-3.48
4	47	152657.48	97.21	0K_H3N4F1/H4N4F1_H7N4F1/H6N4F1NGNA1	152659.15	-10.94
4	48	152819.59	100.00	0K_H4N4F1/H4N4F1_H7N4F1/H6N4F1NGNA1	152821.30	-11.17
4	49	152172.29	11.92	0K_H3N4F1/H3N4F1_H6N4F1/H5N4F1NGNA1	152172.71	-2.76

Chapter 5.

Peak	Isoform number	Average Mass (mean in Da)	Relative Abundance	Variant (XK_FcGly1/FcGly2_FabGly1/FabGly2)	Theoretical mass (Da)	Δ_m (ppm)
4-5	53	153524.65	2.15	0K_H5N4F1/H5N4F1_H7N4F1/H9N5F1	153527.96	-21.58
4-5	54	153202.19	20.42	0K_H4N4F1/H4N4F1_H7N4F1/H9N5F1	153203.67	-9.66
4-5	55	153476.60	8.07	1K_H4N4F1/H4N4F1_H7N4F1/H8N5F1NGNA1	153476.96	-2.34
4-5	56	152951.09	20.42	1K_H4N4F1/H4N4F1_H7N4F1/H6N4F1NGNA1	152949.47	10.63
4-5	57	153150.30	4.88	1K_H3N4F1/H3N4F1_H7N4F1/H8N5F1NGNA1	153152.67	-15.45
4-5	58	152675.10	48.17	0K_H4N4F1/H4N4F1_H7N4F1/H7N4F1	152676.18	-7.08
4-5	59	153315.08	14.86	1K_H3N4F1/H4N4F1_H7N4F1/H8N5F1NGNA1	153314.81	1.74
4-5	60	152628.14	14.83	1K_H3N4F1/H3N4F1_H7N4F1/H6N4F1NGNA1	152625.18	19.43
4-5	61	152789.24	30.29	1K_H3N4F1/H4N4F1_H6N4F1NGNA1/H7N4F1	152787.32	12.57
5	62	152352.42	54.58	0K_H3N4F1/H3N4F1_H7N4F1/H7N4F1 or 0K_H4N4F1/H4N4F1_H6N4F1/H6N4F1	152351.89	3.46
5	63	153361.86	7.45	0K_H4N4F1/H5N4F1_H7N4F1/H9N5F1	153365.82	-25.80
5	64	152878.10	21.85	0K_H3N4F1/H3N4F1_H7N4F1/H9N5F1	152879.38	-8.39
5	65	153040.69	44.53	0K_H3N4F1/H4N4F1_H7N4F1/H9N5F1	153041.53	-5.45
5	66	152286.74	4.00	0K_H3N4F1/M5_H7N4F1/H7N4F1	152285.79	6.24
5	67	152513.74	81.32	0K_H3N4F1/H4N4F1_H7N4F1/H7N4F1	152514.04	-1.92
5	68	152835.70	8.20	0K_H4N4F1_H5N4F1_H7N4F1/H7N4F1	152838.33	-17.18
5	69	152125.73	11.83	0K_H3N4F1/M5_H7N4F1/H7N4F1	152123.64	13.71
5	71	151056.22	9.43	0K_H3N4F1/H3N4F1_H3N4F1/H3N4F1	151054.73	9.86
5	72	152189.77	29.03	0K_H3N4F1/H3N4F1_H7N4F1/H6N4F1 or 0K_H3N4F1/H4N4F1_H6N4F1/H6N4F1	152189.75	0.19
5	73	151863.84	14.03	0K_H3N4F1/H4N4F1_H5N4F1/H5N4F1	151865.45	-10.61
5	74	151539.95	13.21	0K_H5N3F1/H5N3F1_H5N3F1/H5N3F1	151539.10	5.61
5	75	151701.80	21.11	0K_H4N4F1/H4N4F1_H4N4F1/H4N4F1	151703.31	-9.96
5	76	151377.15	14.40	0K_H3N4F1/H4N4F1_H3N4F1/H4N4F1	151379.02	-12.35
6	77	152027.29	15.46	0K_H3N4F1/H3N4F1_H6N4F1/H6N4F1	152027.60	-2.06
6	78	153746.97	5.77	2K_H4N4F1/H4N4F1_H6N4F1NGNA1/H8N5F1NGNA1	153750.24	-21.27
6	79	153061.22	24.78	2K_H3N4F1/H4N4F1_H6N4F1NGNA1/H6N4F1NGNA1	153060.61	3.99
6	80	153222.56	18.02	2K_H4N4F1/H4N4F1_H6N4F1NGNA1/H6N4F1NGNA1	153222.75	-1.24
6	81	152900.02	10.10	2K_H3N4F1/H3N4F1_H6N4F1NGNA1/H6N4F1NGNA1	152898.46	10.21
7	82	153168.39	20.99	1K_H3N4F1/H4N4F1_H7N4F1/H9N5F1	153169.70	-8.51
7	83	152410.94	3.63	1K_H3N4F1/M5_H7N4F1/H7N4F1	152413.96	-19.83
7	84	152480.88	42.47	1K_H3N4F1/H3N4F1_H7N4F1/H7N4F1	152480.06	5.35
7	85	153330.03	15.74	1K_H4N4F1/H4N4F1_H7N4F1/H9N5F1	153331.84	-11.82
7	86	152640.52	61.86	1K_H3N4F1/H4N4F1_H7N4F1/H7N4F1	152642.21	-11.09
7	87	151667.96	7.58	1K_H5N3F1/H5N3F1_H5N3F1/H5N3F1	151667.27	4.54
7	88	152801.02	63.60	1K_H4N4F1/H4N4F1_H7N4F1/H7N4F1	152804.35	-21.80
7	89	152250.95	5.32	1K_H3N4F1/M5_H7N4F1/H7N4F1	152251.81	-5.66
7	90	151828.80	7.00	1K_H4N4F1/H4N4F1_H4N4F1/H4N4F1	151831.48	-17.66
7	91	153006.01	13.29	1K_H3N4F1/H3N4F1_H7N4F1/H9N5F1	153007.55	-10.06
7	93	151506.62	9.11	1K_H3N4F1/H4N4F1_H3N4F1/H4N4F1	151507.19	-3.77
7	94	152317.58	11.20	1K_H3N4F1/H4N4F1_H6N4F1/H6N4F1 or 1K_H3N4F1/H3N4F1_H7N4F1/H6N4F1	152317.92	-2.23
8	97	153076.34	21.64	2K_H4N4F1/H4N4F1_H7N4F1/H6N4F1NGNA1	153077.64	-8.47
8	98	152753.75	17.32	2K_H3N4F1/H3N4F1_H7N4F1/H6N4F1NGNA1	152753.35	2.65
8	99	153445.61	7.72	2K_H3N4F1/H4N4F1_H6N4F1NGNA1/H9N5F1	153442.98	17.10
8	100	152915.40	33.57	2K_H3N4F1/H4N4F1_H7N4F1/H6N4F1NGNA1	152915.49	-0.62
9	102	152934.05	21.32	2K_H4N4F1/H4N4F1_H7N4F1/H7N4F1	152932.53	9.98
9	103	152770.46	32.71	2K_H3N4F1/H4N4F1_H7N4F1/H7N4F1	152770.38	0.49
9	104	152607.18	16.63	2K_H3N4F1/H3N4F1_H7N4F1/H7N4F1	152608.24	-6.90

Table 5.3.2: Summary of most abundant glycans of each peak

Peak #	Isoform #	Average Mass	Theoretical Mass	Glycan Fab	Glycan Fc	K
1	2	153779.42	153784.12	FA2G2Gal1Sg1/FA3G3Gal2Sg2	FA2G1/FA2G1	0
2	5	153110.28	153111.52	FA2G2Gal1Sg2/FA2G2Gal1Sg1	FA2G1/FA2G1	0
3	19	152964.68	152966.41	FA2G2Gal1Sg2/FA2G2Gal1Sg1	FA2G1/FA2G1	0
4	48	152819.59	152821.30	FA2G2Gal2/FA2G2Gal1Sg1	FA2G1/FA2G1	0
5	67	152513.74	152514.04	FA2G2Gal2/FA2G2Gal2	FA2/FA2G1	0
6	79	153061.22	153060.61	FA2G2Gal1Sg1/FA2G2Gal2Sg1	FA2/FA2G1	2
7	88	152801.02	152804.35	FA2G2Gal2/FA2G2Gal2	FA2G1/FA2G1	1
8	100	152915.40	152915.49	FA2G2Gal2/FA2G2Gal1Sg1	FA2/FA2G1	2
9	103	152770.46	152770.38	FA2G2Gal2/FA2G2Gal2	FA2/FA2G1	2

In chapter 4, the most abundant glycoforms detected after de-glycosylation of the charge variant peaks and after IdeS digestion are consistent with what has been detected by CVA-MS. Simultaneous determination of C-terminal lysine clipping with glycan identification was not possible previously. This is one of the key benefits of CVA-MS, which provides whole molecule analysis and a more accurate image of the charge variants. However, prior knowledge of the glycan profile of molecule is required to inform annotation, ideally when considering molecules with more than one glycosylation site.

5.4. Conclusion & Future Directions

This Chapter provided orthogonal confirmation and further information on the constituents of the major isoforms determined in the previous chapter. The major glycoforms were found to be consistent with the site specific and charge fraction specific profile. The C-terminal lysine quantity accuracy was improved upon, as peptide could only describe a quantitative value rather than a qualitative. Chapter 4 provided pieces of

the puzzle, whereas this study, is the picture on the box which makes putting the pieces together far simpler.

The CVA-MS method, has again, shown itself to be useful for garnering information on the whole structure of a mAb and has proven itself further in universality by accurate analysis of one of the most complex mAb therapies.

Further work on this method will include analysis of the in-house variants, to further investigate their charge isoforms. A proposed triage for the molecules, successful transfection permitting, will be initial characterisation with the ultrafast method. Potentially building a CV profile library over-time to fit a processing method which can assign lysine quantity to each peak by retention time. Following ultrafast analysis, the triage will continue to in-depth characterisation by the CVA-MS method, envisaging the statistical program, Mofi, will be incorporated reducing analysis time. This workflow would find application in many areas of production of biopharmaceuticals, particularly in those with multiple glycosylation sites, as this is where the Mofi software analysis is focused.

Potential application of this technology is in the field of biotransformation of mAbs during pharmacokinetic studies to facilitate elucidation of modification of these complex molecules *in vivo*.

5.5. Author Contributions

Research study devised by Florian Füssl, Jonathan Bones and Anne Trappe; Sample analysis performed by Anne Trappe. Chapter written by Anne Trappe and reviewed by Jonathan Bones.

Chapter 5.
5.6. References

1.Thompson, N. J.; Hendriks, L. J. A.; de Kruif, J.; Throsby, M.; Heck, A. J. R., Complex mixtures of antibodies generated from a single production qualitatively and quantitatively evaluated by native Orbitrap mass spectrometry. *mAbs* **2014,** *6* (1), 197-203.

2.Wohlschlager, T.; Scheffler, K.; Forstenlehner, I. C.; Skala, W.; Senn, S.; Damoc, E.; Holzmann, J.; Huber, C. G., Native mass spectrometry combined with enzymatic dissection unravels glycoform heterogeneity of biopharmaceuticals. *Nature Communications* **2018,** *9* (1), 1713.

3.Siuti, N.; Kelleher, N. L., Decoding protein modifications using top-down mass spectrometry. *Nature Methods* **2007,** *4*, 817.

4.Terral, G.; Beck, A.; Cianférani, S., Insights from native mass spectrometry and ion mobility-mass spectrometry for antibody and antibody-based product characterization. *Journal of Chromatography B* **2016,** *1032*, 79-90.

5.Rosati, S.; Thompson, N. J.; Barendregt, A.; Hendriks, L. J. A.; Bakker, A. B. H.; de Kruif, J.; Throsby, M.; van Duijn, E.; Heck, A. J. R., Qualitative and Semiquantitative Analysis of Composite Mixtures of Antibodies by Native Mass Spectrometry. *Analytical Chemistry* **2012,** *84* (16), 7227-7232.

6.Skala, W.; Wohlschlager, T.; Senn, S.; Huber, G. E.; Huber, C. G., MoFi: A Software Tool for Annotating Glycoprotein Mass Spectra by Integrating Hybrid Data from the Intact Protein and Glycopeptide Level. *Analytical Chemistry* **2018,** *90* (9), 5728-5736.

7.Yang, Y.; Liu, F.; Franc, V.; Halim, L. A.; Schellekens, H.; Heck, A. J. R., Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity. *Nature Communications* **2016,** *7*, 13397.

8.Rosati, S.; Yang, Y.; Barendregt, A.; Heck, A. J. R., Detailed mass analysis of structural heterogeneity in monoclonal antibodies using native mass spectrometry. *Nature Protocols* **2014,** *9*, 967.

9.Dyachenko, A.; Wang, G.; Belov, M.; Makarov, A.; de Jong, R. N.; van den Bremer, E. T. J.; Parren, P. W. H. I.; Heck, A. J. R., Tandem Native Mass-Spectrometry on Antibody–Drug Conjugates and Submillion Da Antibody–Antigen Protein Assemblies on an Orbitrap EMR Equipped with a High-Mass Quadrupole Mass Selector. *Analytical Chemistry* **2015,** *87* (12), 6095-6102.

10.Rosati, S.; van den Bremer, E. T. J.; Schuurman, J.; Parren, P. W. H. I.; Kamerling, J. P.; Heck, A. J. R., In-depth qualitative and quantitative analysis of composite glycosylation profiles and other micro-heterogeneity on intact monoclonal antibodies by high-resolution native mass spectrometry using a modified Orbitrap. *mAbs* **2013,** *5* (6), 917-924.

11.Yang, Y.; Barendregt, A.; Kamerling, J. P.; Heck, A. J. R., Analyzing Protein Micro-Heterogeneity in Chicken Ovalbumin by High-Resolution Native Mass Spectrometry Exposes Qualitatively and Semi-Quantitatively 59 Proteoforms. *Analytical Chemistry* **2013,** *85* (24), 12037-12045.

12.Yang, Y.; Franc, V.; Heck, A. J. R., Glycoproteomics: A Balance between High-Throughput and In-Depth Analysis. *Trends Biotechnol.* **2017,** *35* (7), 598-609.

13. Staub, A.; Guillarme, D.; Schappler, J.; Veuthey, J.-L.; Rudaz, S., Intact protein analysis in the biopharmaceutical field. *J.Pharm.Biomed.Anal.* **2011,** *55* (4), 810-822;

14. Berkowitz, S. A.; Zhong, H.; Berardino, M.; Sosic, Z.; Siemiatkoski, J.; Krull, I. S.; Mhatre, R., Rapid quantitative capillary zone electrophoresis method for monitoring the micro-heterogeneity of an intact recombinant glycoprotein. *Journal of Chromatography A* **2005,** *1079* (1), 254-265.

Chapter 6.
6. Overall Conclusion & Future Work

6.1. Conclusion

Monoclonal antibodies are versatile and powerful therapeutics. They are revolutionising treatment for disease of many different classes. They provide hope for orphan diseases, inflammatory conditions and even late stage cancer. From the first mAb therapy in the 1980's, the specificity and safety of these therapies has rapidly improved. The production and manufacturing process is being constantly reviewed and updated. PTMs have been shown to have a significant impact on the function and safety of protein therapies. Biosimilars are required to have from sequence level up, similarity to the reference product and so in-depth characterisation is required. Chapter 1 outlined the process of developing and manufacturing mAb therapies and the effects these processes can have on structure and function, with an emphasis on post translational modifications. The various "state-of-the-art" techniques which are used to characterise biosimilars were also framed with an emphasis on the technology used in this thesis. The analytical method which provides the most information about a protein is arguably LC-MS. In this work an SCX-MS was developed to determine the PTM's of various mAbs.

The study reported in Chapter 2 was designed to evaluate the best SCX approach for use in characterising mAbs. During the course of this work, a greater understanding of the adsorption enthalpies unique to each separation mode was gained. This aided the development and subsequent validation of an ultra-fast SCX method. The method was thoroughly validated by repeat injections of an in-house produced adalimumab variant. Integration with the Cobra Wizard algorithm enabled rapid detection of the major charge isoforms. The method was then applied to a sample set of in-house variants of trastuzumab, which had been produced by transient transfection in HEK and CHO and a stably transfected CHO variant. Differences in the charge variant patterns were detected and with reference to the drug product, higher levels of acidic variants were present in the transiently expressed mAbs. Peptide mapping of the variants yielded interesting results showing that the stable variant had amino acids still intact from the

leader sequence, resulting in a more acidic profile. This method is envisioned to be applied to a PAT framework, where biotherapeutics can be rapidly analysed, directly from either protein A elution buffer or cell culture media. This would have significant impact on enhancing purification workflows, particularly from a perspective of decision making upon out-of-specification events. From the perspective of biosimilar development, it can be a useful tool for screening molecules which have unfavourable CV profiles, such as the in-house variants analysed here. Rapid methods have been previously reported, however, the high throughput potential was not realised, and this was achieved here.

In Chapter 3, to help understand the complex heterogeneity at play in the in-house variants, a CVA-MS method, developed by Füssl *et al.,* (2018) was modified and tested using stressed adalimumab drug product. The adjustment of buffer pH enabled a broader sample pI to be analysed and optimisation of the MS resolution settings enabled more accurate peak annotation. The resulting method was tested on stressed and unstressed adalimumab, to determine the increasing levels of degradation of the molecule over time. Lysine clipping, deamidation and succinimide formation were all shown to have a significant impact on the molecules charge heterogeneity. The major glycoforms of the charge fractions and extensive levels of glycation were also found to be contributing to heterogeneity and illustrated as a point of common isobaric interference which could be avoided by using this method. Fragmentation of the antibody was found to also significantly contribute to charge heterogeneity, using a specialised post-acquisition filtering method, fragments of the mAb were identified, and determined to not be caused by conventional fragmentation means. The implication of HCP in this fragmentation was investigated by peptide mapping analysis of common HCPs with sequence searching of the adalimumab drug product. Cathepsin L1 was found to most likely be contributing to the fragmentation observed.

The major finding of this study was the detection of HCP in the drug product itself. With more sensitive instrumentation it is now possible to detect well below, previously defined

limits. This has implications for regulatory agencies, where the ultimate effects of these low abundant impurities on patients has not previously been studied. This then falls to the manufacturer and clinicians to ensure that these impurities do not significantly impact the product. However, as was observed in this study, over time, the product degraded significantly. Therefore, this study highlights an area, in current regulatory processes, where LOD and LOQ need to be reviewed with the most current instrumentation and more regularly than at present.

In terms of the analytical field, the findings of this work represent an exciting new method for in-depth characterisation. As the limitations of the previously reported method have been overcome and the expansive knowledge which can be obtained has been illustrated, a new avenue of native separations has been opened up.

To further investigate the heterogeneity of in-house produced mAbs, an in-depth investigation in PTMs of HEK and CHO cetuximab was undertaken in Chapter 4. The CV profiles of these molecules were more complex than anticipated. SEC analysis did not detect any overt aggregation occurring and so glycan analysis of the whole molecule with orthogonal site-specific annotation was performed. The resulting profiles showed varying degrees of sialylated glycoforms known to contribute to charge variation. No non-human glycan residues were found on the in-house variants contrasting to the RDP which historically has glycan forms that have been known to induce anaphylaxis. Analysis of the glycans released form charge variant fractions showed a decrease of sialylated form going from the acidic variants to the basic variants. Levels of glycosylation also decreased indicating that glycan structures as a whole could contribute to negative charge on the molecule. Peptide mapping of the in-house variants was performed, to gain further information on the acidic nature of the molecules. The resulting data showed an additional glycosylation site, N41, on the light chain of both HEK and CHO produced cetuximab as well as significant deamidation and oxidation. These modifications implicated the cell culturing process in the acidic nature of the molecules. The in-house variants also had additional C-terminal Lys clipping compared to the RDP.

Chapter 6.

Chapter 4 is potentially the most comprehensive study into cetuximab, both of the reference product and of a biosimilar, to date. This molecule has an infamous glycan profile and in this work, that profile was examined from all potential views, illustrating the link between amino acid sequence, tertiary structure and glycosylation. The effects of uninhibited cell culturing were unexpected. As the molecules produced in-house, exhibited wild type glycosylation sites which were previously thought to be unoccupied. Interestingly, these N-glycan sites were not detected through routine glycan workflows but from the glycopeptide. This highlights the importance of native analysis, as this shows the molecule as a whole and highlights the potential gaps which sample preparation creates.

The purpose of this chapter was ultimately to feed into the annotation of data obtained from the native CVA-MS analysis of cetuximab.

In Chapter 5, the CVA-MS method was applied to the cetuximab drug product. Using the information gained in Chapter 4 on the site-specific glycosylation pattern present on the molecule, nine major isoforms were accurately identified. The empirical formula of the glycan structures were reported and 104 unique combination of Lys and glycans, across two glycosylation sites were shown. The trend of no Lys/highly sialylated rapidly flipped to 2 Lys/low sialylation going from acidic to basic peaks. This method provided an excellent global view into the intact dynamic at work in the charge heterogeneity of cetuximab, one of the most complex biotherapeutics on market. Essentially, this study is the tip of the iceberg in revealing what information can be gleaned using the CVA-MS. From Chapter 3, it is evident that there are many more discoveries to be found with this molecule, particularly in the field of glycomics. Potentially, for the first time, it will be possible to achieve site-specific, statistical data on the abundance of glycans found at each glycosylation site. This will have significant impact on drug design and efficacy/safety studies. The biggest hurdle is data analysis time, however, with the incorporation of statistical modelling software, Mofi, into the workflow this should decrease significantly. Currently, the main application of this method is in comparability

studies; however there are many avenues to be explored. These are discussed further in section 6.2.

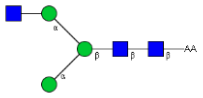
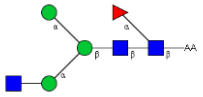
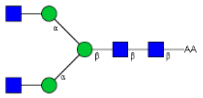
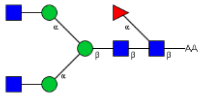
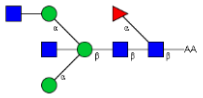
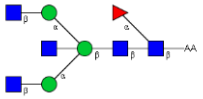
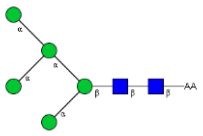
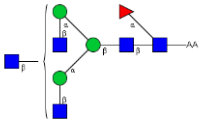
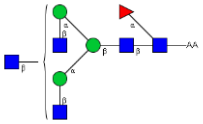
Chapter 6.
6.2. Future Work

The research presented in this thesis has demonstrated the potential for hyphenating native separations of intact proteins direct to native high resolution mass spectrometry. This ideal ‘native-native’ combination is ideally suited for investigating further problems encountered in biopharmaceutical production such as the characterisation of more complex proteins such as Fc fusion proteins, antibody drug conjugates or modification of the method to deal with heavily glycosylated proteins such as erythropoietin. The ability to rapidly characterise the molecule as an intact entity is exciting, combination with MS/MS as used in top-down proteomics may facilitate localisation of identified modifications. Combination of the developed technology with clinical sample analysis may facilitate biotransformation and metabolic modification studies of these therapeutic proteins upon administration.

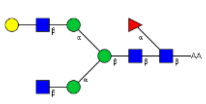
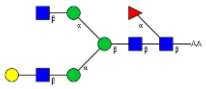
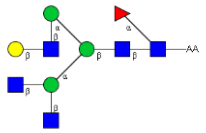
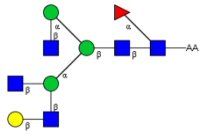
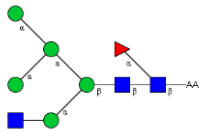
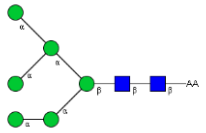
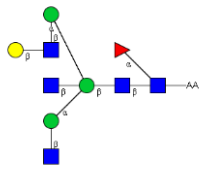
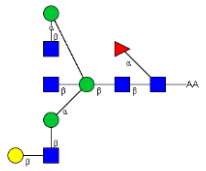
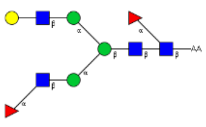
Additionally, the presented research opens up the potential for more fundamental research in the area of structural biology. The advent of native mass spectrometry has facilitated an in depth understanding of molecular complexes involved in a variety of biochemical events within the cell. However, as mentioned in the introduction to Chapter 5, many of these studies rely on buffer exchange of the sample into ammonium acetate followed by static nano-infusion into the mass spectrometer. The potential of the developed native LC-MS approach offers exciting potential to deepen our understanding within structural biology by providing a technological solution to separate various forms of these protein complexes but while maintaining them as an intact entity, that may by differentially involved in a biological process, information that is lost due to time averaging when using nano-infusion. Such a platform could be applied to many fundamental research questions to deepen our understanding of cellular biology in a variety of physiological states.

Appendix A
Supplementary Data

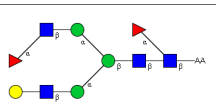
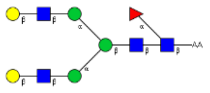
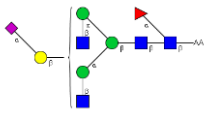
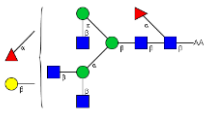
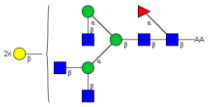
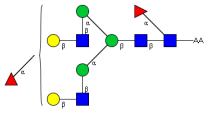
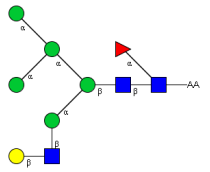
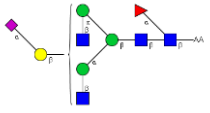
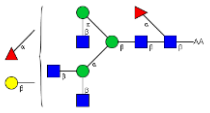
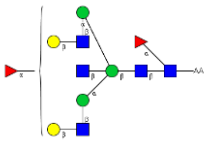
Table 1-C: Glycan assignments for each HILIC peak with corresponding m/z and proposed structures for HEK variant

RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
7.63	1	616.2235	616.2227	1.30	A1	
8.65	2	1379.5101	1379.5105	0.29	FA1	
		689.2528	689.2516	1.74		
8.95	3	717.7638	717.7624	1.95	A2	
9.85	4	1582.5883	1582.5899	1.01	FA2	
		790.7917	790.7913	0.51		
10.18	5	790.7921	790.7913	1.01	FA2/FA1b	
10.68	6	892.3323	892.3310	1.46	FA3	
10.85	7	1354.4788	1354.4789	0.07	M5	
		676.7366	676.7358	1.18		
11.30	8	892.3319	892.3310	1.01	FA3	
11.48	9	892.3320	892.3310	1.12	FA3	

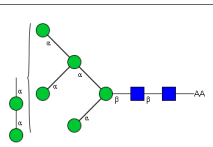
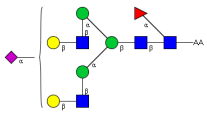
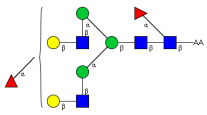
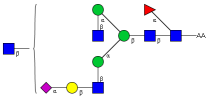
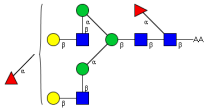
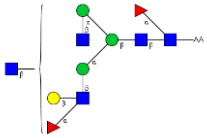
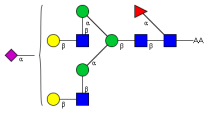
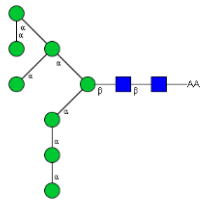
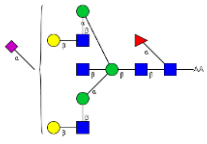
Appendix A

RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
11.77	10	871.8187	871.8177	1.15	FA2[6]G1	
		1744.6374	1744.6427	3.04		
12.08	11	871.8188	871.8177	1.26	FA2[3]G1	
		1744.6407	1744.6427	1.15		
12.33	12	973.3587	973.3574	1.34	FA3[6]G1	
12.60	13	973.3586	973.3574	1.23	FA3[3]G1	
12.83	14	851.3056	851.3045	1.29	FM5A1	
13.00	15	757.7635	757.7622	1.72	M6	
13.06	16	973.3584	973.3574	1.03	FA2b[6]G1	
13.30	17	973.3588	973.3574	1.44	FA2b[3]G1	
13.60	18	944.8477	944.8467	1.06	FA2[6]G1F1	

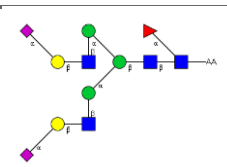
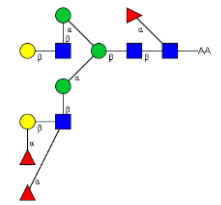
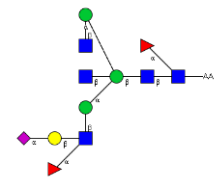
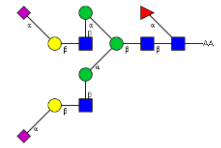
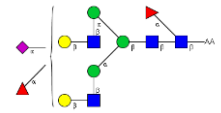
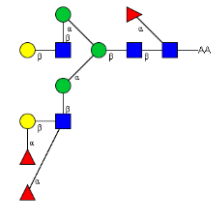
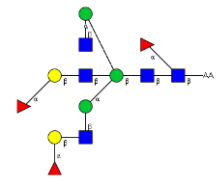
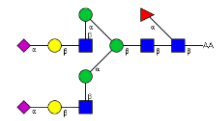
Appendix A

RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
13.83	19	944.8480	944.8467	1.38	FA2[3]G1F1	
13.89	20	952.8455	952.8441	1.47	FA2G2	
		1017.3668	1017.3654	1.38	FA2G1S1	
14.06	21	1046.3880	1046.3864	1.53	FA3G1F1	
14.17	22	1054.3847	1054.3838	0.85	FA3G2	
14.58	23	1025.8746	1025.8731	1.46	FA2G2F1	
14.82	24	932.3300	932.3309	0.97	FM5A1G1	
		1017.3668	1017.3654	1.38	FA2G1S1	
		1046.3890	1046.3864	2.48	FA3G1F1	
15.04	25	1127.4136	1127.4128	0.71	FA3G2F1*	

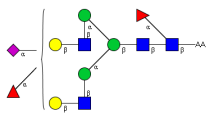
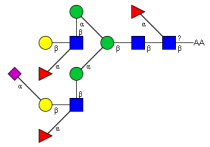
Appendix A

RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
		838.7901	838.7886	1.79	M7	
15.46	26	1098.3930	1098.3919	1.00	FA2G2S1	
		1025.8751	1025.8731	1.95	FA2G2F1	
		1118.9059	1118.9051	0.71	FA3G1S1	
15.53	27	1025.8751	1025.8731	1.95	FA2G2F1	
15.78	28	1127.4142	1127.4128	1.24	FA3G2F1	
16.39	29	1098.3925	1098.3919	0.55	FA2G2S1	
16.92	30	919.8175	919.8151	2.61	M8	
		1199.9322	1199.9315	0.58	FA3G2S1	

Appendix A

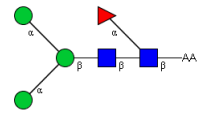
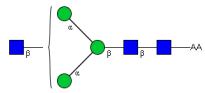
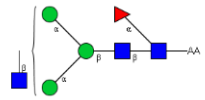
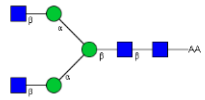
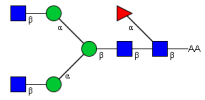
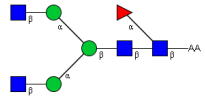
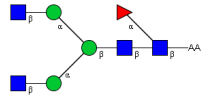
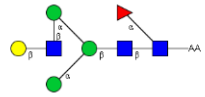
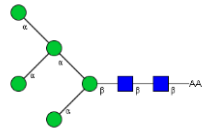
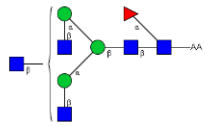
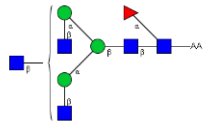
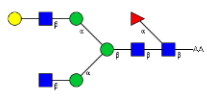
RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
		1243.9413	1243.9396	1.37	FA2G2S2	
		1098.9017	1098.9021	0.36	FA2G2F2	
		1191.9354	1191.9341	1.09	FA3G1F1S1*	
		1264.4543	1264.4528	1.19	FA2G2S2	
		1171.4218	1171.4208	0.85	FA2G2F1S1	
17.16	31	1098.9039	1098.9021	1.64	FA2G2F2	
17.40	32	1200.4431	1200.4417	1.17	FA3G2F2	
17.75	32	1243.9401	1243.9396	0.40	FA2G2S2	

Appendix A

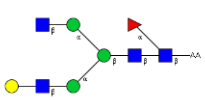
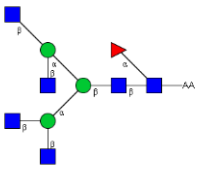
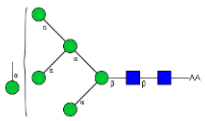
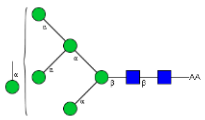
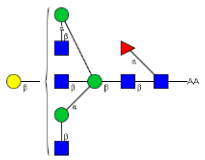
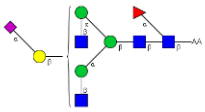
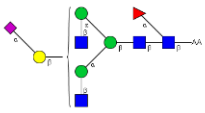
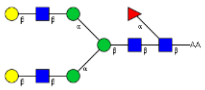
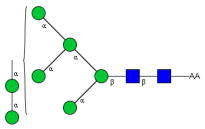
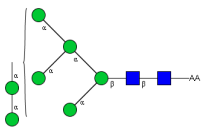
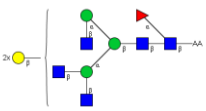
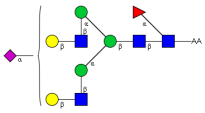
RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
17.81	33	1171.4214	1171.4208	0.51	FA2G2F1S1	
18.62	34	1244.4449	1244.4498	3.94	FA2G2F2S1	

Appendix A

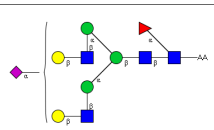
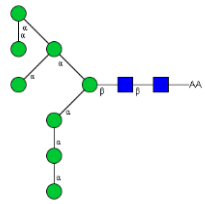
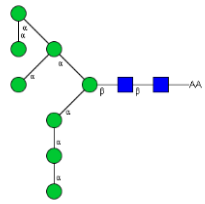
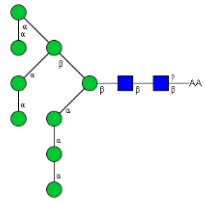
Table 1-D: Glycan assignments for each HILIC peak with corresponding m/z and proposed structures for CHO variant

RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
7.27	1	1176.4331	1176.4312	1.62	FM3	
7.62	2	1233.4508	1233.4526	1.46	A1	
		616.2235	616.2227	1.30		
8.61	3	1379.5093	1379.5105	0.87	FA1	
		689.2522	689.2516	0.87		
8.93	4	717.7632	717.7624	1.11	A2	
9.81	5	1582.5876	1582.5899	1.45	FA2	
		790.7913	790.7913	0.00		
10.15	6	790.7925	790.7913	1.52	FA2	
10.42	7	790.7919	790.7913	0.76	FA2	
10.65	8	770.2798	770.2780	2.34	FA1G1	
10.83	9	676.7362	676.7358	0.59	M5	
		1354.4781	1354.4789	0.59		
11.22	10	892.3321	892.3310	1.23	FA3	
11.44	11	892.3321	892.3310	1.12	FA3	
11.78	12	871.8188	871.8177	1.26	FA2[6]G1	
		1744.6431	1744.6427	0.23		

Appendix A

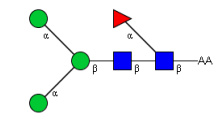
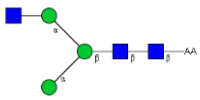
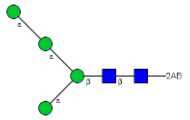
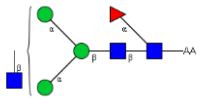
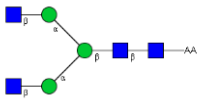
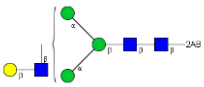
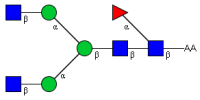
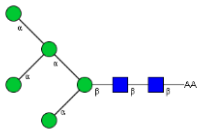
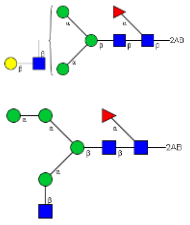
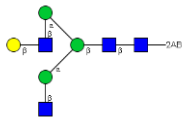
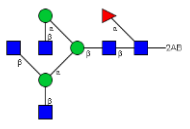
RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
12.08	13	871.8187	871.8177	1.15	FA2[3]G1	
		1744.6409	1744.6427	1.03		
12.28	14	993.8718	993.8707	1.11	FA4	
12.78	15	757.7632	757.7622	1.32	M6	
12.98	16	757.7631	757.7622	1.06	M6	
13.34	17	973.3581	973.3574	0.62	FA2bG1	
		1017.3661	1017.3654	0.69	FA2G1S1	
13.72	18	1017.3662	1017.3654	0.79	FA2G1S1	
13.85	19	952.8453	952.8441	1.26	FA2G2	
14.69	20	838.7897	838.7886	1.31	M7	
15.03	21	838.7895	838.7886	1.07	M7	
		1054.3854	1054.3838	1.52	FA3G2	
15.25	22	1098.3934	1098.3919	1.37	FA2G2S1	

Appendix A

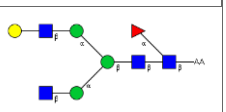
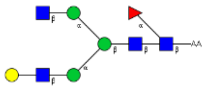
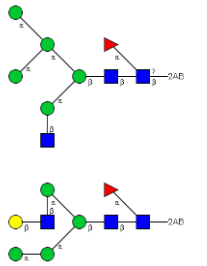
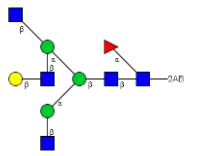
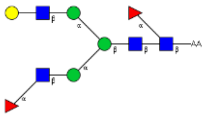
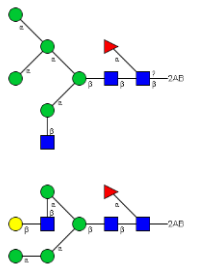
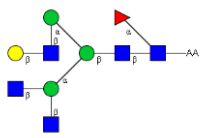
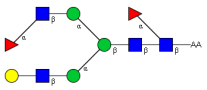
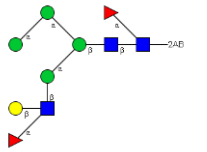
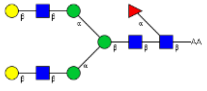
RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
16.28	23	1098.3938	1098.3919	1.73	FA2G2S1	
16.6	24	919.8167	919.8151	1.74	M8	
16.84	25	919.8169	919.8151	2.07	M8	
18.19	26	1000.8435	1000.8415	2.00	M9	

Appendix A

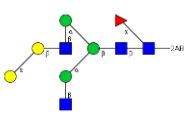
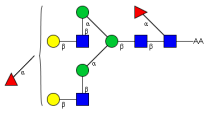
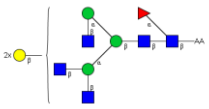
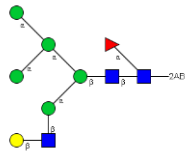
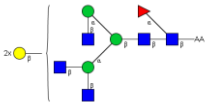
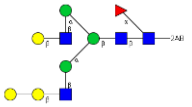
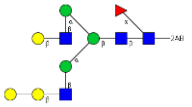
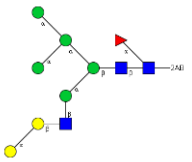
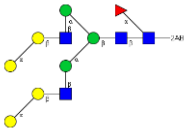
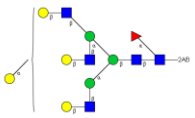
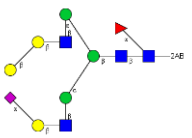
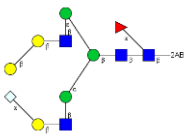
Table E: Glycan assignments for each HILIC peak with corresponding m/z and proposed structures for RDP

RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
3.8	1	1175.4612	1175.4472	11.91	FM3	
3.9	2	1232.4615	1232.4686	5.76	A1	
4.5	3	1191.4353	1191.4421	5.71	M4	
4.8	4	1378.5200 688.7577	1378.5265 688.7596	4.72 2.76	FA1	
5.0	5	717.2684 1435.5331	717.2704 1435.5480	2.79 10.38	A2	
5.6	6	696.7526 1394.5275	696.7571 1394.5214	6.46 4.37	A1G1/ M4A1	
5.9	7	790.2933 1581.5984	790.2993 1581.6059	7.59 4.74	FA2	
6.6	8	676.2357 1353.4972	676.2438 1353.4949	11.98 1.70	M5	
6.8	9	769.7750 1540.5703	769.7860 1540.5794	14.29 5.91	FA1G1/ FM4A1	
		798.2877 1597.5686	798.2968 1597.6008	11.40 20.16	A2G1	
		891.8237 1784.6915	891.8390 1784.6853	17.16 3.47	FA3	
7.8	10	871.3199	871.3257	6.66	FA2[6]G1	

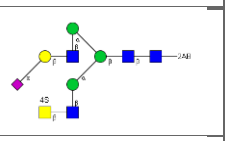
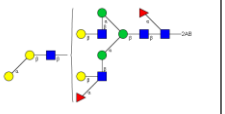
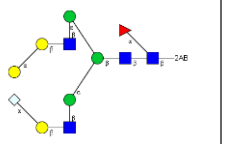
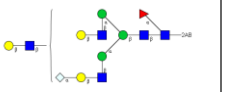
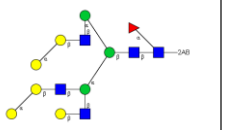
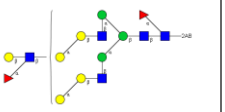
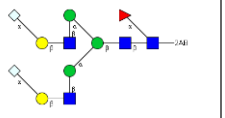
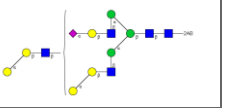
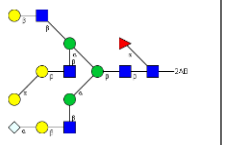
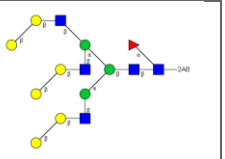
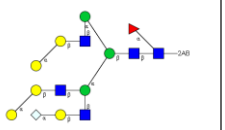
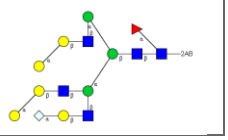
Appendix A

RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
		1743.6433	1743.6587	8.83		
8.2	11	871.3199	871.3257	6.66	FA2[3]G1	
		1743.6348	1743.6587	13.71		
9.0	12	850.7973	850.8125	17.87	FM5A1/ FM4A1G1	
		1702.6240	1702.6322	4.82		
		972.8490	972.8654	16.86	FA3G1	
		1946.7131	1946.7381	12.84		
		944.3312	944.3547	24.88	FA2G1F1	
9.3	13	850.8093	850.8125	3.76	FM5A1/ FM4A1G1	
		1702.5902	1702.6322	24.67		
		972.8490	972.8654	16.86	FA3G1	
		1946.7131	1946.7381	12.84		
		944.3312	944.3547	24.88	FA2G1F1	
10.0	14	923.8398	923.8414	1.73	FM4A1G1F1	
10.4	15	952.3463	952.3521	6.09	FA2G2	
		1905.6665	1905.7116	23.67		

Appendix A

RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
10.7	16	952.3463	952.3521	6.09	FA2G1Gal1	
		1905.7113	1905.7116	0.16		
11.4	17	1025.3698	1025.3811	11.02	FA2G2F1	
		1053.8849	1053.8918	6.55	FA3G2	
11.6	18	931.8303	931.8389	9.23	FM5A1G1	
		1864.6675	1864.6850	9.38		
11.9	19	1053.8783	1053.8918	12.81	FA3G2	
12.9	20	1033.3781	1033.3785	0.39	FA2G2Gal1	
13.9	21	1106.4052	1106.4075	2.08	FA2G2F1Gal1	
14.3	22	1012.8443	1012.8653	20.73	FM5A1G1Gal1	
15.6	23	1114.3948	1114.4050	9.15	FA2G2Gal2	
16.6	24	1215.9491	1215.9446	3.70	FA3G3Gal1	
		1178.9164	1178.9263	8.40	FA2G2Ga1S1	
16.9	25	1186.9087	1186.9237	12.64	FA2G2Gal1Sg1	

Appendix A

RT	Peak #	Experimental mass (m/z)	Theoretical mass (m/z)	Error (ppm)	Assn.	Proposed Structure
		1085.3777	1085.3626	13.91	A2G1S1GalNAc1	
17.4	26	1288.9551	1288.9736	14.35	FA3G3Ga1F1	
18.0	27	1186.9087	1186.9237	12.64	FA2G2Gal1Sg1	
18.8	28	1288.4470	1288.4637	12.96	FA3G3Sg1	
19.0	29	1296.9569	1296.9711	10.95	FA3G3Gal2	
19.8	30	1369.9891	1370.0000	7.96	FA3G3Gal2F1	
20.6	31	1259.4216	1259.4425	16.59	FA2G2Sg2	
21.0	32	1369.5033	1369.4898	9.86	A3G3Gal2S1	
		1369.5033	1369.4898	9.86	FA3G3Gal1Sg1	
21.4	33	1377.9794	1377.9975	13.14	FA3G3Gal3	
22.3	34	1450.5078	1450.5162	5.79	FA3G3Gal2Sg1	
23.2	35	1450.5000	1450.5162	11.17	FA3G3Gal2Sg1	

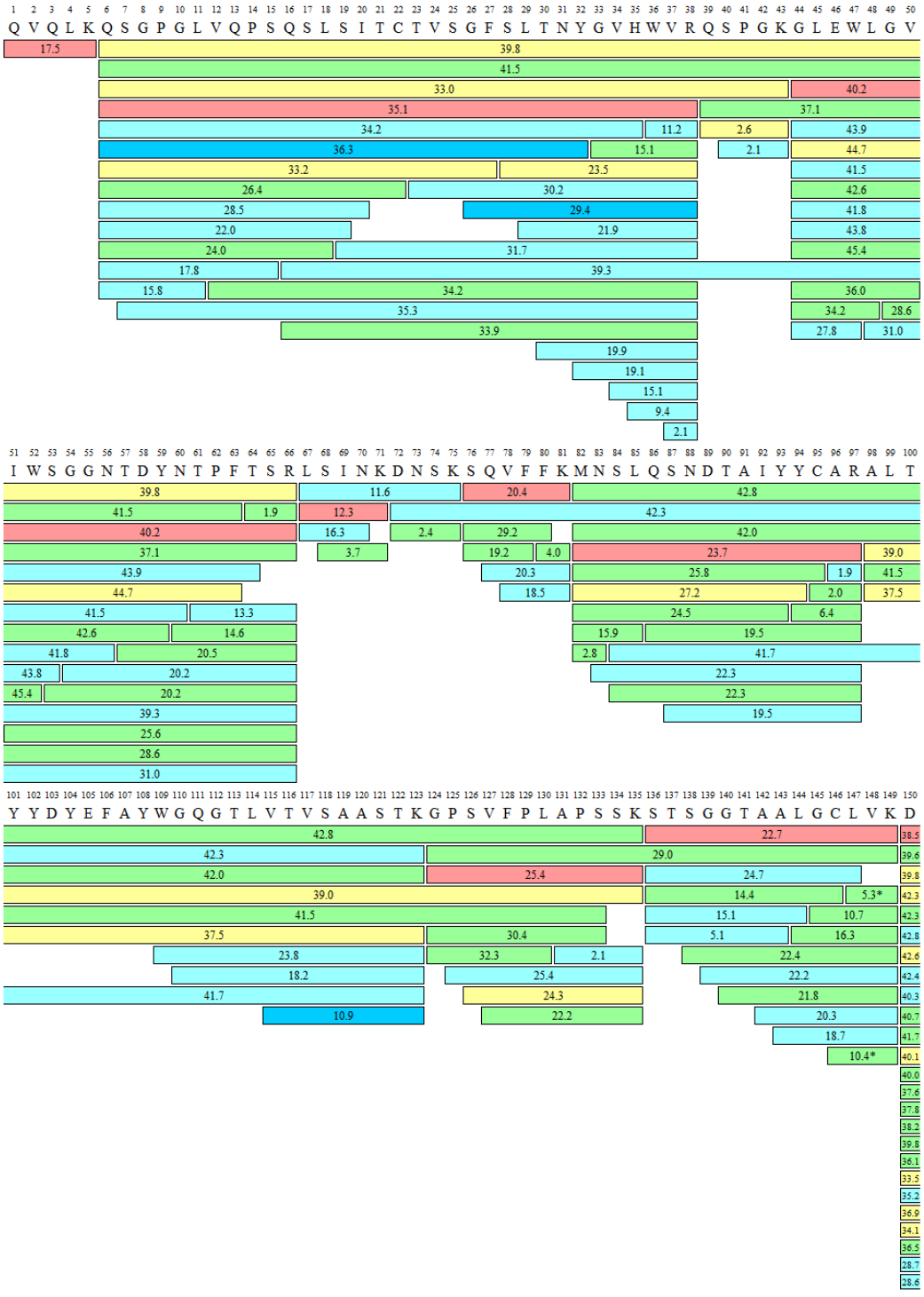
Appendix B

Figure 1-D: Peptide map coverage of the HEK variant

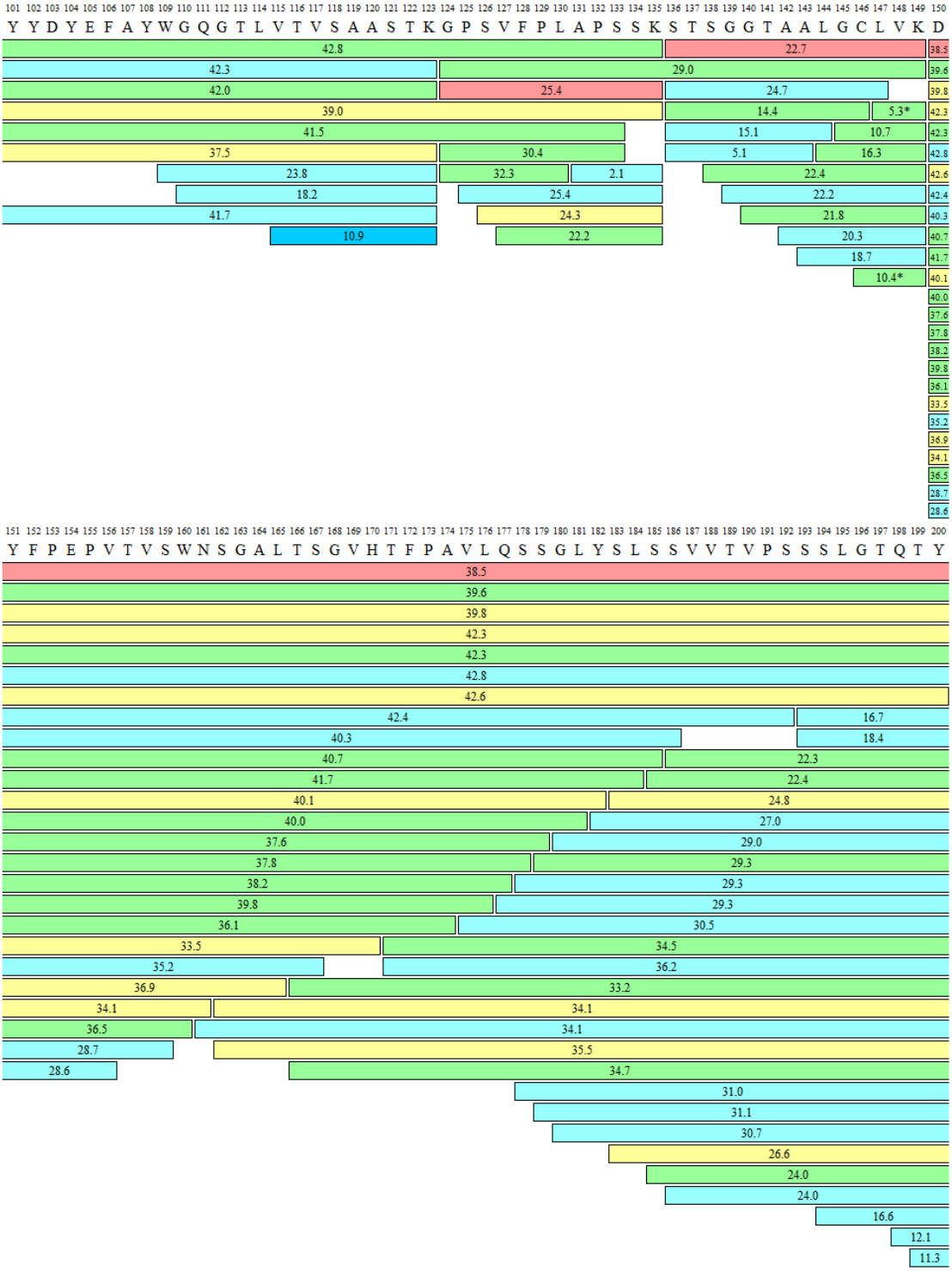
Proteins	Number of MS Peaks	MS Peak Area	Sequence Coverage	Abundance (mol)
1:Heavy Chain	1681	31.8%	100.0%	54.29%
2:Light Chain	759	15.9%	100.0%	45.71%
Unidentified	5006	52.3%		

Minimum Peptide Mass = 7000.0
Minimum Confidence Score = 0.80
Minimum Peak Area = 1.2e+04
Relative Peak Area Threshold = 0%

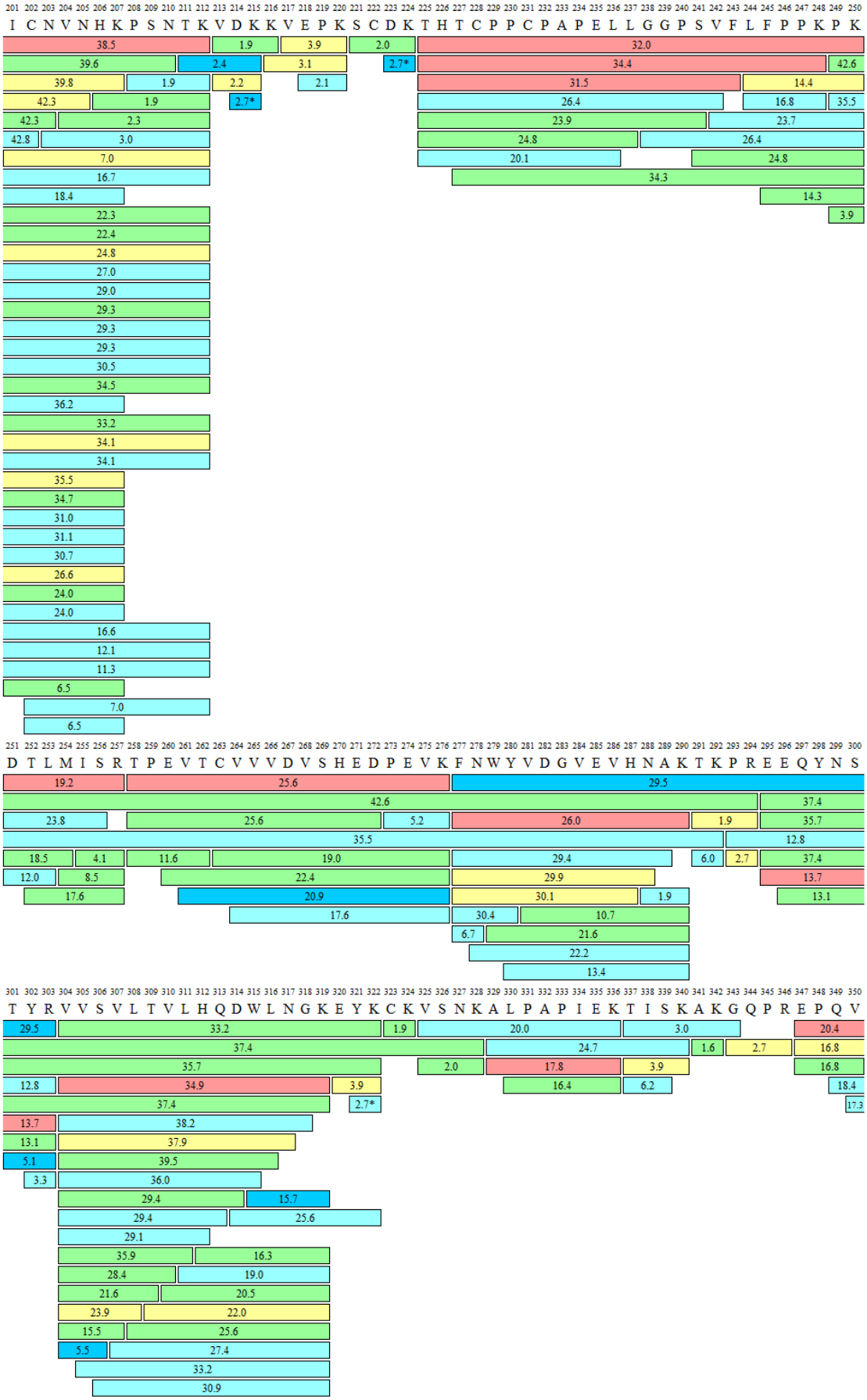
Heavy Chain



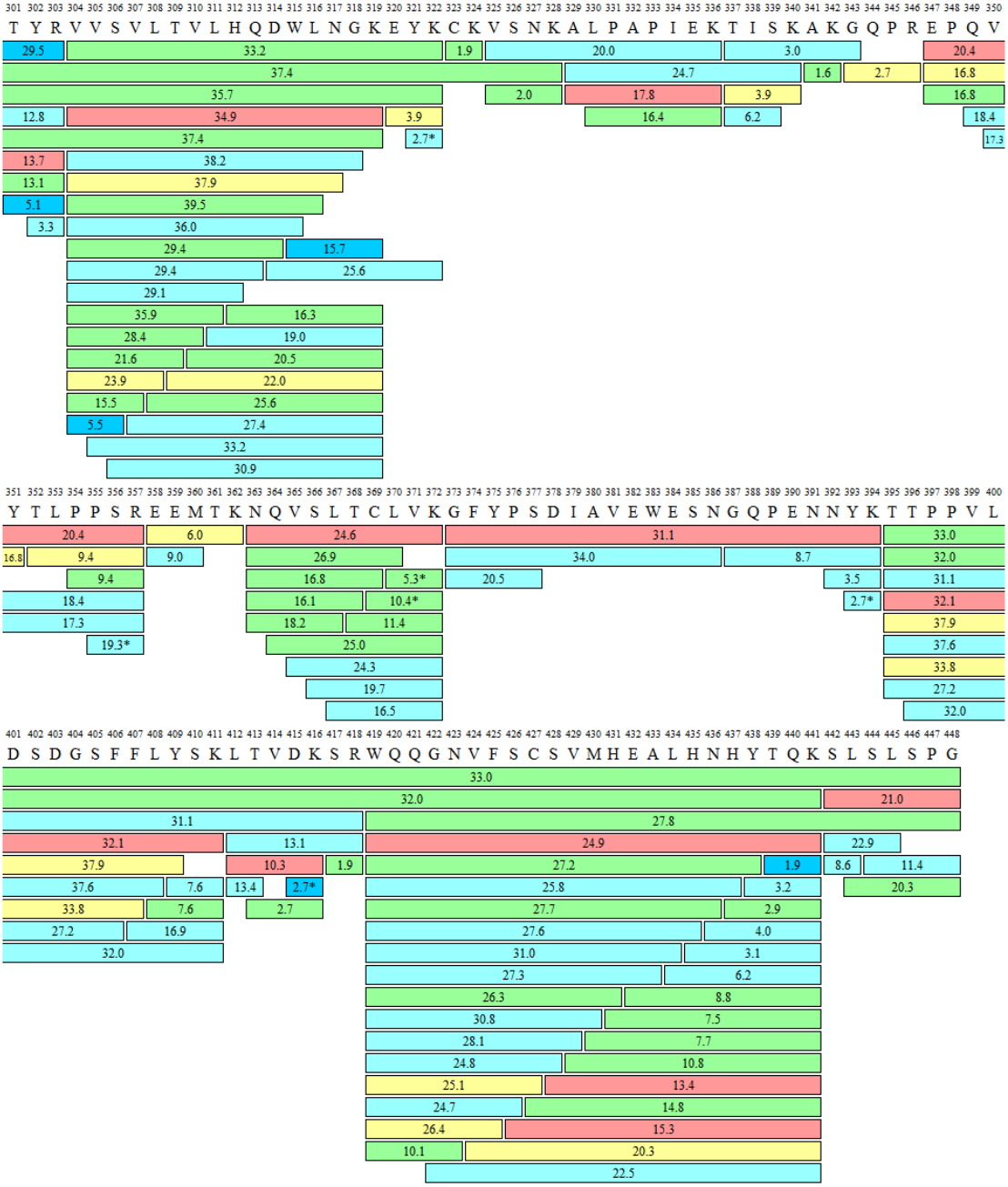
Appendix B


101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
Y Y D Y E F A Y W G Q G T L V T V S A A S T K G P S V F P L A P S S K S T S G G T A A L G C L V K D
42.8 22.7 38.5
42.3 29.0 39.6
42.0 25.4 24.7 39.8
39.0 14.4 5.3* 42.3
41.5 15.1 10.7 42.3
37.5 30.4 5.1 16.3 42.8
23.8 32.3 2.1 22.4 42.6
18.2 25.4 22.2 42.4
41.7 24.3 21.8 40.3
10.9 22.2 20.3 40.7
18.7 41.7
10.4* 40.1
40.0
37.6
37.8
38.2
39.8
36.1
33.5
35.2
36.9
34.1
36.5
28.7
28.6
151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
Y F P E P V T V S W N S G A L T S G V H T F P A V L Q S S G L Y S L S S V V T V P S S S L G T Q T Y
38.5
39.6
39.8
42.3
42.3
42.8
42.6
42.4 16.7
40.3 18.4
40.7 22.3
41.7 22.4
40.1 24.8
40.0 27.0
37.6 29.0
37.8 29.3
38.2 29.3
39.8 29.3
36.1 30.5
33.5 34.5
35.2 36.2
36.9 33.2
34.1 34.1
36.5 34.1
28.7 35.5
28.6 34.7
31.0
31.1
30.7
26.6
24.0
24.0
16.6
12.1
11.3

Appendix B

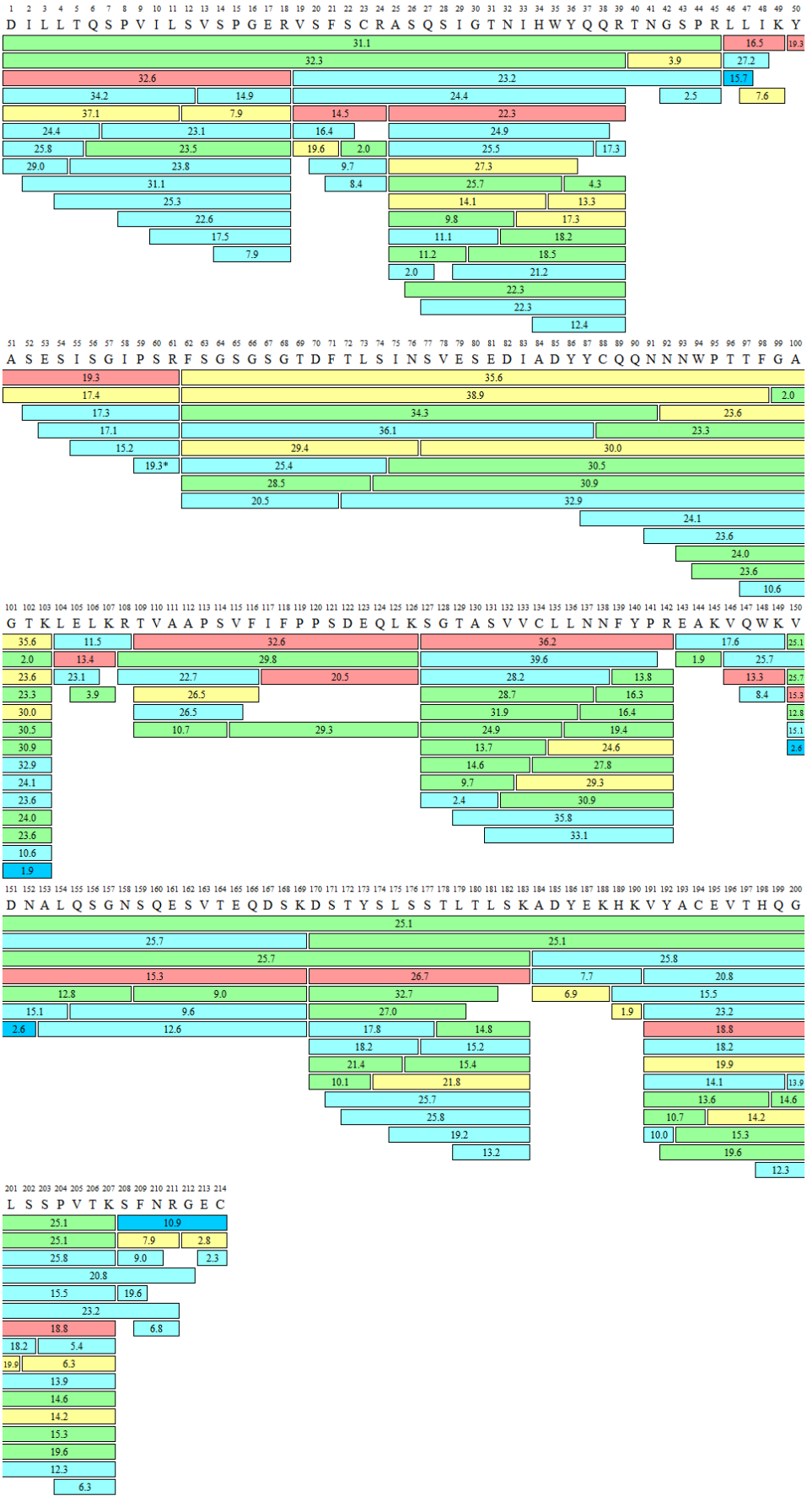

201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250
I C N V N H K P S N T K V D K K V E P K S C D K T H T C P P C P A P E L L G G P S V F L F P P K P K
38.5 1.9 3.9 2.0 32.0
39.6 2.4 3.1 2.7* 34.4 42.6
39.8 1.9 2.2 2.1 31.5 14.4
42.3 1.9 2.7* 26.4 16.8 35.5
42.3 2.3 23.9 23.7
42.8 3.0 24.8 26.4
7.0 20.1 24.8
16.7 34.3
18.4 14.3
22.3 3.9
22.4
24.8
27.0
29.0
29.3
29.3
29.3
30.5
34.5
36.2
33.2
34.1
34.1
35.5
34.7
31.0
31.1
30.7
26.6
24.0
24.0
16.6
12.1
11.3
6.5
7.0
6.5
251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
D T L M I S R T P E V T C V V V D V S H E D P E V K F N W Y V D G V E V H N A K T K P R E E Q Y N S
19.2 25.6 29.5
42.6 37.4
23.8 25.6 5.2 26.0 1.9 35.7
35.5 12.8
18.5 4.1 11.6 19.0 29.4 6.0 2.7 37.4
12.0 8.5 22.4 29.9 13.7
17.6 20.9 30.1 1.9 13.1
17.6 30.4 10.7
6.7 21.6
22.2
13.4
301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350
T Y R V V S V L T V L H Q D W L N G K E Y K C K V S N K A L P A P I E K T I S K A K G Q P R E P Q V
29.5 33.2 1.9 20.0 3.0 20.4
37.4 24.7 1.6 2.7 16.8
35.7 2.0 17.8 3.9 16.8
12.8 34.9 3.9 16.4 6.2 18.4
37.4 2.7* 17.3
13.7 38.2
13.1 37.9
5.1 39.5
3.3 36.0
29.4 15.7
29.4 25.6
29.1
35.9 16.3
28.4 19.0
21.6 20.5
23.9 22.0
15.5 25.6
5.5 27.4
33.2
30.9

Appendix B


301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350
T Y R V V S V L T V L H Q D W L N G K E Y K C K V S N K A L P A P I E K T I S K A K G Q P R E P Q V
29.5 33.2 1.9 20.0 3.0 20.4
37.4 24.7 1.6 2.7 16.8
35.7 2.0 17.8 3.9 16.8
12.8 34.9 3.9 16.4 6.2 18.4
37.4 2.7* 17.3
13.7 38.2
13.1 37.9
5.1 39.5
3.3 36.0
29.4 15.7
29.4 25.6
29.1
35.9 16.3
28.4 19.0
21.6 20.5
23.9 22.0
15.5 25.6
5.5 27.4
33.2
30.9
351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
Y T L P P S R E E M T K N Q V S L T C L V K G F Y P S D I A V E W E S N G Q P E N N Y K T T P P V L
20.4 6.0 24.6 31.1 33.0
16.8 9.4 9.0 26.9 34.0 8.7 32.0
9.4 16.8 5.3* 20.5 3.5 31.1
18.4 16.1 10.4* 2.7* 32.1
17.3 18.2 11.4 37.9
19.3* 25.0 37.6
24.3 33.8
19.7 27.2
16.5 32.0
401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448
D S D G S F F L Y S K L T V D K S R W Q Q G N V F S C S V M H E A L H N H Y T Q K S L S L S P G
33.0
32.0 21.0
31.1 27.8
32.1 13.1 24.9 22.9
37.9 10.3 1.9 27.2 1.9 8.6 11.4
37.6 7.6 13.4 2.7* 25.8 3.2 20.3
33.8 7.6 2.7 27.7 2.9
27.2 16.9 27.6 4.0
32.0 31.0 3.1
27.3 6.2
26.3 8.8
30.8 7.5
28.1 7.7
24.8 10.8
25.1 13.4
24.7 14.8
26.4 15.3
10.1 20.3
22.5

Appendix B

Light Chain



* indicates peptide not unique

Color code for signal intensity

>1.1e+08	>1.2e+07	>1.4e+06	>1.6e+05	>1.8e+04

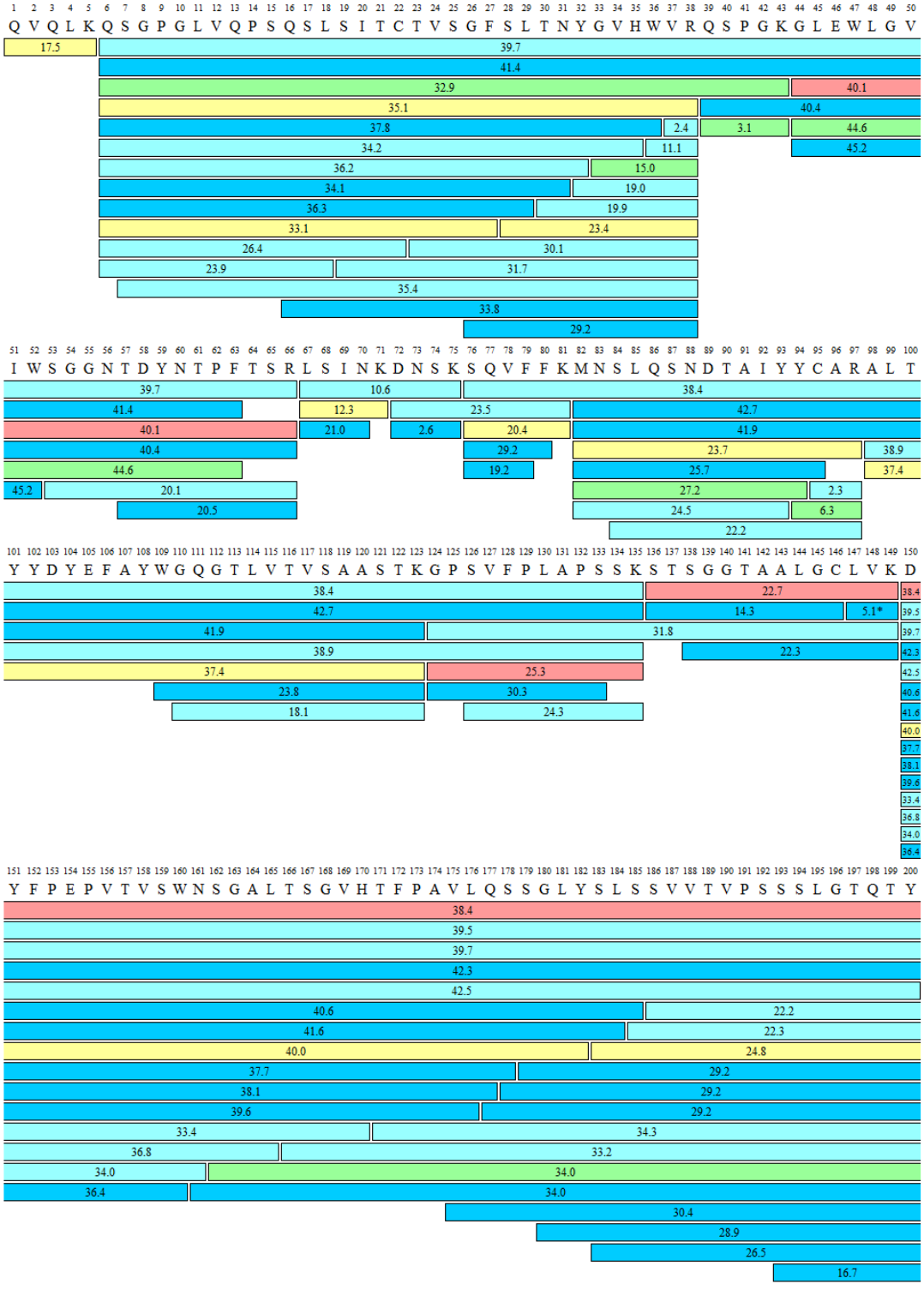
Appendix B

Figure 1-E: Peptide map coverage of the CHO variant

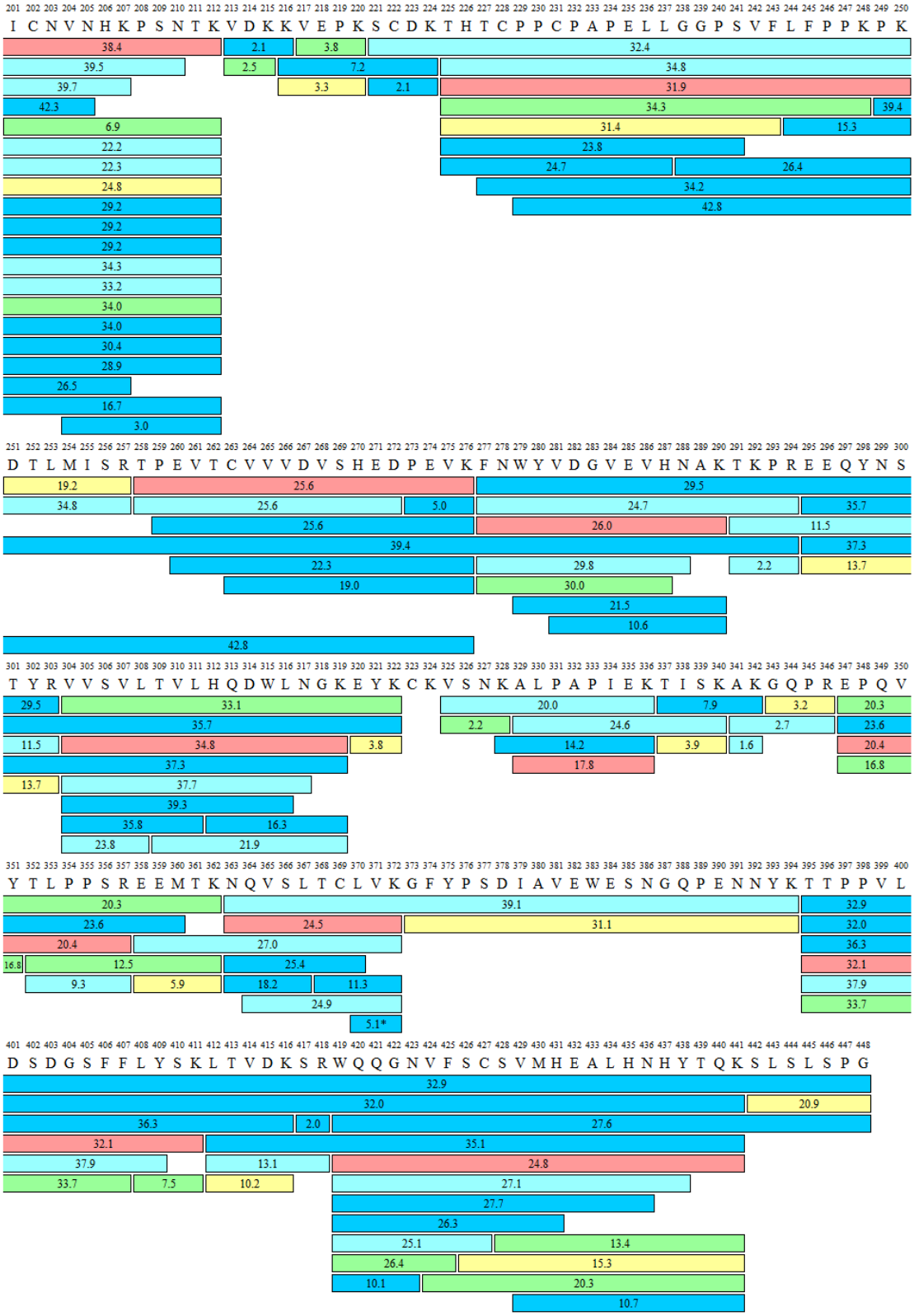
Proteins	Number of MS Peaks	MS Peak Area	Sequence Coverage	Abundance (mol)
1:Heavy Chain	1238	40.6%	99.6%	44.82%
2:Light Chain	586	25.8%	100.0%	55.18%
Unidentified	3821	33.6%		

Minimum Peptide Mass = 7000.0
Minimum Confidence Score = 0.80
Minimum Peak Area = 1.6e+05
Relative Peak Area Threshold = 0%

Heavy Chain

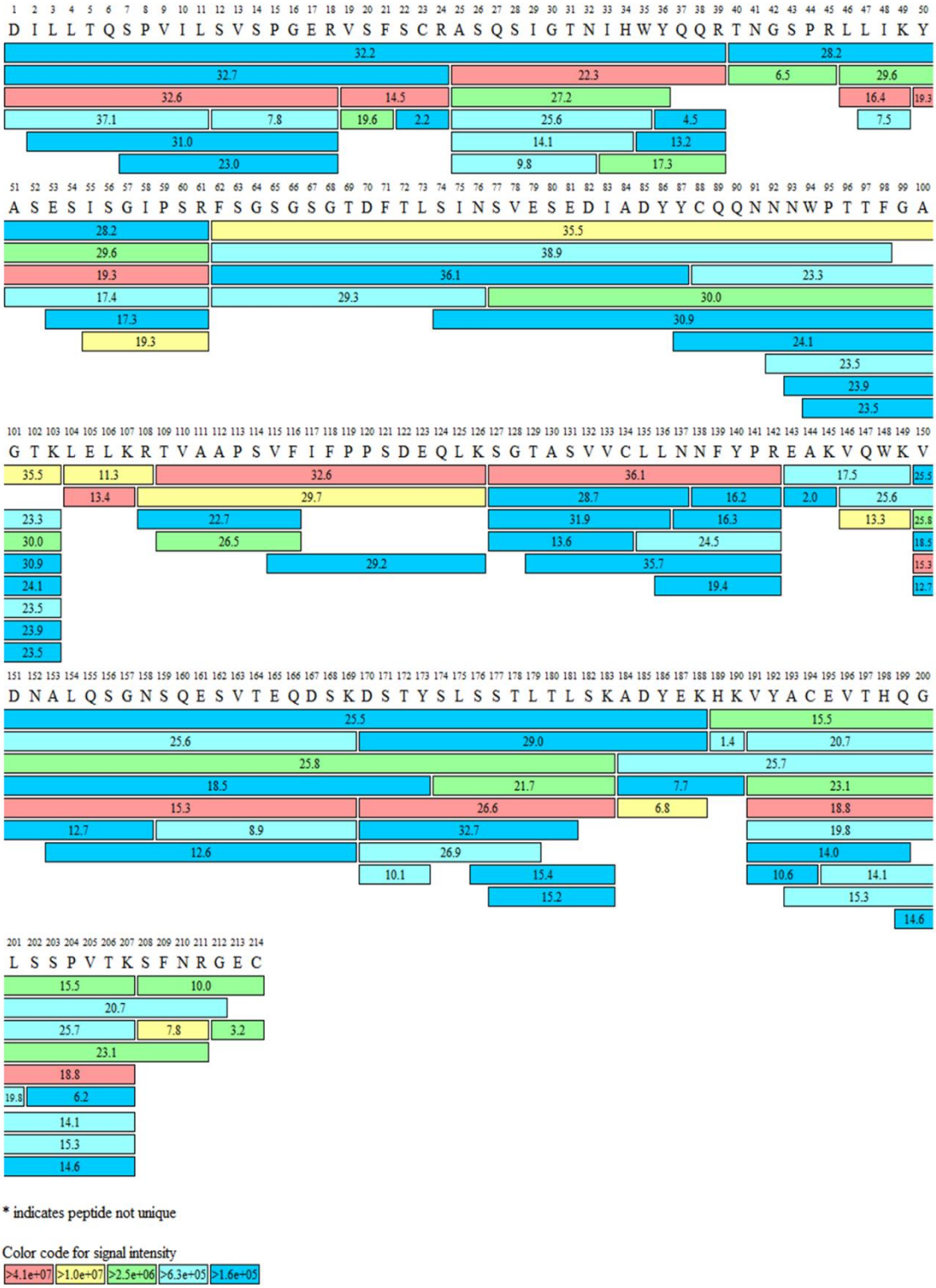


Appendix B


201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250
I C N V N H K P S N T K V D K K V E P K S C D K T H T C P P C P A P E L L G G P S V F L F P P K P K
38.4 2.1 3.8 32.4
39.5 2.5 7.2 34.8
39.7 3.3 2.1 31.9
42.3 34.3 39.4
6.9 31.4 15.3
22.2 23.8
22.3 24.7 26.4
24.8 34.2
29.2 42.8
29.2
29.2
34.3
33.2
34.0
34.0
30.4
28.9
26.5
16.7
3.0
251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
D T L M I S R T P E V T C V V V D V S H E D P E V K F N W Y V D G V E V H N A K T K P R E E Q Y N S
19.2 25.6 29.5
34.8 25.6 5.0 24.7 35.7
25.6 26.0 11.5
39.4 37.3
22.3 29.8 2.2 13.7
19.0 30.0
21.5
10.6
42.8
301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350
T Y R V V S V L T V L H Q D W L N G K E Y K C K V S N K A L P A P I E K T I S K A K G Q P R E P Q V
29.5 33.1 20.0 7.9 3.2 20.3
35.7 2.2 24.6 2.7 23.6
11.5 34.8 3.8 14.2 3.9 1.6 20.4
37.3 17.8 16.8
13.7 37.7
39.3
35.8 16.3
23.8 21.9
351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
Y T L P P S R E E M T K N Q V S L T C L V K G F Y P S D I A V E W E S N G Q P E N N Y K T T P P V L
20.3 39.1 32.9
23.6 24.5 31.1 32.0
20.4 27.0 36.3
16.8 12.5 25.4 32.1
9.3 5.9 18.2 11.3 37.9
24.9 33.7
5.1*
401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448
D S D G S F F L Y S K L T V D K S R W Q Q G N V F S C S V M H E A L H N H Y T Q K S L S L S P G
32.9
32.0 20.9
36.3 2.0 27.6
32.1 35.1
37.9 13.1 24.8
33.7 7.5 10.2 27.1
27.7
26.3
25.1 13.4
26.4 15.3
10.1 20.3
10.7

Appendix B

Light Chain



* indicates peptide not unique

Color code for signal intensity

>4.1e+07	>1.0e+07	>2.5e+06	>6.3e+05	>1.6e+05

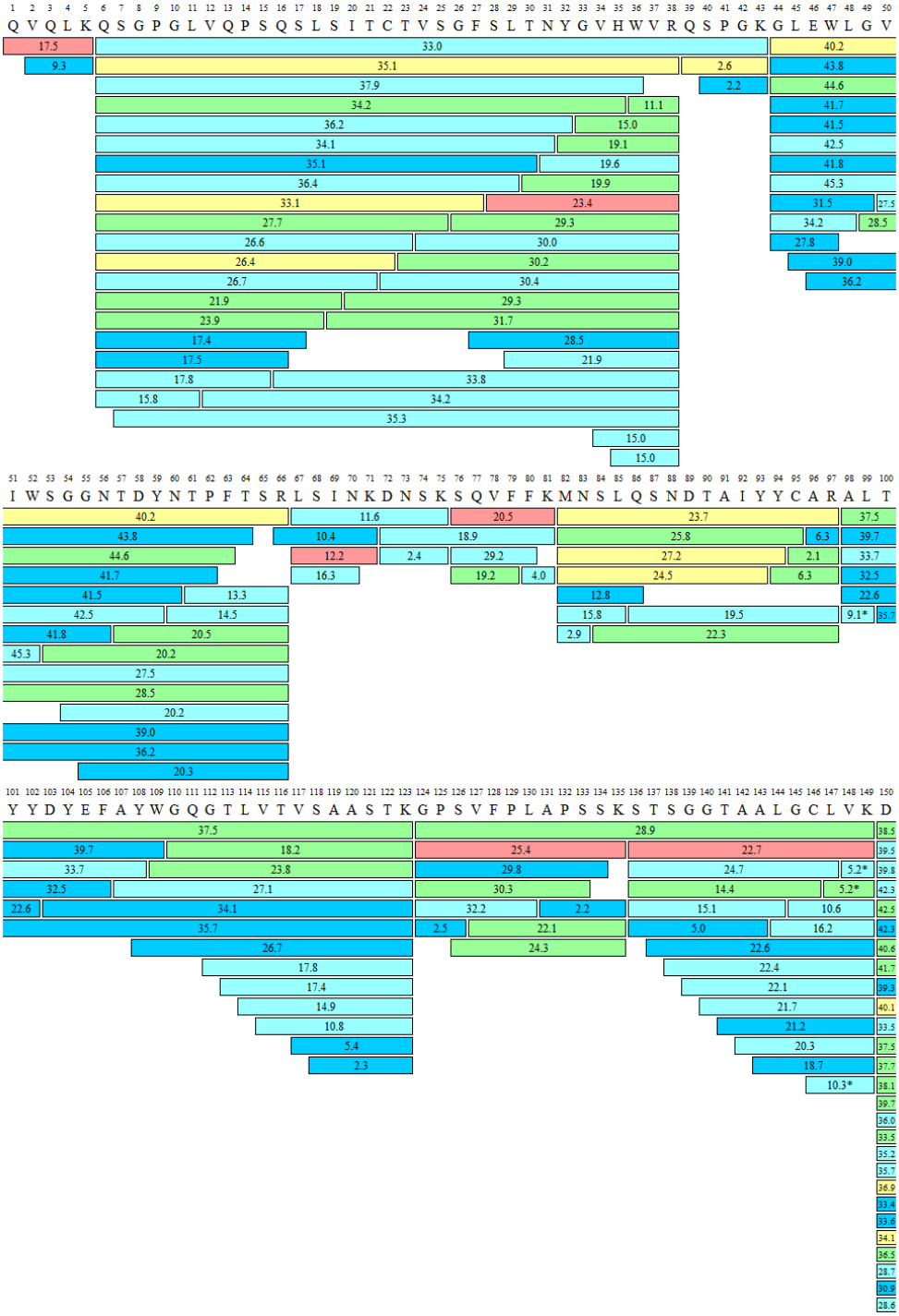
Figure 1-F: Peptide map coverage of the RDP

Appendix B

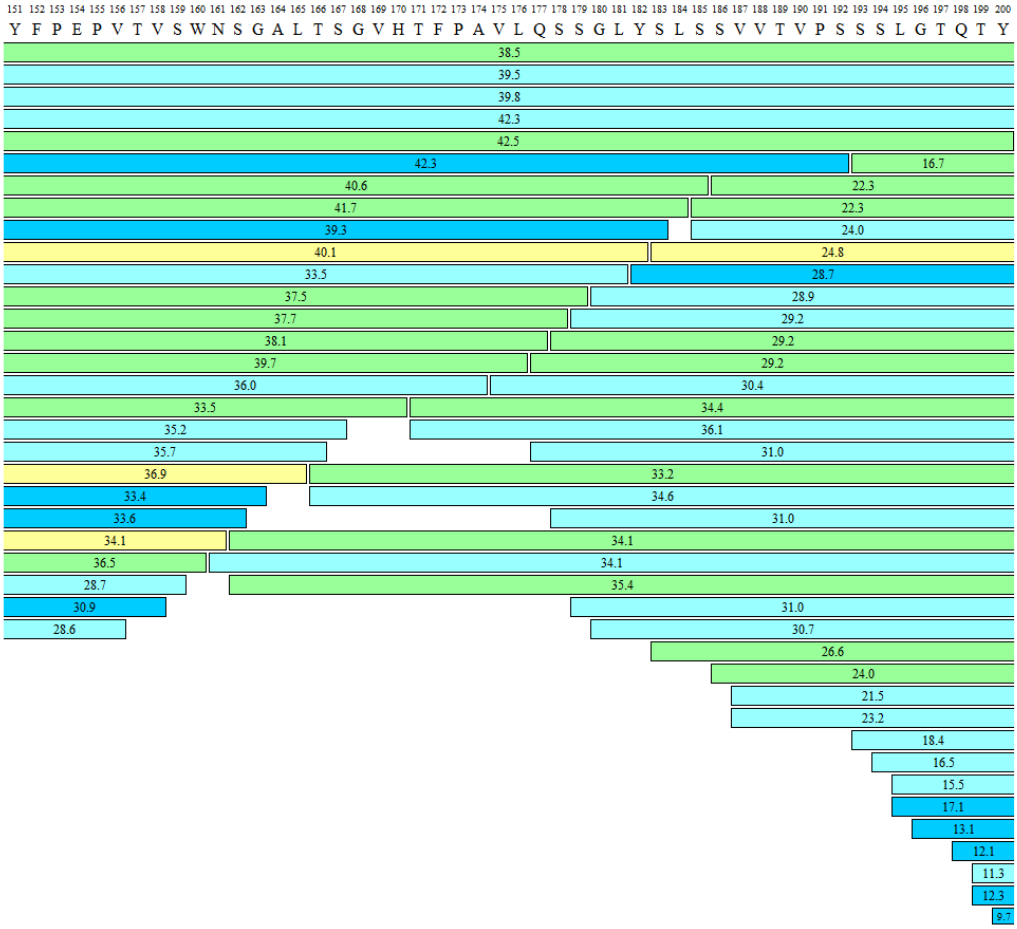
Proteins	Number of MS Peaks	MS Peak Area	Sequence Coverage	Abundance (mol)
1:Heavy	2145	54.3%	100.0%	51.67%
2:Light	855	27.9%	100.0%	48.33%
Unidentified	14826	17.8%		

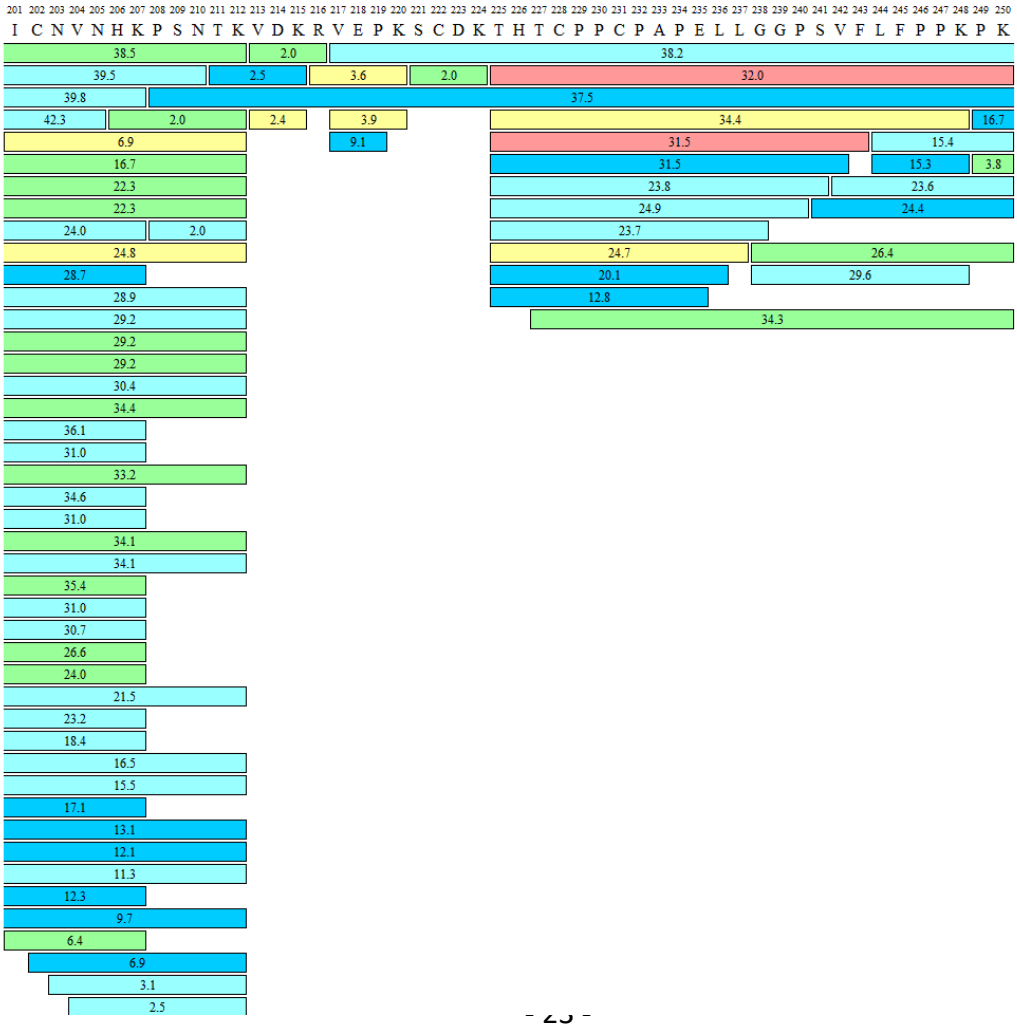
Minimum Peptide Mass = 7000.0
Minimum Confidence Score = 0.80
Minimum Peak Area = 1.4e+04
Relative Peak Area Threshold = 0%

Heavy

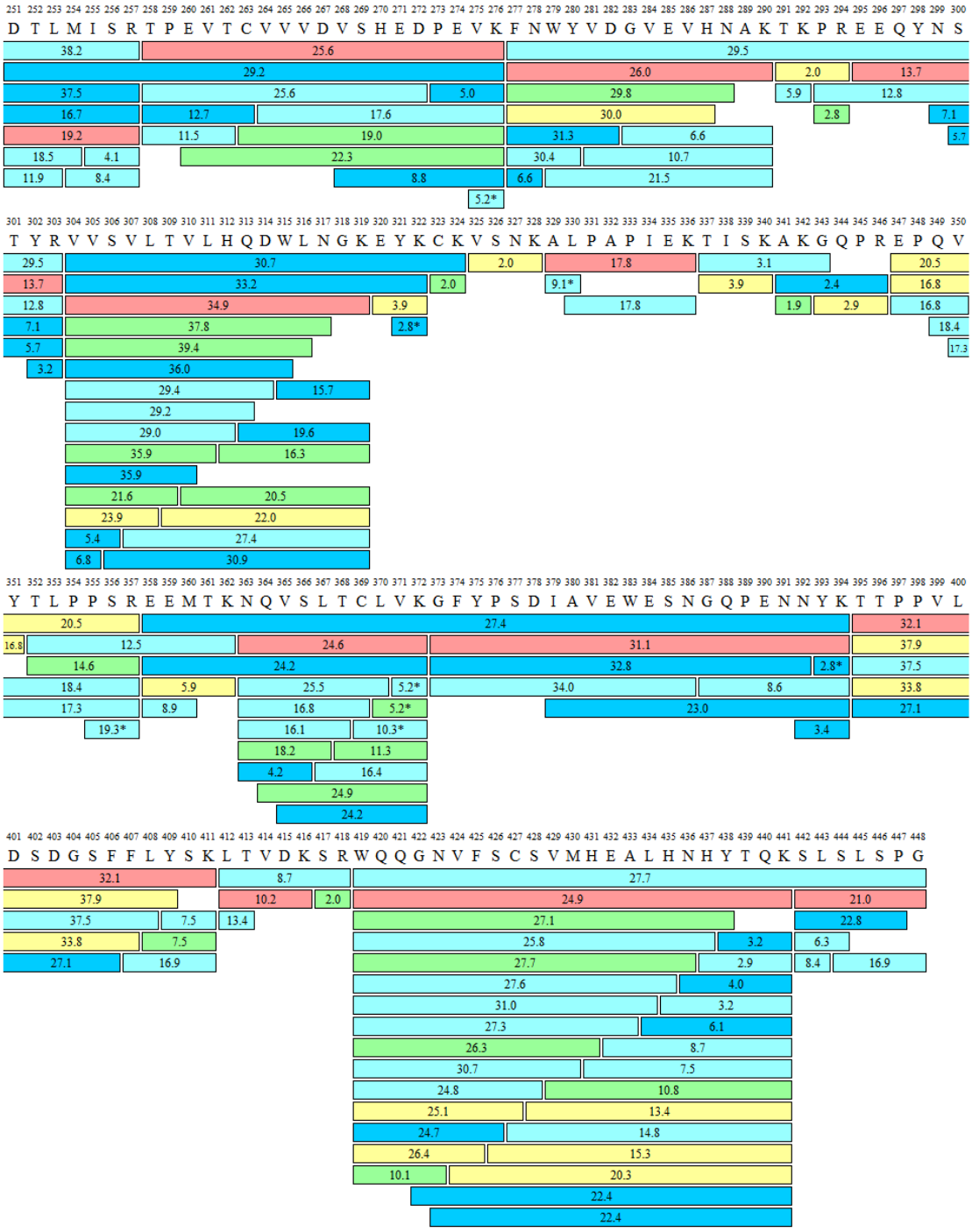


Appendix B


151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
Y F P E P V T V S W N S G A L T S G V H T F P A V L Q S S G L Y S L S S V V T V P S S S L G T Q T Y
38.5
39.5
39.8
42.3
42.5
42.3 16.7
40.6 22.3
41.7 22.3
39.3 24.0
40.1 24.8
33.5 28.7
37.5 28.9
37.7 29.2
38.1 29.2
39.7 29.2
36.0 30.4
33.5 34.4
35.2 36.1
35.7 31.0
36.9 33.2
33.4 34.6
33.6 31.0
34.1 34.1
36.5 34.1
28.7 35.4
30.9 31.0
28.6 30.7
26.6
24.0
21.5
23.2
18.4
16.5
15.5
17.1
13.1
12.1
11.3
12.3
9.7

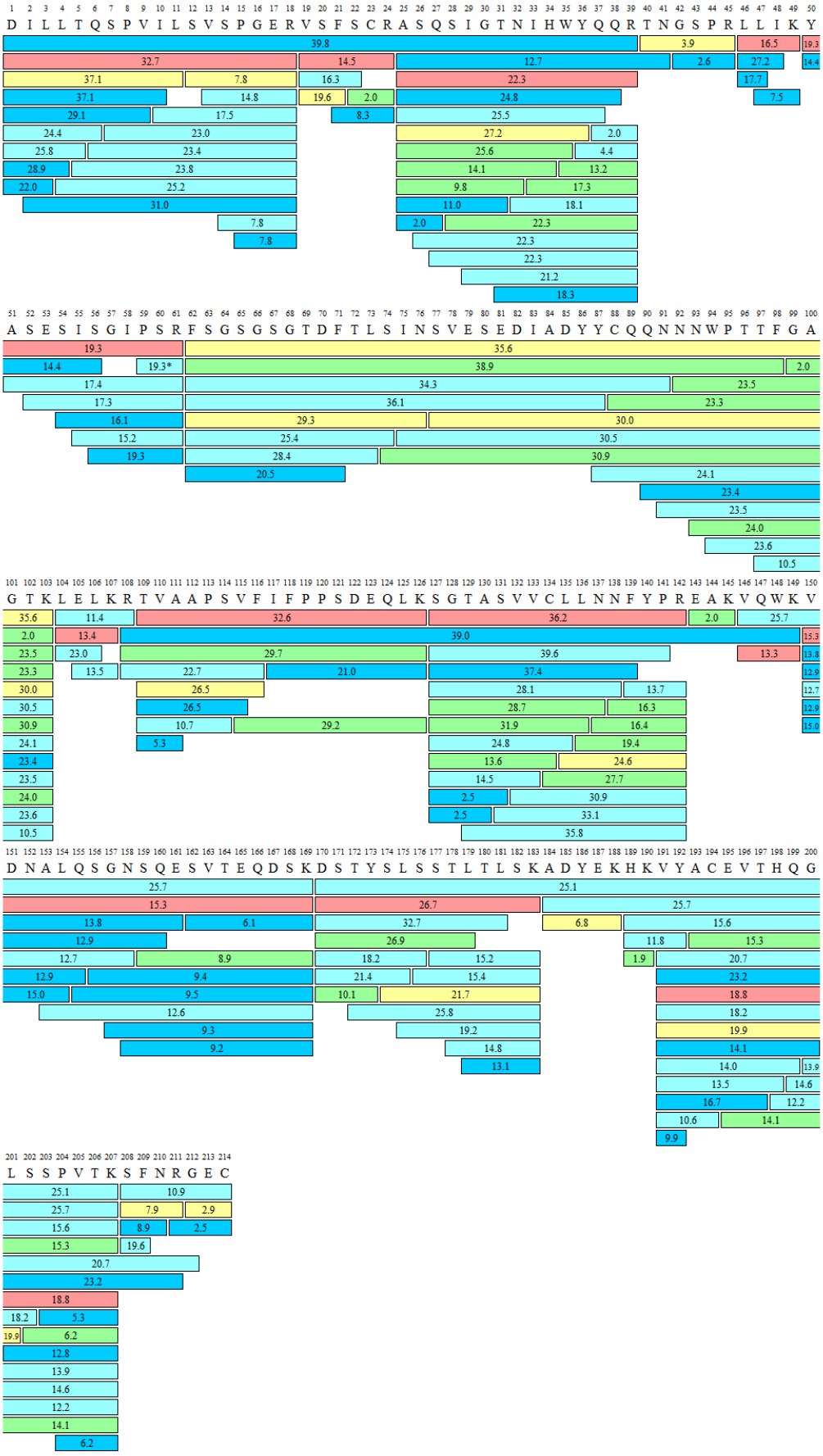

201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250
I C N V N H K P S N T K V D K R V E P K S C D K T H T C P P C P A P E L L G G P S V F L F P P K P K
38.5 2.0 38.2
39.5 2.5 3.6 2.0 32.0
39.8 37.5
42.3 2.0 2.4 3.9 34.4 16.7
6.9 9.1 31.5 15.4
16.7 31.5 15.3 3.8
22.3 23.8 23.6
22.3 24.9 24.4
24.0 2.0 23.7
24.8 24.7 26.4
28.7 20.1 29.6
28.9 12.8
29.2 34.3
29.2
29.2
30.4
34.4
36.1
31.0
33.2
34.6
31.0
34.1
34.1
35.4
31.0
30.7
26.6
24.0
21.5
23.2
18.4
16.5
15.5
17.1
13.1
12.1
11.3
12.3
9.7
6.4
6.9
3.1
2.5

Appendix B


251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
D T L M I S R T P E V T C V V V D V S H E D P E V K F N W Y V D G V E V H N A K T K P R E E Q Y N S
38.2 25.6 29.5
29.2 26.0 2.0 13.7
37.5 25.6 5.0 29.8 5.9 12.8
16.7 12.7 17.6 30.0 2.8 7.1
19.2 11.5 19.0 31.3 6.6 5.7
18.5 4.1 22.3 30.4 10.7
11.9 8.4 8.8 6.6 21.5
5.2*
301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350
T Y R V V S V L T V L H Q D W L N G K E Y K C K V S N K A L P A P I E K T I S K A K G Q P R E P Q V
29.5 30.7 2.0 17.8 3.1 20.5
13.7 33.2 2.0 9.1* 3.9 2.4 16.8
12.8 34.9 3.9 17.8 1.9 2.9 16.8
7.1 37.8 2.8* 18.4
5.7 39.4 17.3
3.2 36.0
29.4 15.7
29.2
29.0 19.6
35.9 16.3
35.9
21.6 20.5
23.9 22.0
5.4 27.4
6.8 30.9
351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
Y T L P P S R E E M T K N Q V S L T C L V K G F Y P S D I A V E W E S N G Q P E N N Y K T T P P V L
20.5 27.4 32.1
16.8 12.5 24.6 31.1 37.9
14.6 24.2 32.8 2.8* 37.5
18.4 5.9 25.5 5.2* 34.0 8.6 33.8
17.3 8.9 16.8 5.2* 23.0 27.1
19.3* 16.1 10.3* 3.4
18.2 11.3
4.2 16.4
24.9
24.2
401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448
D S D G S F F L Y S K L T V D K S R W Q Q G N V F S C S V M H E A L H N H Y T Q K S L S L S P G
32.1 8.7 27.7
37.9 10.2 2.0 24.9 21.0
37.5 7.5 13.4 27.1 22.8
33.8 7.5 25.8 3.2 6.3
27.1 16.9 27.7 2.9 8.4 16.9
27.6 4.0
31.0 3.2
27.3 6.1
26.3 8.7
30.7 7.5
24.8 10.8
25.1 13.4
24.7 14.8
26.4 15.3
10.1 20.3
22.4
22.4

Appendix B

Light



* indicates peptide not unique

Color code for signal intensity

>2.5e+07	>3.8e+06	>5.9e+05	>9.1e+04	>1.4e+04