Mass spectrometry platforms for higher order structural measurements of protein biotherapeutics

Reference procedures for establishing the reproducibility and sensitivity of measurements

Kate Joanna Groves

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Molecular and Cellular Biology

January 2020

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Author Contributions

The description of the HDX-MS statistical model described in Chapter 2 (section 2.3.1.5) plus HDX-MS and IMS-MS measurements of NISTmAb presented in Chapter 3 are largely based on the following jointly authored publication;

Groves. K.; Cryar. A.; Cowen, S.; Ashcroft, A.E.; Quaglia. M.; Mass Spectrometry Characterization of Higher Order Structural Changes Associated with the Fc-glycan Structure of the NISTmAb Reference Material, RM 8761, *Journal of the American Society for Mass Spectrometry*, Accepted Dec 2019,

dx.doi.org/10.1021/jasms.9b00022

All experimental work and measurement in the publication were performed by the candidate. Dr. S. Cowen assisted with the notations described in Equations 32-34, defined in section 2.3.1.5 for the HDX-MS statistical model. Dr A. Cryar at times provided technical support. Prof. A.E. Ashcroft and Dr. M. Quaglia are the candidate's academic and industrial supervisors respectively.

In Chapter 2, repeat HDX-MS measurements of the rhGH control, using the method developed by the candidate, were performed by Dr. A. Cryar and Dr. D. Cooper-Shepherd. These were used to understand the robustness and reproducibility of HDX-MS measurements performed across a yearlong period (section 2.3.1.4), facilitating the contribution to measurement repeatability as a result of analyst.

Acknowledgements

I would like to thank both my supervisors Prof. Alison Ashcroft and Dr. Milena Quaglia for their guidance and support during my PhD. Milena in particular, thank you for your day-by-day support, discussion, laughter and reality checks! I would also like to thank my colleagues, past and present, from both LGC and University of Leeds, for their training, assistance, patience and contributions towards this and other work, it has been a pleasure learning so much from all of you.

I would also like to thank my family. To my parents, Fran and David, who have always been loving, supportive and who have raised us with the confidence to follow any path we should wish. To my siblings, Tom, Freia and Helen (honorary), thank you for always supporting me with unrelenting love, encouragement and without judgment. To Uncle Rich, for always making us see the humour in life and for making sure my glass is never empty!

I cannot explain how grateful I am to have such great friends, who from near or afar, never fail to pick me up and make me smile when times are tough, or simply remind me that no-one in the real world cares what a mass spectrometer is anyway!

Some marathons are tougher than others, give me 26.2 miles any day!

Abstract

Monitoring changes in protein higher order structure (HOS) is an important regulatory requirement for reasons of safety, efficacy and potency of the resultant therapeutic. In particular, as monoclonal antibodies (mAbs) emerge rapidly as a dominant class of therapeutics, so does the need for suitable analytical technologies to monitor for changes in HOS of these complex large biomolecules. Reference materials (RM) serve a key analytical purpose of benchmarking the suitability and robustness of analytical procedures for both drug producers and regulators. Here two different model systems, of increasing complexity based on commercially available RMs, have been developed and used to illustrate the assessment of repeatability and reproducibility of a range of mass spectrometry based analytical platforms for protein HOS measurements. The simplest model protein RM, recombinant human growth hormone, rhGH, was used to illustrate the development and systematic evaluation of more established platforms hydrogen deuterium exchange-mass spectrometry (HDX-MS) and ion mobility spectrometry-mass spectrometry (IMS-MS) and a more novel platform, fast photochemical oxidation of proteins (FPOP-MS) for protein HOS measurements. Both global and localised changes in rhGH HOS, induced due to the presence of zinc ligands were identified using the methods developed and cross-platform measurements compared. Structural conclusions were underpinned using a statistical approach developed for HDX-MS measurements in which measurement variability was differentiated from small but significant changes in HOS. These HDX-MS and IMS-MS methods approaches were validated using a mAb-based RM from the National Institute of Science and Technology (NISTmAb) and two Fc-glycan variants generated using a simple enzymatic protocol and structural changes characterised. Measurements reproducibly demonstrated decreases in structural stability as a result of loss of Fc-glycan structure. These data promote the use of these, rhGH and NISTmAb RM based model systems, for both validating and establishing the sensitivity of analytical methods for the detection of HOS changes of mAbs and other protein therapeutics. These data also demonstrate the suitability of HDX-MS, IMS-MS and FPOP-MS as valuable analytical methods within the biopharma toolbox.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Chapter 1
Introduction

The introduction of this thesis aims to present established scientific theory to support the experimental data presented herein in the later parts of this thesis. The introduction can be separated into two sections; mass spectrometry (MS) theory and the application of MS methodologies for the measurement of proteins and protein higher order structure.

1.1 Mass Spectrometry theory

MS is an analytical technique that facilitates the study of gas-phase ions, measuring them based on their mass-to-charge (*m/z*) ratio and intensity. Measurements can be used to abstract masses of single elements, molecules or used to gain structural insights into the composition of a molecular ion. It is the ability to provide both qualitative and quantitative insights regarding an analyte that makes MS such powerful analytical technique.

In light of the scope of this thesis, the theory presented here will focus on MS techniques relating to the analysis of proteins and peptides and the specific MS instrumentation used to produce the data described herein. A general overview of the history of MS and the principles of the technique will be provided before discussing the technical details of the main components of a mass spectrometer including ionisation, mass analysis and detector stages of the instrumentation, and finally the hyphenation of ion mobility spectrometry (IMS) apparatus within the MS workflow.

1.1.1 History and overview

The development of MS goes back over a century, to J.J. Thomson's experiments in where he used magnetic and electric fields to deflect positively charged canal rays within a discharge tube. By evaluation of the deflected parabolas, collected on a photographic plate situated at the end of the tube, Thomson was able to obtain the first mass spectrum of several elemental ions including H^+, H^{2+} and He ions and the isotopes Ne^{20} and Ne^{22} [1].

Developing the instrumentation further, his research assistant, F. Aston, who introduced superior velocity focusing, giving birth to the first focusing mass

spectrometers, which would later provide the first proof of the existence of isotopes [2].

Mass spectrometers have come a long way since the days of Thomson and Aston, though, essentially those founding characteristics of Aston's first mass spectrometer remain. Mass spectrometers can be simplified to three essential elements, a) an ionisation source, used to produce gas-phase ions, b) a mass analyser, which sort ions based on their mass-to-charge ratio (*m/z*, Th) via magnetic and/or electric fields and a c) detector, measuring the ions in proportion to their abundance (Figure 1). The electrical output from the detector is then converted via data analysis algorithms, into a mass spectrum, reflecting the *m/z* of ions present vs relative intensity.

To maximise the ion free path, both the mass analyser and detector stages are operated at high vacuum (low pressure) allowing ions to reach the detector without undergoing collisions with other gaseous molecules preventing unwanted reactions, fragmentation or charge loss of the analyte ion. An electrical gradient applied across the instrumentation transmits the ions from the source to the detector, though often additional ion guides are used.

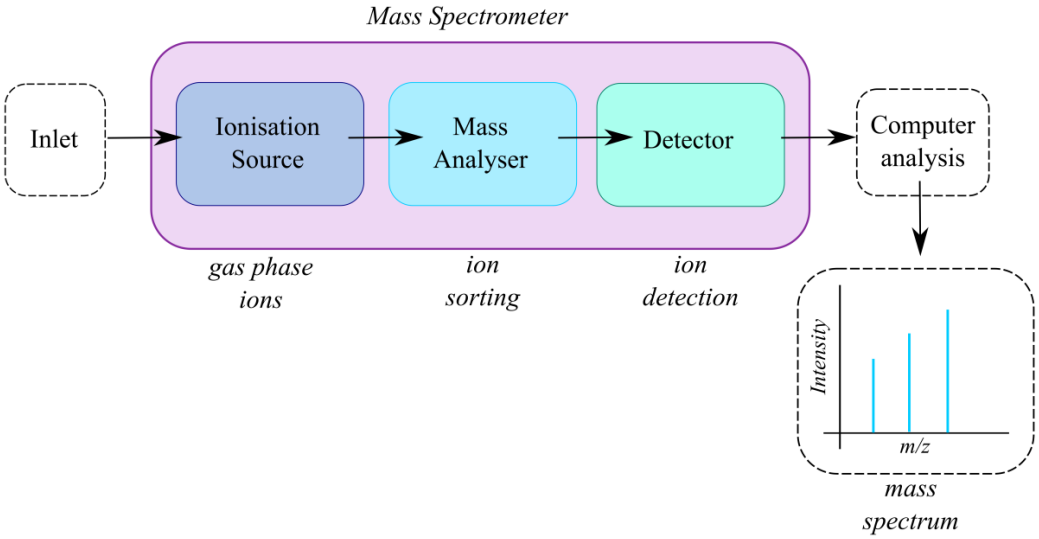


Figure 1. Schematic of the components found in all mass spectrometers, an ionisation source, a mass analyser and a detector.

Though these components described represent the foundation of every mass spectrometer, modern day mass spectrometers vary widely in the types of components used and the configurations adopted. Such configuration can include the presence of

multiple mass analysers, carefully selected to fulfil resolution, mass accuracy, mass range and sensitivity requirements for a given analysis of a chemical species.

1.1.2 Ionisation sources

As a previously mentioned, the first stage of the mass spectrometer is that on the ionisation source. Many types of ionisation sources are available and are typically grouped into what is referred to as “hard” and “soft” ionisation sources and are selected based on the properties of the analyte being measured. For small volatile molecules, “hard” sources such as electron impact (EI) or atmospheric pressure chemical ionisation (APCI) are often applied. However for less volatile and larger molecules, such as the analytes discussed in this thesis, “softer” ionisation techniques are required to produce intact gas phase ions and avoid over fragmentation of the analyte mol

ecules; these techniques include matrix-assisted laser desorption ionisation (MALDI) and electrospray ionisation (ESI). The theory of ESI will now be discussed in greater detail reflecting the use of the technique to produce the data presented in this thesis.

1.1.2.1 Electrospray ionisation (ESI)

Early MS ionisation sources required a gas phase analyte, restricting MS analysis to small volatile and stable compounds. However, the analytical space covered by MS was revolutionised by the development of electrospray ionisation (ESI) in 1984 by Fenn and Yamashita [8], which allowed the analysis of analytes directly from the solution phase at atmospheric pressures. Importantly, this “soft” method of ionisation induced little or no fragmentation during ionisation, allowing large bio-macromolecules to survive the transformation from the liquid- into the gas-phase intact and facilitating their subsequent MS detection [3].

The applicability of ESI to wide mass ranges, from several hundred Da up to ~MDa [4, 5], combined with the technique’s ability to produce multiply charged ions, i.e. $[M+zH]^{z+}$, has revolutionised the world of bioanalysis giving birth to the “omics” fields.

The production of gas-phase ions during ESI can simply be broken down into three main stages a) the production of charged droplets, b) shrinkage of these droplets due to evaporation and c) the production of gas-phase ions.

During ESI, an analyte solution is flowed through a charged capillary, ~ kV, and into an electric field applied at the entrance of the mass spectrometer. Charge accumulation occurring at the liquid surface of the capillary tip, causes repulsive forces to build, forming what is known as a Taylor cone. These repulsive forces combine with the pulling force of the electric field and disperse the liquid into a fine aerosol [6, 7] containing smaller droplets with high excessive surface charges and neutral interiors. Evaporation caused by applied drying gases (often N_2), cause the surface tension and coulombic repulsion of these droplets to increase until the Rayleigh limit is exceeded by the surface charge [8], resulting in fission of the droplet into smaller successive droplets. These cycles continue until a single desolvated charged analyte ion is produced, though the exact mechanism of this final ion formation step is thought to be analyte dependent and can be described by the following models and in Figure 2, below;

The **ion evaporation model (IEM)** [9], describes when the radius of a droplet becomes $\leq$10 nm, a small solvated ion is ejected from the droplet surface as a result of the high surface charge. Excess solvation of this ion is then lost through collisions with background gas ions resulting in a single fully desolvated ion. This model is thought to be the dominating ion release mechanism for small low molecular weight species [10, 11].

The **charge residue model (CRM)** [12], is widely accepted as the dominant ion production mechanism for large globular species, such as natively folded proteins. The process describes how nano-droplets containing single analyte ions evaporate to dryness. As droplet size decreases during the evaporation process, protons are ejected from the surface via IEM, shedding charge as the radii decreases, so as to maintain the surface charge within the Rayleigh limit. This process repeats itself until no solvent buffer remains and the analyte ion adopts any residual charge on the surface as its own.

The **chain ejection model (CEM)**, was recently proposed by Konermann et al [13], as a mechanism for ion production for intrinsically disordered proteins. During the drying stages of the droplet formation, a charged hydrophobic region of the disordered poly-peptide chain is ejected from the droplet to reduce surface charge, as with in IEM. This process continues until the whole disordered polypeptide chain is ejected as a charged ion.

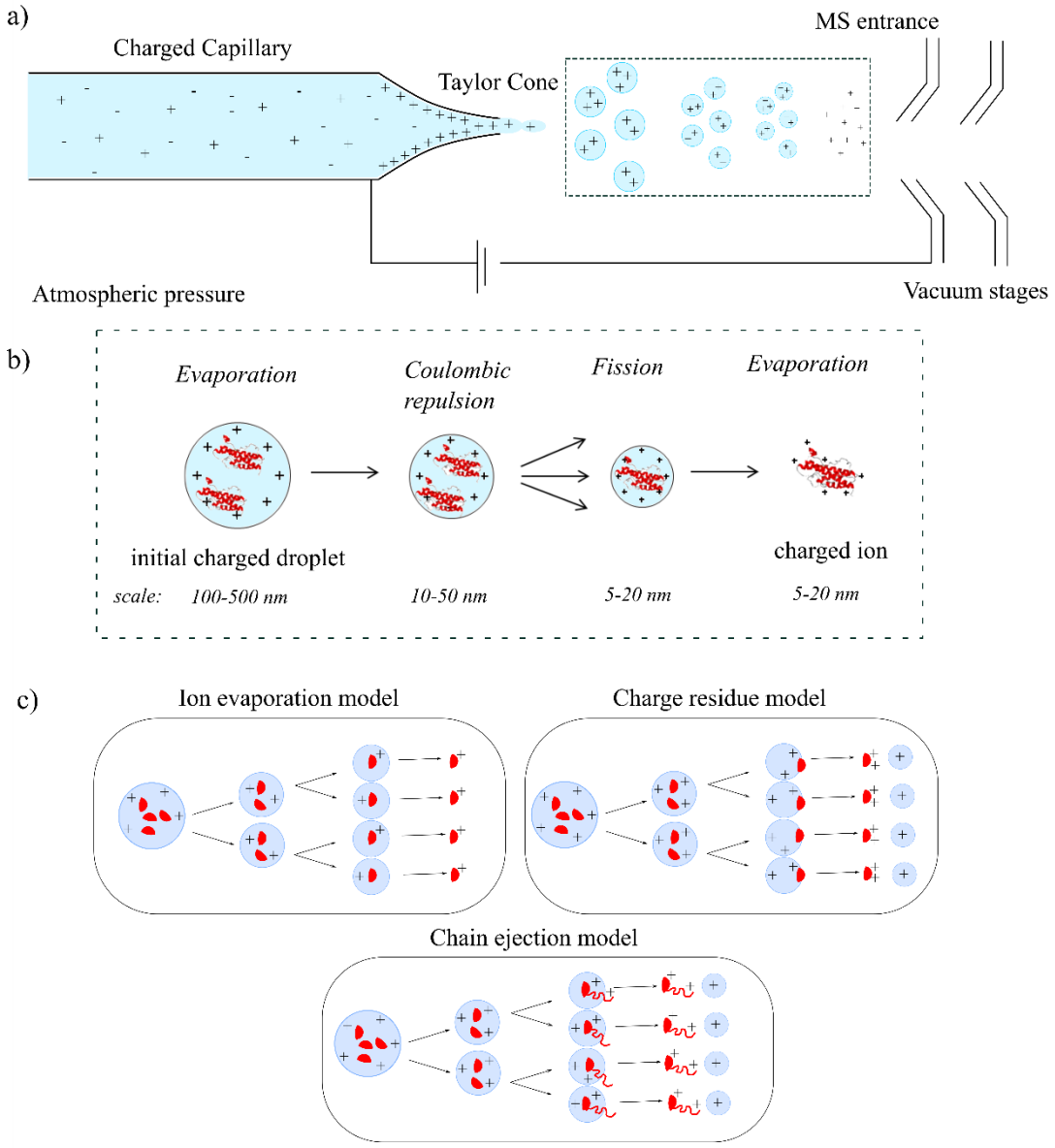


Figure 2. The production of gas-phase ions during ESI. Charged droplets are produced from an analyte containing solution by a) flow through a charged capillary into an electrical field and dispersed via a Taylor cone b) these charged droplets then undergo shrinkage due to multiple rounds of evaporation and fission until finally producing an analyte ion; c) three different models exist for describing the mechanisms of ion release from charged droplets, the ion evaporation, charge residue and chain ejection models and vary based on analyte type, generally existing for small molecular weight molecules, large globular species such as proteins and intrinsically disordered proteins respectively. The exact mechanisms are explained in the text, section 1.1.1.

The introduction of nano ESI (nESI), using nL/min flow rates, has further progressed "soft" ionisation. The smaller capillary orifice produces smaller droplets, with more favourable surface area to volume ratios, requiring less evaporation and removing the need for denaturing source conditions [14]. nESI also exhibits higher ionisation efficiencies and salt tolerances permitting the use of less volatile buffers, allowing more biologically native-like environments for analysis [15].

Following the generation of ions, the ions are then guided into the mass spectrometer to undergo m/z separation in the mass analyser.

1.1.3 Mass analysers

A extensive range of mass analysers exist, each operating with an individual ion separation principle, varying in mode of action to transmit and separate analyte ions based on their m/z. The selection of the most appropriate mass analyser type is a key step in the analytical process, dependent on the properties of the analyte post ionisation and the context for which the analysis is occurring.

Mass analysers can be best described and compared based on five fundamental characteristics, mass range limit, analysis speed, transmission, mass accuracy and mass resolving power;

Mass range limit, describes the *m/z* range across which the mass analyser can separate ions across and defined in unit of, Th, or u, if an elemental charge is defined, such as z =2. This limit can be theoretically limitless, such as with time-of-flight (ToF) mass analysers, or be defined to a specific upper mass limitation such is the case with a quadrupole mass analyser.

Analysis speed, is the speed at which a mass analyser can perform measurements across a given mass range, and is typically expressed in units per second (u s^{-1}).

Transmission is the ability of the mass analyser to transmit ions from the entrance of the analyser to the detector and is calculated by the number of ions entering the mass analyser divided by the number of ions reaching the detector. The transmission of an analyser is related to the sensitivity of the mass spectrometer, though ionisation efficiency, duty cycle and detector sensitivity are also contributing factors.

Mass accuracy, describes the accuracy at which the mass analyser is able to measure the *m/z* of an ion (m/z_{measured}) compared with its theoretical *m/z* value ($m/z_{\text{theoretical}}$), and is often expressed, as parts per million (ppm) as;

Equation 1

$$\text{Mass accuracy (ppm)} = \left[\frac{m/z\ \text{measured} - m/z_{\text{theoretical}}}{m/z_{theoretical}}\right] x\ 10^6$$

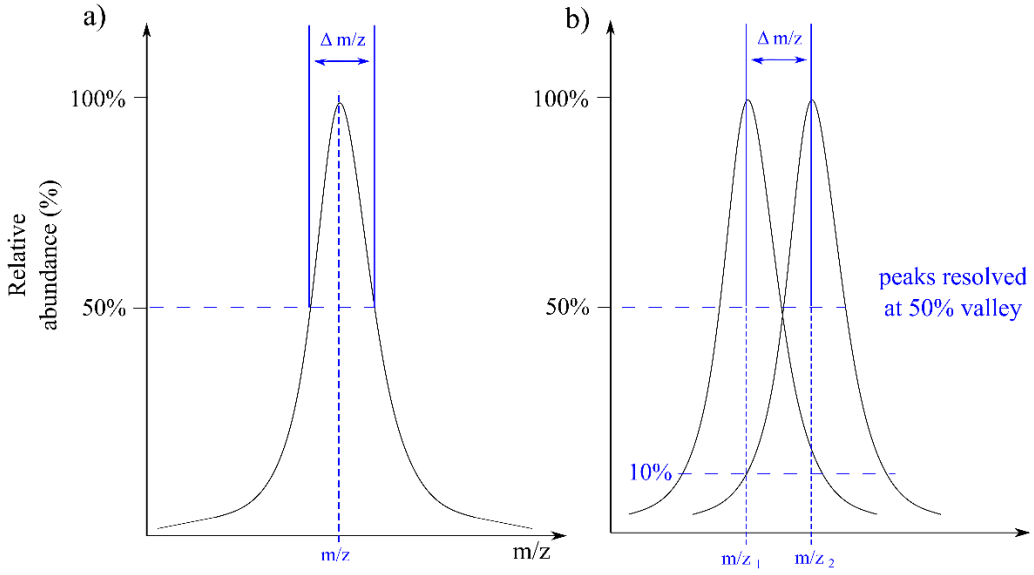


Figure 3. Illustration of the concepts of mass resolution. Peak width at 50% a) full width half maximum (FWHM) and b) two peaks are considered resolved if the mass difference between the valley between them ≤ 50% of their height. For Orbitrap instruments the valley must be ≤ 10% of their height.

Mass resolution and/or resolving power, describes the mass analyser's ability to distinguish between signals of two ions of similar *m/z* (Figure 3 b). Though these expressions of mass resolution and resolving power are often used interchangeably, they do in fact define two slightly different properties of a mass analyser. According to IUPAC guidelines[16], mass resolution refers to the m/Δm, where m is the mass of the ion of interest and Δm is the spacing between two adjacent ions of similar *m/z* with the valley of a defined % threshold between. This percentage threshold is often mass analyser specific, for example a ToF mass analyser, a percentile maximum of 50% is used (Figure 3 b), whereas for an Orbitrap mass analyser this value is 10% [17]. Resolving power, defines the ability of the mass spectrometer to distinguish between two peaks of very similar *m/z*, and is expressed in terms of the peak width in mass units.

Mass accuracy is intrinsically linked to the mass resolution of the mass analyser, such that a low resolution analyser will give poor precision on the $m/z_{measured}$, resulting in an inaccuracy of the centroid $m/z_{measured}$.

Mass analysers can be separated into two main classes based on their ion separation principles, scanning devices such as quadrupoles and non-scanning devices such a time-of-flight and Orbitrap mass analysers.

1.1.3.1 Quadrupole mass analysers

Quadrupole mass analysers [18–20] are scanning devices, using dynamic electric fields to filter and direct the transmission of ions, of a particular *m/z* value, towards the detector at a given time. They are generally considered low resolution instruments, typically operated at unit resolution, but have high scanning speeds and high sensitivity.

Quadrupoles consist of four cylindrical rods, assembled into two electrically linked perpendicular pairs (Figure 4). To each pair, opposing electric potentials, consisting of both direct current (DC), U, potentials and an oscillating radio-frequency (RF), V, voltages are applied. Each of these pair of rods are within phase (+ or -), but in opposing phase with the other perpendicular pair, and at any given time, t, the total potential applied, φT, to each set of poles can be described by;

Equation 2

$$\phi T = U + V\cos\omega t$$

Equation 3

$$-\phi T = -(U + V\cos\omega t)$$

It is these potentials applied to these poles, that dictates if an ion of a given *m/z* will have a stable trajectory through the quadrupole to the detector. Half the distance between each pair of hyperbolic rods, is termed r_o.

As an ion enters the mass analyser, it traverses towards the detector along the z axis, however, additional oscillation also occurs in both the x and y trajectories in response to the applied fields to the rod pairs. These oscillation strengths are dependent on the particular *m/z* of the ion, if there is an unstable trajectory in either x or y planes,

collision of the ion with these rods occurs, causing the ion to be neutralised and prevented from reaching the detector. In general, lighter ions experience larger effects from the oscillating V potential, whilst the U potential dominates for heavier ions [21].

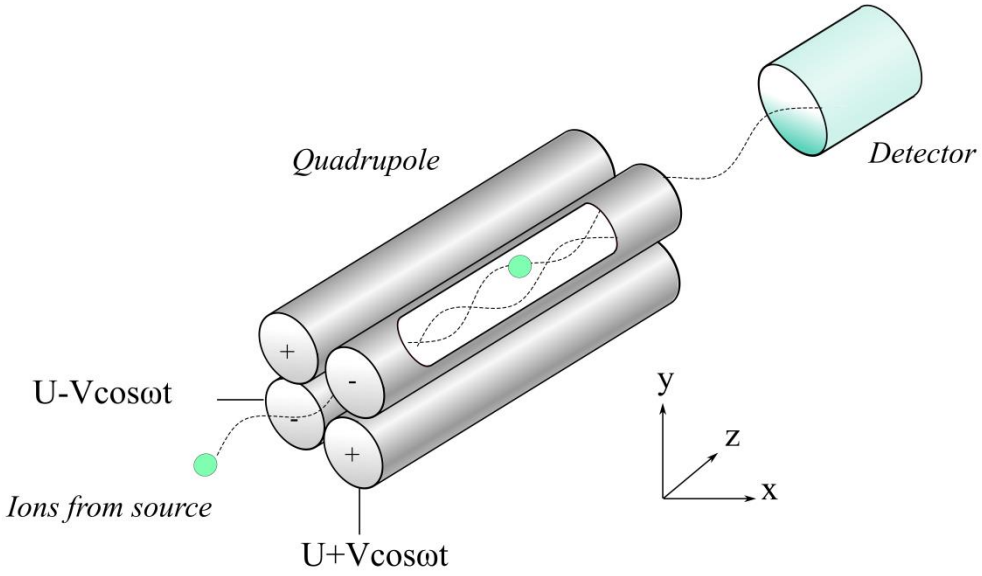


Figure 4. Schematic of a quadrupole mass analyser and detector. As ions, generated at the source, enter the quadrupole they traverse towards the detector along the z axis whilst also experiencing oscillations in the x and y axis under the influence of the opposing electric potentials applied to the two pairs of hyperbolic rods that make up the quadrupole.

To understand whether an ion will experience an unstable or stable trajectory we can consider the Matthieu's differential equations, which models the theoretical stable trajectories of a series of ions should no collision with quadrupole rods occur (Figure 5). For a quadrupole, increasing U and V voltages whilst maintaining them at a constant U/V ratio, facilitates the transmission of multiple ions. Figure 5, illustrates two scanning lines, presented in blue and red, each line maintaining a constant U/V ratio yet the linear rate of potential increase (or slope) is changed between red and blue lines respectively. The area underneath each prospective slope indicates those ions which have stable trajectories in both the x-z and y-z dimensions, so that their trajectories are $< r_{o,}$ preventing collision with the hyperbolic rods and their successful transmission to the detector.

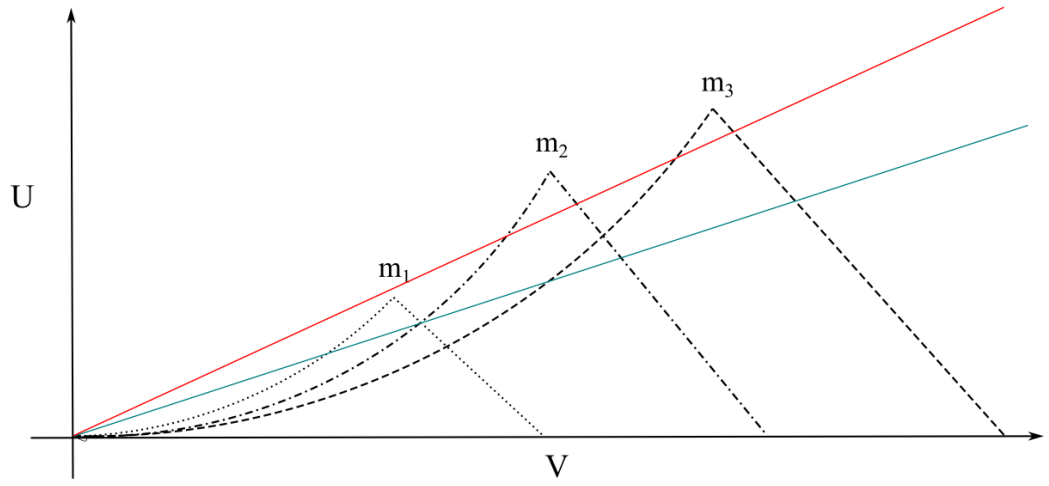


Figure 5. Stability areas of a series of ions ($m_1 < m_2 < m_3$) as a function of U (DC) and V (RF) potentials. Multiple ions are transmitted through the mass analyser by a linear increase of potential whilst maintaining a constant ratio of U/V. The red and green lines represent two differing scanning modes, stable ion trajectories for each slope are indicated by the area underneath the respective scanning line. A lower resolution mode is represented by the green line, in which all m_1, m_2 and m_3 ions are transmitted, however due to overlapping profiles, m_2 and m_3 ions are unresolved. The red line represents a higher resolution mode in which m_2 and m_3 are resolved yet m_1 is no longer transmitted. Figures adapted from de Hoffman and Stroobart [22].

If we consider the green slope, m_1, m_2 and m_3 are all transmitted, yet, as a result of their overlapping profiles m_2 and m_3 ions are unresolved. Increasing the scan slope from green to red, increases the resolution of the scan, that is the ability to separate of the m_2 and m_3 ions, however prevents the transmission of the m_1 ion. Therefore it is the scan slope which defines the resolution of a quadrupole MS experiment.

In contrast, the absence of the applied U potential allows the transmission of all ions, yet resolution becomes zero. If we increase V from 0 to past m_1, then all ions with mass < m_1 will experience unstable trajectories, therefore it is the applied V potential that dictates the minimum transmission mass of an ion, thus dictating the scanning range of the experiments.

1.1.3.2 Time-of-flight (ToF) mass analysers

In contrast to quadrupole based mass analysers, ToF mass analysers are other non-scanning devices transmitting all ions within a *m/z* range from a single pulse of ions

and recording a mass spectrum of all ions present. Theoretically with no upper mass limitation, they are particularly amenable to analysing large molecular weight ions, and hence ToFs are particularly amenable the analysis of intact large biomolecules.

The principle of ToF analysers was first described by 1946 by Stephens[23], but it was not until 1955, when the designs of Wiley and McLaren [24] were used to produce the first commercially available instrumentation that their first practical application was realised.

ToF mass analysers separate ions of differing *m/z* based on their varying velocities as they drift though a field free region, known as a flight tube, following their acceleration by an applied electrical field. The ions are accelerated by an applied potential gradient between source and extraction grid, and the ToF mass analyser measures the time, t, taken for an ion, of *m/z*, to move through the flight tube and hit the detector situated at the end of the flight tube. Early generation ToF instrumentation were based were based on linear ToF mass analysers, such as represented in Figure 6.

Considering an ion of mass, m, and a total charge of q=ze, as it leaves the source and is accelerated by a potential V_s, its electronic potential energy E_{el} is converted into kinetic energy E_k, as described by;

Equation 4

$$E_k = \frac{mv^2}{2} = qV_s = zeV_s = E_{el}$$

This can be rearranged to calculate an ions velocity, v;

Equation 5

$$v = (\frac{2zeV_s}{m})^{1/2}$$

revealing a velocity dependence of the m and z of an ion. So therefore, everything being equal, a lighter ion will reach the detector faster than a heavier ion.

Following the initial acceleration of the ion, the ion travels in straight line at constant velocity until it reaches the detector. The time, t, needed to travel the length of the path, L, to the detector is given by;

Equation 6

$$t = \frac{L}{v}$$

Combining Equation 5 and Equation 6, gives us;

Equation 7

$$t^2 = \frac{m}{z}\left(\frac{L^2}{2eV_s}\right)$$

If path length and applied potentials are constant, a value for the *m/z* of an ion can be calculated from the measurements of t^2.

The resolution of a ToF mass analyser is defined by;

Equation 8

$$R = \frac{L}{2\Delta x}$$

where Δx is equal to the thickness of the packet of the ions approaching the detector, and resolution is proportional to flight time and path length. Increasing path length is one way to increase the resolution of the instrument, although this can have a negative effect on sensitivity of the mass analyser as a result of reduced ion transmission due to ion scattering or angular dispersion of ion beam.

Earlier generation linear ToF instruments were plagued by poor resolution, caused by several factors resulting in apparent flight time distributions among ions of the same *m/z*. These factors included; variations in ion formation pulse lengths resulting in the formation of ions at differing times (temporal distribution), the formation of ions at slightly different positions within the source (spatial distribution) and variations existing in the initial E_k of the ions (kinetic energy distributions). To overcome these factors leading to significant resolution improvements of ToF instruments, two methods in particular, delayed pulsed extraction and use of a reflectron, were developed.

Continuous extraction (Figure 6 a), used in conventional instrumentation, involves the immediate acceleration of ions towards detector following ionisation. Any variations in E_k, will result in isobaric ions reaching the detector at different times, causing peak broadening in the resultant mass spectra.

Delayed pulsed extraction, however, allows for the correction of E_k distributions amongst isobaric ions by the application of a delay between ion formation and ion extraction towards the detector (Figure 6 b). Those ions with lower E_k, situated closer to the pulse, experience greater acceleration than those ions with an initially higher E_k. This difference in velocities allows both these ions to hit the detector at the same time, improving peak resolution of the resultant mass spectra. Delayed extraction can be thought of in terms of energy focusing of the ion beam, achieved either by adjusting the amplitude of the voltage pulse or by the time delay between ion formation and extraction phases. Hence resolution can be optimised by optimisation of both pulse amplitude and delay.

The **reflectron** consists of a series of ring electrodes and grids which are used to create a retarding electric field (Figure 7) [25]. Positioned at the opposing end of the pusher/source within the flight tube, the reflectron reflects the ions back through the flight tube towards the detector now situated near the source. The depth at which an ion will penetrate the reflectron is dependent on the E_k of the ion. Ions with higher E_k penetrate the reflectron further than an isobaric ion of lower E_k before being reflected back towards the detector, hence the higher velocity ion will exhibit a longer flight time. Tuning of the reflector voltage allows for the point at which two isobaric ions of differing E_k react the detector to be aligned, minimising the impact of E_k distributions on resolution.

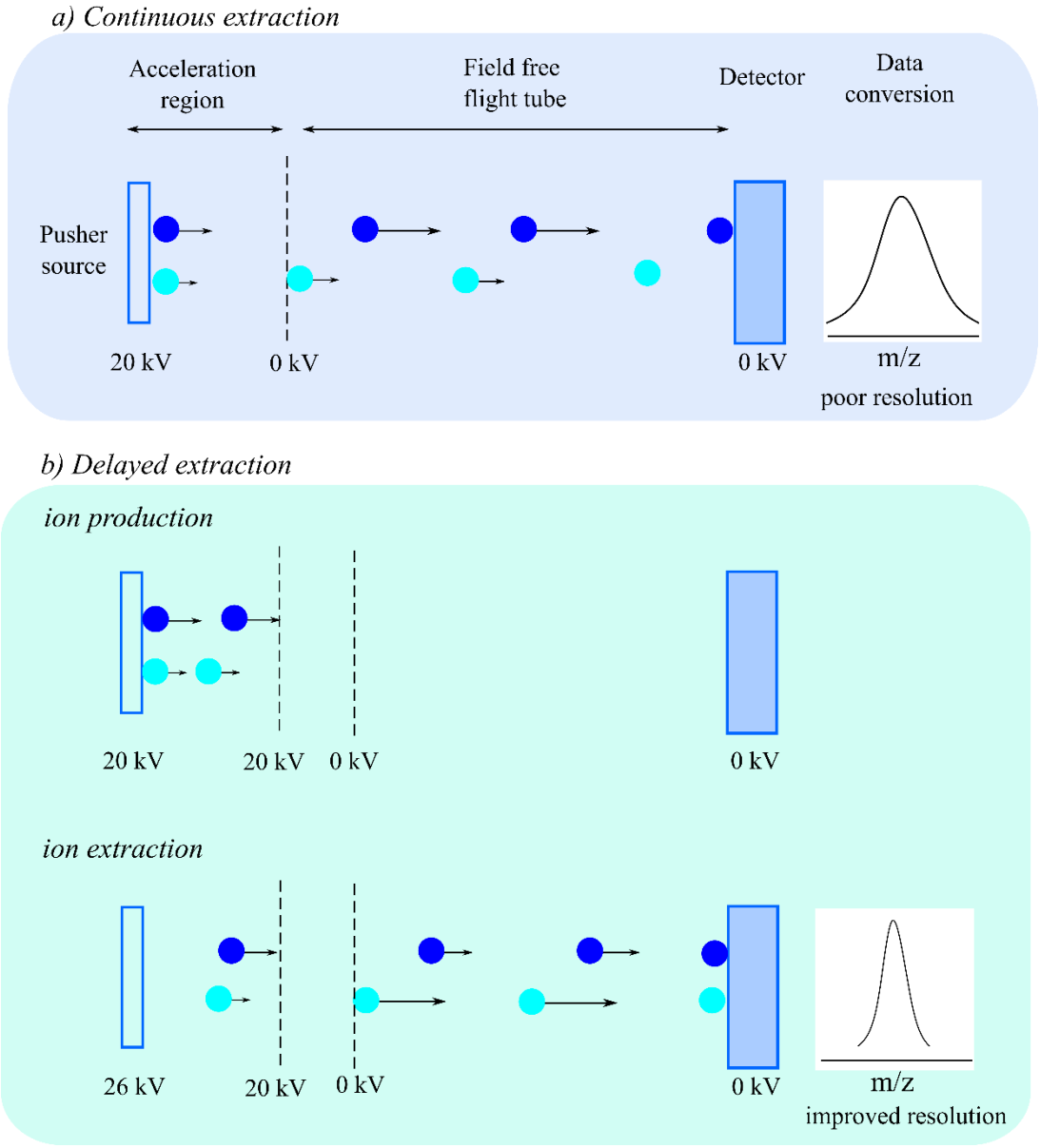


Figure 6. Schematic describing two extraction modes within a linear ToF mass analyser. Two isobaric ions that differ only in kinetic energy, one high (dark blue) and one low (light blue) energy are represented. During an a) continuous extraction mode, this disparity in kinetic energy of the ions, leads to variations in flight time and ultimately poor resolution. Using a b) delayed extraction mode, a delayed pulse is applied to the ions prior to them reaching the flight tube, leading to an increased acceleration towards the detector. The lower energy ion situated closer to the source of the pulse, experiences greater acceleration resulting in an increased velocity and both ions reaching the detector at same time. This leads to subsequent increased resolution of the resulting mass spectra.

1.1.3.3 Orthogonal-acceleration ToF mass analysers

The coupling of ToF mass analysers with pulsed ionisation techniques, such as MALDI, is directly compatible, due to their provision of short, defined ionisation times and small localised ionisation region. Yet, the requirement of the pulsed injection of ions into the flight tube, makes the pairing of linear ToF mass analysers with continuous ion sources (i.e. ESI) incompatible.

However the development of orthogonal-acceleration ToF (oa-ToF) mass analysers by Dawson et al. [26], included adaptations to bring the two technologies together. These adaptations allow for a continuous ion beam to be transformed into pulsed packets of ions, with precisely defined injections times, that are introduced into a perpendicularly positioned ToF flight tube (Figure 8). As the continuous ion beam is directed from the source it passes into the orthogonal accelerator and across an electrical pusher plate, situated opposite to the entrance to the flight tube. Periodic application of an injection pulse voltage creates packets of ions which are pushed towards the flight tube in a field free region. As they reach the entrance of the flight tube these ions are then accelerated towards the detector using acceleration voltages as with conventional ToF. The combination of these two technologies has been revolutionary in maximising the high efficiency and compatibility of chromatography techniques with continuous ionisation sources whilst achieving the high resolution mass analysis of ToF mass analysers. There is however a price to pay for this combination, however, and this if in the form of the reduced duty cycle of the oa-ToF instrument due to the need to wait for one packet on ions to reach the detector before the next pusher event.

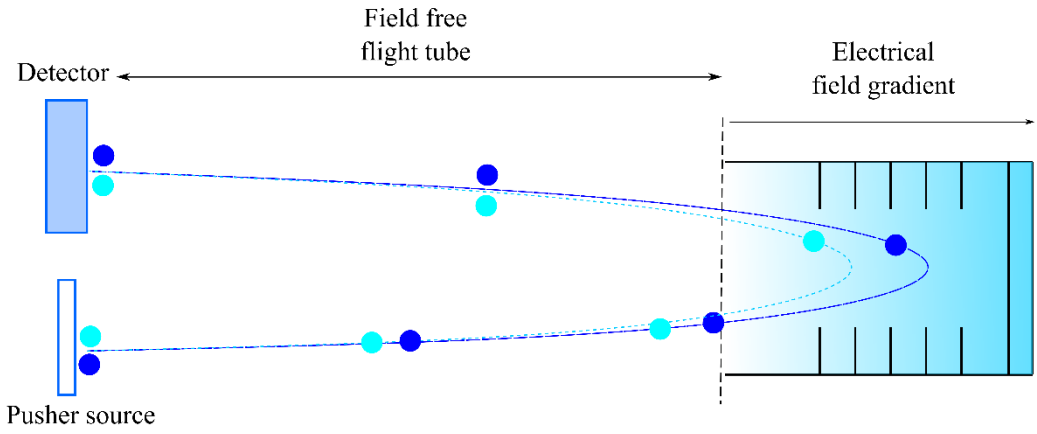


Figure 7. Schematic of a reflectron ToF instrument. Two types of ion, of the same *m/z* but differing kinetic energy, high energy (dark blue) and low energy (light blue) are produced at the source. The greater velocity of the higher energy ion, causes the ion to reach the reflectron first and penetrate it further. The lower energy ion is reflected back towards the detector at a shorter distance and thus overall has a shorter flight path. Appropriate tuning of the reflectron voltage allows both ions to reach the detector simultaneously, improving resolution of the mass analyser.

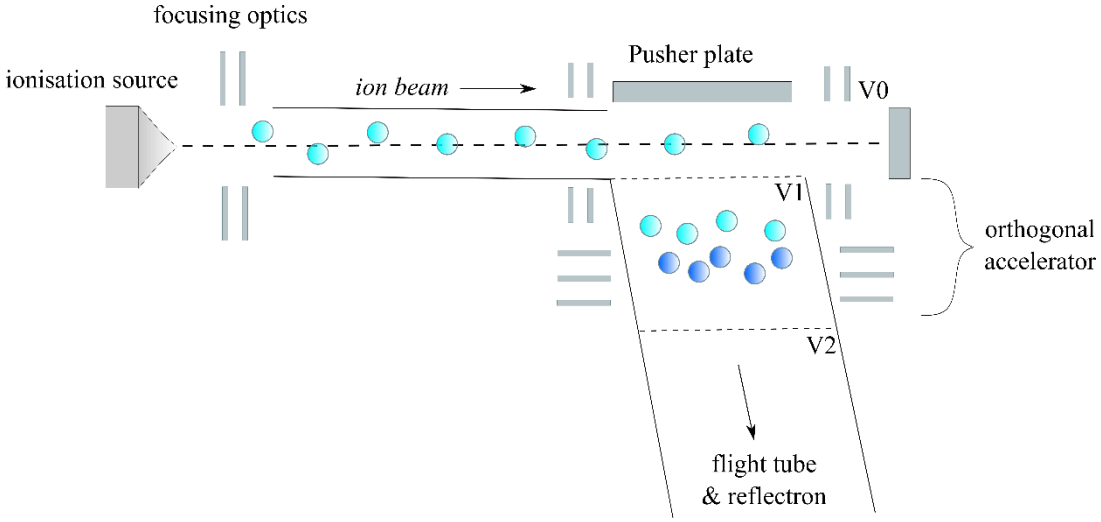


Figure 8. Schematic of an orthogonal accelerator-ToF instrument. Ions from a from a continuous ionisation source are pulsed into a ToF mass analyser via an orthogonal accelerator.

1.1.3.4 Orbitrap mass analysers

Orbitrap mass analysers are electrostatic ion trapping instruments, that operate on a principle of trapping ions based on their attractions to an electrically charged spindle. Unlike ToF and quadrupole based instrumentation, they are devoid of a physical detector but instead convert the frequency of an ions oscillation within the trap into a *m/z* value via Fourier Transform (FT). The concept of electrostatically trapping ions around a central spindle was first described by Kingdon [27] as early as 1920s, yet the orbitrap mass analyser was not developed until 1990s by Alexander Makarov [28] and not commercialised by Thermo Fisher Scientific in 2005. Like other FT-MS instruments, it demonstrates high resolution, high mass accuracy and wide mass range.

The Orbitrap itself consists of two electrodes, an outer barrel shaped electrode of which a second central spindle shaped electrode is located at its heart (Figure 9). The electrodes are separated by a small gap and a DC potential of several kV applied to the central electrode, with the outer electrode grounded. As ions are injected into the side of the outer electrode, they begin to rotate around the central electrode. The ions also start to oscillate up and down the central electrode, forming spindle like orbital oscillation due to the potential and the Orbitrap's specific geometry. This harmonic oscillation along the axis of the applied field, is defined by a specific frequency which in turn in converted into the *m/z* of the ion by FT.

To describe this process in more detail, we consider the potential inside the trap, U, where r and z are cylindrical coordinates of the trap, C is constant, k is field curvature, and R_m is the characteristic radius we can obtain:

Equation 9

$$U(r,z) = \frac{k}{2}\left(z^2 - \frac{r^2}{2}\right) + \frac{k}{2}(R_m)^2 ln\left(\frac{r}{R_m}\right) + C$$

From this equation it is observed that the voltage gradient in the z direction, is proportional to k;

Equation 10

$$\frac{\partial U(r,z)}{\partial z} = kz$$

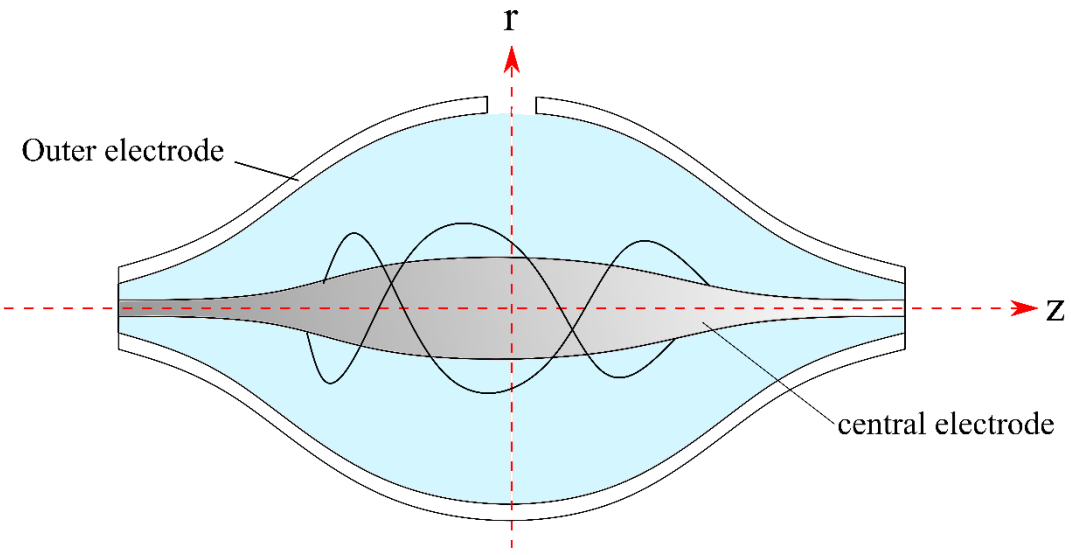


Figure 9. Schematic of an Orbitrap mass analyser. As ions are entered into the Orbitrap, the applied potential to the central electrode causes the ions to orbit the electrode whilst also oscillating in the z axis, forming a spiral like motion. These spiral like motions are of a specific frequency which can be related to the *m/z* of the ions by Fourier Transform.

Accelerating an ion of mass, m and total charge q=ze, along the z axis, using the force from an electrical field, we can consider that electrical force, -qkz, equal to the mass times the acceleration, $m(d^2/dt^2)$:

Equation 11

$$F_z = m\frac{d^2}{dt^2} = -q\frac{\partial U}{\partial z} = -qkz \quad or \quad \frac{d^2z}{dt^2} = -\frac{q}{m}kz$$

This equation describes a simple harmonic oscillator, to which an energy factor E_z, can be introduced, which is defined by:

Equation 12

$$qE_z = \frac{m}{2}\left(\frac{d_{zo}}{dt^2}\right)^2$$

Solving Equation 11 gives;

Equation 13

$$z(t) = z_0 \cos \omega t + \sqrt{(2E_Z/k \sin \omega t)}$$

Where frequency, ω, is equal to

Equation 14

$$\omega = \sqrt{(q/m)k}$$

This solutions demonstrates the dependence of the frequency on the m/q of an ion, whilst remaining independent of the kinetic energy of the injected ions.

The oscillation of the ions within the trap induce an image current, which is subsequently measured, converted by FT to give individual frequencies of ions and intensities. These are solved using

Equation 13 forming the resultant mass spectrum.

The wide mass range, high resolution and high mass accuracy of Orbitrap mass analysers make Orbitrap MS an attractive technique for the analysis of intact proteins.

Early generation Orbitrap instruments, however, demonstrated reduced resolution for ion mass increases even at constant *m/z* ratio, as a result of increased ion scattering due to collisions with background molecules within the Orbitrap, leading to transient decay [29]. Modern Orbitrap instrumentation now have alternative reduced pressure modes (~ low 10^{-10} $_{mbar}$), which can be operated for intact protein analysis, so as to reduce the statistical probability of such unwanted collisions, increasing resolution, mass accuracy and sensitivity for these higher *m/z* ions. Increasingly Orbitrap instruments have been utilised in the area of structural biology, where intact complexes of proteins have been reported in 4 MDa range at high mass resolution [5].

1.1.3.5 Hybrid mass spectrometers

Multiple mass analysers and/or types can be combined within a single instrument, to give hybrid instrumentation, facilitating flexibility of analysis and control of multiple MS experiments (known as tandem MS, described later in section 1.1.4) simultaneously. Two hybrid instruments, used in this thesis, based on combinations of quadrupole, ToF and Orbitrap type mass analysed will now be discussed.

1.1.3.5.1 Waters Synapt G2Si

A common hyphenated instrument, combines quadrupole and ToF mass analysers, termed a Q-ToF instrument. This allows for preselection or focusing/guiding of ions using a quadrupole mass analyser before *m/z* separation by oa-ToF. An example of one such instrument is that of a Waters Synapt instrument series [30]. Several generations of this instrumentation exist (G1, G2, GS and G2Si), and the G2Si in particular has been utilised in this thesis (Figure 10). An extra component of this instrumentation is that of the travelling wave ion mobility spectrometry (TWIMS) cells housed between the Q and ToF mass analysers which shall be discussed further in section1.1.6.2, but add further analytical capability to the instrumentation.

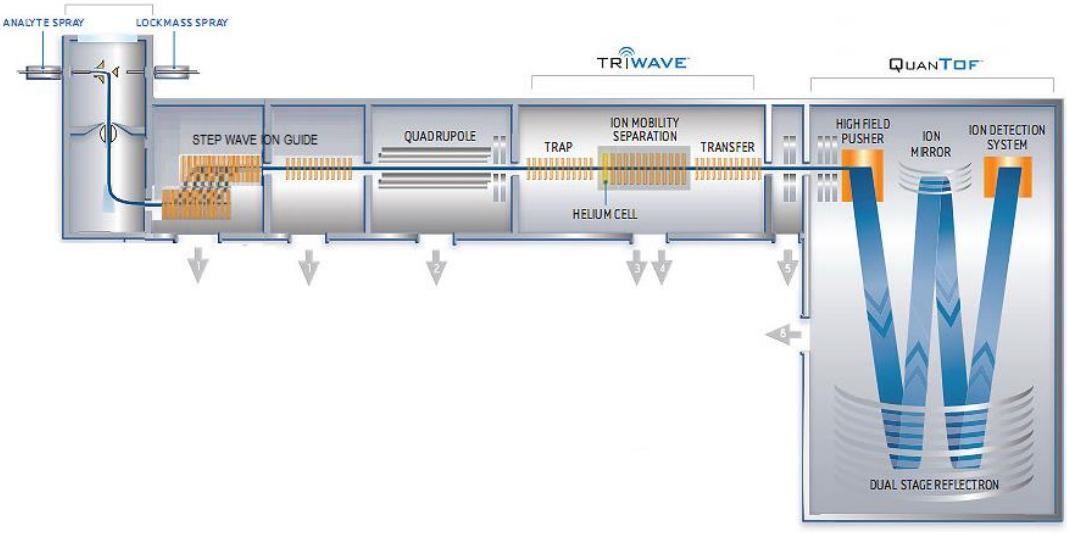


Figure 10. Schematic of Synapt G2Si Q-ToF instrument. Image taken from Waters [31].

1.1.3.5.2 Q-Exactive Plus

The Q-Exactive range of MS instruments are hybrid instrumentation that include a quadrupole mass analyser positioned prior to the Orbitrap [32]. This feature facilitates the preselection of ions in space and capitalises on the fast scanning speeds of quadrupoles. The selected ions are then transferred into the C-trap storage device, where they can be stored or undergo fragmentation before the final high resolution mass analysis in the Orbitrap. Figure 11, gives a schematic of the Q-Exactive plus instrument used to produce data in this thesis.

This hybrid nature, also extends the variety of tandem MS experiments that can be performed using the instrumentation, which is particularly attractive for proteomic type MS analysis. The theory of tandem MS will now be discussed.

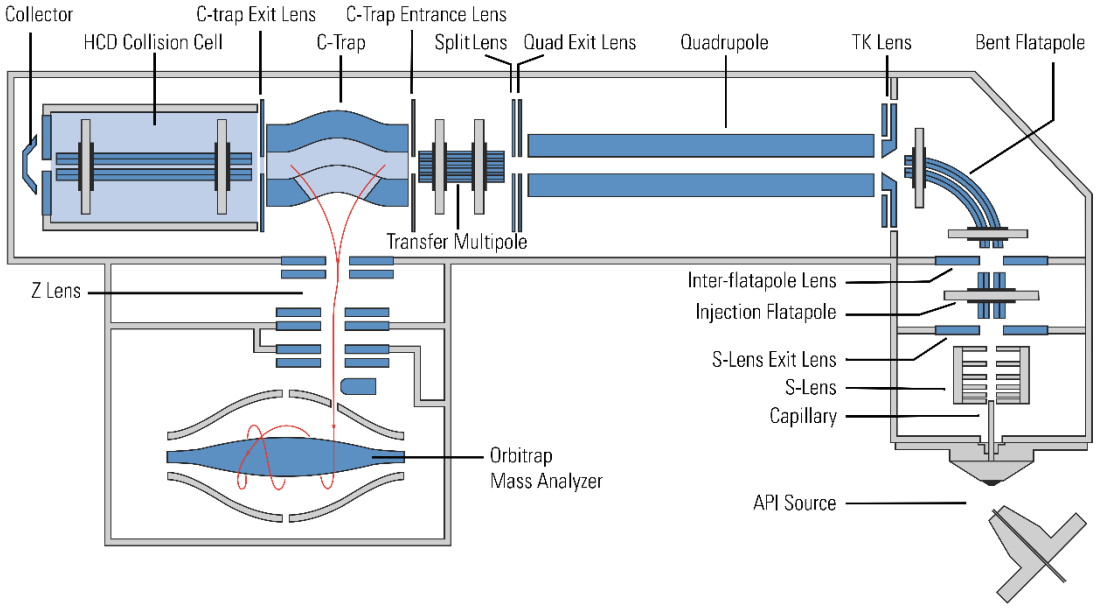


Figure 11. Schematic of Q-Exactive plus Orbitrap MS instrument. Image taken from Thermo Scientific [33].

1.1.4 Tandem mass spectrometry

Tandem mass spectrometry (MS/MS or MS^n) experiments involve the use of multiple mass analysers and refers to when multiple rounds of *m/z* separation are performed within a single mass spectrometry experiment, typically including a dissociation process between the rounds of separation [34].

Typically tandem MS, involves the isolation of an ion (precursor ion) during the first round of mass analysis, after which some form of ion activation process is then initiated, producing fragment ions (product ions), which are then subsequently measured during another round of mass analysis. These processes can be repeated giving MS^n rounds of experiments. These types of experiments facilitate specificity and vital structural elucidation information of parent ions and are applicable to small and large molecules alike.

MS^n experiments can be performed either by separation of the mass analysis stages in space or time (Figure 12). Space, refers to the physical coupling of multiple mass

analysers separated by a collision cell, such as with triple quadrupole (QqQ) instruments or Q-ToFs, and time, refers to when sequential events occur by trapping ions within a device, such as with ion trap instruments.

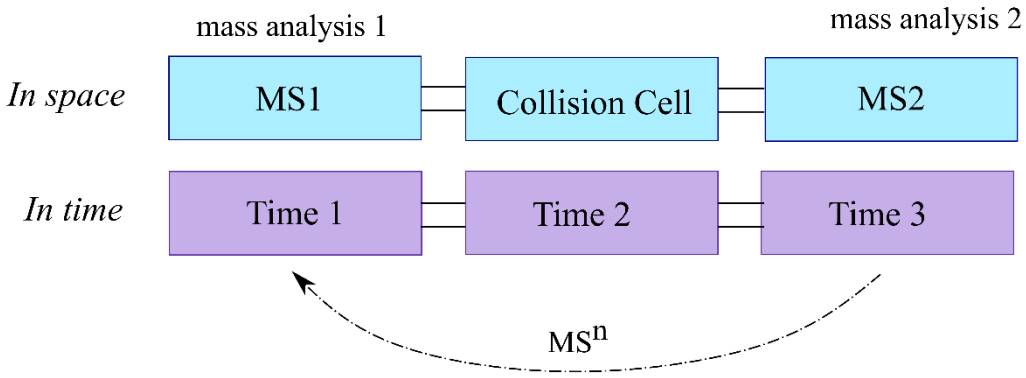


Figure 12. Tandem MS experiments can be performed by separation in space or in time depending on the instrumentation. Multiple rounds of tandem MS can be performed, termed MS^n.

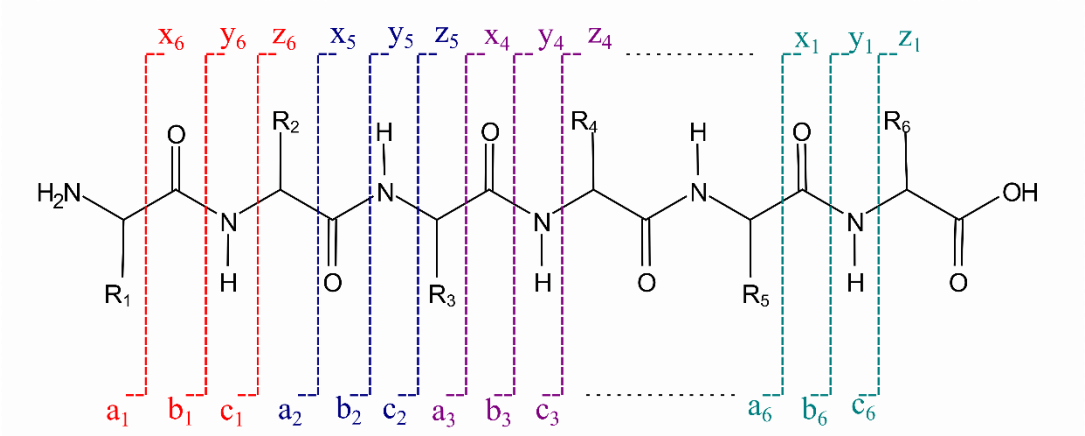


Figure 13. Schematic of the range of ions produced by MS/MS fragmentation of a peptide ion.

A whole plethora of activation processes have been developed to fragment an ion, these include collision-induced dissociation (CID) [35], electron-transfer dissociation (ETD) [36], electron-capture dissociation (ECD) [37], surface-induced dissociation (SID)[38], and photo-dissociation (PD) [39]. These techniques differ in the mechanisms of fragment ion production and therefore differ in the type of product ions produced. Table 1 describes the types of product ions produced from protein and peptide ions by these differing fragment techniques [40], highlighting some of the advantages and disadvantages of the techniques within this context [41].

As a result of its high efficiency and sequence coverage yields, CID is the most commonly used fragmentation techniques for protein and peptide MS. It is also the main fragmentation technique used in the Synapt, Xevo QToF and Orbitrap instruments described herein. During CID , a parent ion is collided with inert collision gas molecules, within the trap (Orbitrap) or the collision cell (Synapt/Xevo) of the instrument. These collision result in the transfer of kinetic energy to the parent ion, exciting the parent ion to a higher energy state. This excess internal energy is then dissipated via the characteristic fragmentation of the ion, which during CID occurs most commonly at peptidic bond, N-C, yielding b and y ions as described in Figure 13. If these fragments then undergo further collisions and subsequent fragmentations at the peptide bonds, the resulting mass spectra consists of a series of b and y ions differing in mass to the next corresponding fragment along the chain by one amino acid. Analysis of the mass spectra of these fragmentation series allows the sequencing of the corresponding polypeptide parent ion.

The fragmentation of a polypeptide ion into a series of heterogeneous fragment ions, thus facilitating sequencing, is in part directed by the polypeptide sequence and can be explained by the model proton theory [42]. The model describes how for protonated peptides produced during soft ionisation, the protons are initially localised to basic sites of the molecule such as the N-terminus and side chains of basic amino acids (arginine, lysine and histidine). The protons are then able to relocate from these sites to less basic sites of the peptide linkages following ion activation processes. Following this migration, charge-site-initiated mechanisms occur providing the desired fragment ions.

One major disadvantage of CID is its incompatibility with the analysis of post-translational modifications (PTMs). The relatively slow transfer of energy during CID, causes fragmentation to typically occur at the weakest point, such as is the case with PTM groups, which are typically lost as neutral fragments.

One of the most attractive features of MS/MS, is the diversity of experiments, or scanning modes, that can be conducted to gain structural insights into the parent ion and hence solve very specific biological questions or provide specificity when performing quantitative measurements.

Table 1. Summary of the main peptide/protein MS/MS fragmentation techniques.

Fragmentation technique	**Ions produced**	**Advantages**	**Disadvantages**
Collision-induced dissociation (CID)	b^{n+} and y^{n+}	High efficiency, high sequence coverage and robust	Labile groups, such as post-translational modifications (PTMs) typically lost
Electron-transfer dissociation (ETD)	c^{n+} and $z^{\bullet+}$ $c^{\bullet+}$ and z^{n}	PTMs retained and can be used localised.	Low efficiency, dependence on multiply charged species
Electron-capture dissociation (ECD)	c^{n+} and $z^{\bullet+}$ $c^{\bullet+}$ and z^{n+}	PTMs retained and can be used localised.	Low efficiency, dependence of multiply charged species
Surface-induced dissociation (SID)	b^{n+} and y^{n+}	High sequence coverage	Labile groups, such as post-translational modifications (PTMs) typically lost
Ultra-violet photodissociation (UVPD)	a^{n+} and x^{n+} b^{n+} and y^{n+} c^{n+} and $z^{\bullet+}$ $c^{\bullet+}$ and z^{n+}	Tune ability of fragmentation energies, diverse fragmentation pathways available	Complicated resulting fragment spectra and requirement of a chromophore

1.1.4.1 MS/MS scanning modes

The most common and relevant MS/MS modes are here described and illustrated in Figure 14 ;

Product ion scans, the precursor ion of a specific *m/z* ratio is selected during the first mass analysis stage of the experiment, this parent ion is then subsequently fragmented and all the product ions separated in a second mass analysis stage, before being detected.

Precursor ions scan, in contrast to a product ion scan, all precursor ions are allowed to traverse the first mass analyser and fragmented, but only a specific product ion is selected in the second mass analyser. Hence all precursor ions contributing to the production of the specific product ion are therefore detected. This type of experiment is only applicable for space-based mass spectrometers, such as multiple quadrupole instruments.

Neutral loss scan, both mass analysers are scanned simultaneously, with a constant mass offset, corresponding to the mass of a specific neutral loss fragment, applied to the scan of second mass analyser. Therefore only product ions which have undergone this neutral loss during fragmentation are detected. This type of experiment can be useful for determining the parent ion of a PTM-containing peptide, by selection of the appropriate neutral loss ion. Similarly this experiment is only applicable space-based mass spectrometers.

Selected reaction monitoring, both precursor and product ions, of a specific fragmentation pathway are specifically selected ion the first and second mass analysis stages respectively and no scanning of the instrumentation occurs. The removal of the requirement of time to scan through ions allows for more time to be spent on the detection of specific precursor and product ions. This combined with the removal of background ions increasing the sensitivity towards these selected ions, whilst maintain specificity of the analysis via the specific fragmentation pathways. These types of experiment are particularly useful for quantitative analysis.

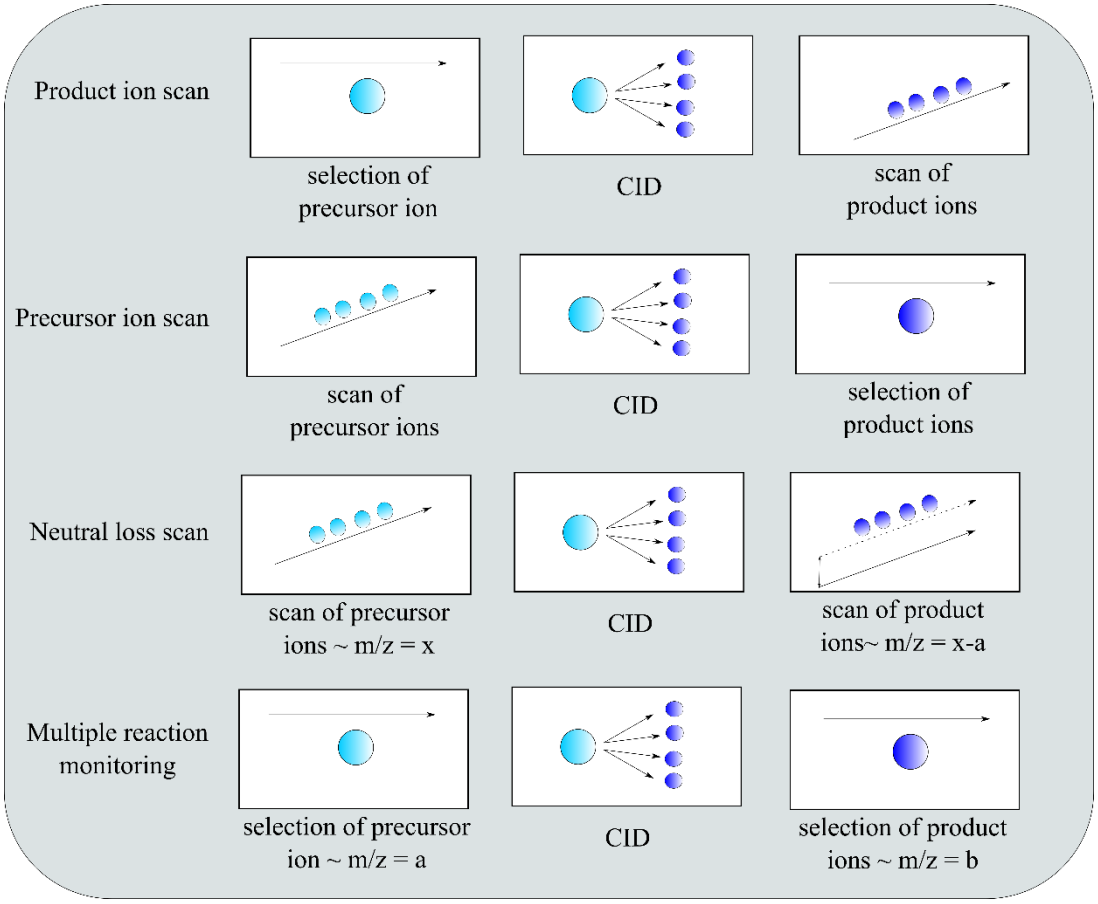


Figure 14. Tandem MS scanning modes. Collisional-induced dissociation of the precursor ions is indicated by CID, figure adapted from de Hoffman and Stroobart [22].

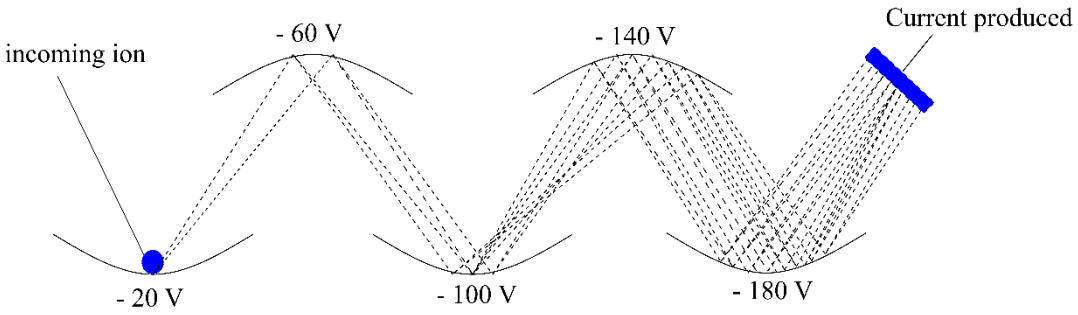


Figure 15. Schematic of ion detection in an electron multiplier detector.

1.1.5 Detectors

The final stage of an ions journey through a mass spectrometer, following production at the source and separation in the mass analyser, is its detection and transformation into a signal in the detector. Generally, detectors operate on the principle that an incident ion hits the detector, which generates secondary electrons, which are then amplified producing a signal of detectable current that is proportional to ion abundance.

There are several types of detectors including electron multiplier (EM), photomultiplier, Faraday cup and microchannel plate (MCP) type detectors. The Synapt G2Si instrumentation used in this thesis has a "Ultra-fast Electron multiplier and ''hybrid ADC'' type detector, for this reason only the principles of EM type detectors will be discussed. It is worth noting that the Orbitrap Q-Exacutive Plus instrument also used in this thesis, does not have a physical detector in the traditional sense, in contrast ions signals are inferred by the image current produced as ions within the mass analyser pass through a pair of photographic plates within the mass analyser and generate a circuit connecting the plates.

EMs consist of a series of dynodes, which are held at sequentially increasing potential. As an incident ion hits the first dynode, secondary electrons are produced, these electrons are then incident on next dynode forming further electrons and so on in a cascade effect, continuing on until the electrons reach the end of the detector. By which point, depending on dynamic range of the specific detector, the signal has been amplified by up to ~ 10^7 and these electrons are now detected as a current. Figure 15 gives a schematic of an EM type detector.

1.1.6 Ion mobility spectrometry (IMS) - MS

Ion mobility spectrometry (IMS) is an orthogonal analytical separation technique which allows for the separation of gaseous ions based on the size and shape of ion in addition to their charge, yielding increased analytical information of an ion.

In brief, during IMS, ions are pulled through a drift cell, that is filled with neutral buffer molecules, and the time taken for them to traverse the cell, or drift time (t_D) is measured. As a result of collisions with buffer gas molecules, ions exhibit different velocities through the drift cell. These velocities in turn can be related to their size and

shape via their collisional cross section (CCS) area, facilitating insights into the three dimensional shape of an ion.

The requirement of a gaseous analyte ion, enables the direct hyphenation of the technique with MS (IMS-MS), facilitating a powerful tool for obtaining both molecular and three dimensional information of analyte ion. The ability of IMS to separate out isobaric ions that differ in CCS, and applicability to a wide range of analyte molecular weight including ionic clusters [43], small organic molecules [44], peptides [45] and large protein complexes [46] , make the coupling to MS particularly attractive. The mode of separation can be used as an alternative to chromatographic separation or as an additional orthogonal separation step in the forms of LC-IMS-MS.

Several instrumental modes of IMS separation exist including drift time IMS (DT-IMS), differential mobility spectrometry (DMS), travelling wave IMS (TWIMS) and trapped IMS (TIMS), the two most common types of ion mobility devices, DT-IMS and TWIMS will be discussed herein.

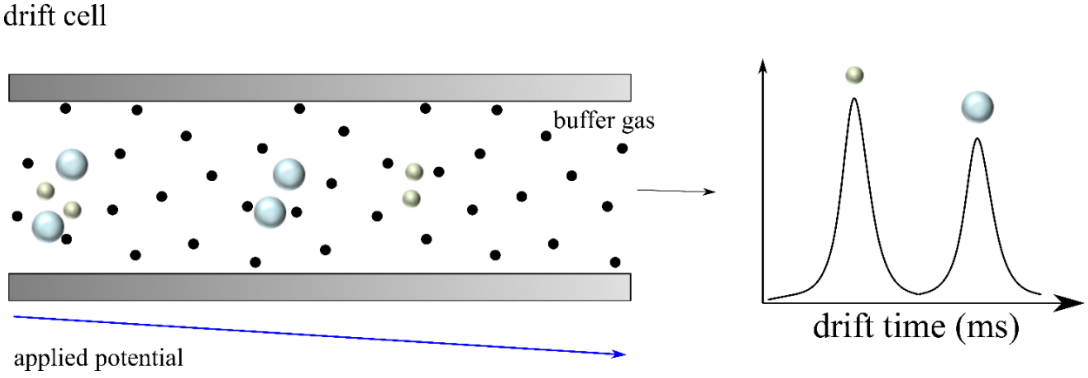


Figure 16. Schematic diagram of the separation of a large ion (blue) and a smaller compact ion (green) in a drift tube IMS (DT-IMS) instrument. A pulsed packet of ions are propelled by a low electrostatic field (5-100V) across the drift tube whilst encountering collisions with buffer gas ions. Larger ions will experience more collisions and drag resulting in subsequent decreased mobility and longer t_Ds across the cell.

1.1.6.1 Linear drift tube IMS-MS

Linear drift tube ion mobility spectrometry (DT-IMS) is the simplest form of IMS. In DT-IMS a packet of ions are pulsed into a gas-filled drift cell under a low field electrical potential, and the t_D measured. As the ions are injected into the cell they experience a constant electrostatic force pulling them through the cell, a force that is then counteracted by collisions with buffer gas molecules within the cell. Separation of two isobaric ions of differing size occurs as the larger ion with a greater CCS value will undergo more interactions with the buffer gas molecules, experiencing a greater reduction in drift velocity, v_d, and ultimately reduced t_D, than an smaller ion (Figure 16). The motion of the ion through the drift cell is termed the ions mobility, K, and is inversely proportional to the CCS of an ion [47, 48]. This relationship between K and CCS only exists at the low field limit (E/N), at higher fields, the mobility of the ion can become distorted as a result of alignment between the ion and the field.

It is worth noting that for a particular *m/z* ion, a single $_{t}D$ value is not measured during IMS experiments, but rather an arrival time distribution (ATD) which describes the t_Ds across the ensemble of isobaric ions pulsed into the cell, the shape and width of these ATDs can also be indicative of structure [49].

An ions mobility, K, can be expressed in terms of its velocity, v_D and the applied electric field, E.

Equation 15

$$K = \frac{v_D}{E}$$

If the drift cell length, L, is fixed and known and the t_D of an ion measured we can express mobility as;

Equation 16

$$K = \frac{L}{t_D E}$$

Temperature and pressure can influence mobility of an ion, hence generally mobility is expressed as reduced mobility, K_0, normalised for standard temperature (273.2 K)

and pressure (760 torr), where T and P are the gas temperature and pressures respectively.

Equation 17

$$K_0 = \frac{L}{t_D E}\frac{273.2}{T}\frac{P}{760}$$

The Mason-Schamp equation [48], demonstrates how mobility of an ion is inversely proportional to its CCS (Ω);

Equation 18

$$\Omega = \frac{3ze}{16N}\left(\frac{2\pi}{\mu k_B T}\right)^{1/2}\frac{1}{K_0}$$

where z is the charge of the analyte ion, e is the elementary charge (1.6022 x 10^{19} C), N is buffer gas density, μ is the reduced mass of the ion-neutral drift gas pair, k_B is the Boltzmann constant. Substituting reduced mobility and rearranging Equation 18 the in terms of CCS (Ω)we obtain;

Equation 19

$$\Omega = \frac{(18\pi)^{\frac{1}{2}}}{16}\frac{ze}{(k_B T)^{\frac{1}{2}}}\left[\frac{1}{m_i} + \frac{1}{m_n}\right]^{\frac{1}{2}}\frac{760}{P}\frac{T}{273.2}\frac{1}{N}\frac{t_D E}{L}$$

Where m_i and m_n are mass of the buffer gas and the ion respectively.

DT-IMS allows direct measurement of t_D, therefore under controlled conditions of temperature and pressure, CCS values can be directly abstracted from t_D measurements. Allowing structural insights into the gaseous ion which can be compared with molecular simulations and projections. However, DT-IMS based instruments have long since suffered from low duty cycle and poor transmission, and in recent years researchers have turned to TWIMS an alternative form of IMS separation in order to overcome these issues.

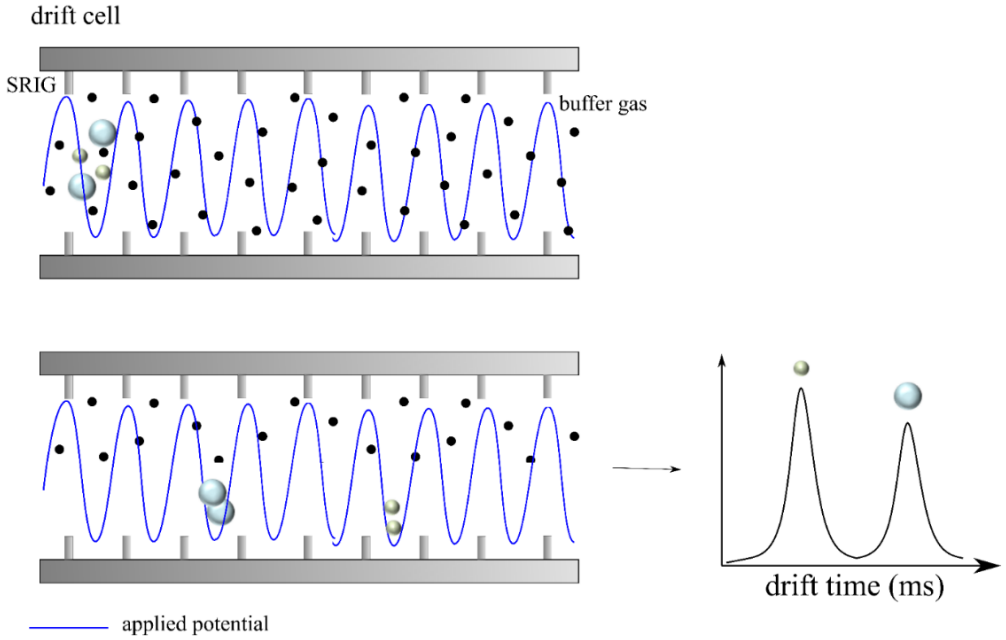


Figure 17. Schematic diagrams of travelling wave IMS (TWIMS) for ion mobility separation of a large ion (blue) and a smaller compact ion (green). The drift cell consists of a stacked ring ion guide (SRIG), this SRIG consists of a series of stacked ring electrodes with an applied confining radio-frequency field onto which a moving DC pulse is applied to consecutive electrode pairs. This DC pulse creates the travelling wave upon which the ions can surf. As ions are propelled by the wave they will experience an opposing drag force from collisions with the buffer gas and roll over the top of the wave. Ions with a large CCS (blue) will roll over the waves more frequently than low CCS (green) ions as they experience more drag force, resulting in mobility separation and elution from the tube later than their smaller counterparts.

1.1.6.2 Travelling wave-IMS (TWIMS)

First introduced in 2006 [50], the Synapt instrument series incorporated IMS separation in the form of a TWIMS device, situated between the quadrupole and ToF mass analysers, allowing for potential pre-IMS ion selection and activation. This region consists of three sections of stacked ring electrodes, or stack ring ion guides (SRIG), termed the 'trap', TWIMS and 'transfer' regions, known collectively as the 'Tri-wave', Figure 10.

Unlike DT-IMS where a uniform electric field is applied, TWIMS involves the application of a travelling wave of oscillating potential that propels the ions through the SRIG [51]. A confining RF-field is applied to this SRIG onto which a moving DC field is applied onto consecutive electrode pairs. This DC pulse creates the travelling wave upon which the ions can surf. The TWIMS drift region is filled with a neutral buffer gas, and as ions are propelled by the wave within this region they will experience an opposing drag force from collisions with the buffer gas and roll over the top of the wave. Ions with a large CCS will roll over the waves more frequently than low CCS ions as they experience more drag force, resulting in mobility separation and elution from the drift tube later than their smaller counterparts, as demonstrated in Figure 17

The main caveat of TWIMS is that the very nature of the travelling wave, and lack of uniform electric field alters the inverse relationship between t_D and CCS [48], meaning CCS measurements cannot be obtained directly from observed t_D. Instead an experimentally derived calibration curve must be generated, using calibrant ions with known CCS as measured by DT-IMS [52, 53].

This can be achieved by modifying Equation 19, for a non-uniform electric field;

Equation 20

$$\Omega = \frac{(18\pi)^{\frac{1}{2}}}{16}\frac{ze}{(k_B T)^{\frac{1}{2}}}\left[\frac{1}{m_i}+\frac{1}{m_n}\right]^{\frac{1}{2}}\frac{760}{P}\frac{T}{273.2}\frac{1}{N}At_D{}^{B}$$

Where A and B are correction factors for the electrical field and non-linear effects of the TWIMS device respectively [53]. A value of reduced CCS (**Ω**') can then be expressed as a charge and mass independent quantity, by dividing by reduced mass and charge and by grouping constants into a single constant term (A'):

Equation 21

$$\Omega' = A' {t_D}^B$$

This equation can then be plotted in the form $y=A't_D^B$, and a calibration curve generated using experimentally measured TWIMS t_D values of calibrant ions against know **Ω**' values from DT-IMS measurements [52]. This allows the subsequent abstraction of experimental **Ω**' for unknown analyte ions from TWIMS measurements, which in turn can then be back calculated to give CCS values using the reduced mass and charge:

Equation 22

$$\Omega = ze\left[\frac{1}{m_i} + \frac{1}{m_n}\right]^{\frac{1}{2}} \Omega'$$

As with DT-IMS measurements, comparisons can then be made between experimentally derived CCS values and simulated CCS values generated from potential 3D structures via a series of different projection algorithm[54–56].

Although DT-IMS instruments have superior IMS resolution of around 100 Ω/ ΔΩ compared to the 40Ω/ΔΩ of travelling wave instruments, the reduced ion transit time due to the faster mode of separation in TWIMS combined with the ability to store and pulse ions means an enhanced duty cycle resulting in higher sensitivity compared with traditional DT-IMS instruments.

1.2 MS and protein structural analysis

Following the introduction of MS theory, the many ways in which MS can be applied specifically for the analysis of protein structure will now be discussed.

1.2.1 Protein structure

Proteins play a pivotal role in many biological processes *in vivo*. Gaining an understanding of their role and function is key to the expansion of our knowledge of health and disease. To fully understand this, we must consider not just the presence, absence or amount of a protein in question but also the impact of the protein's higher order structures (HOS) and the role that these structures have on the function and activity of a protein. The different levels of protein HOS are illustrated in Figure 18.

A proteins primary structure, consisting of the amino acid sequence of the polypeptide chain, is the simplest form of its structure. This polypeptide chain can then fold locally, through the formation of hydrogen bonds across the amide backbone, forming localised three dimensional structures, such as α-helices, β-sheets, which are termed secondary structure. The presence of these localised structures causes the polypeptide chain to bend or twist supported by interactions between amino acid side chains. These interactions occur so as to achieve a maximum stability or energetic minimum of the structure, giving a resulting globular three dimensional shape to the protein that is termed the tertiary structure. If multiple subunits of proteins combine via non-covalent interactions to form a protein complex this structure is termed quaternary structure. Finally further complexity can be present in the form of post translation modifications (PTMs), and structural and conformational isomers. Variation in any of these including primary, secondary, tertiary and quaternary structures can have a large impact on the activity and function of a protein.

In the last 20 years, the application of the field of proteomics, which involves identification and quantification of proteins at a peptide level by MS, has revolutionised the way we perform clinical assays and understand complex pathways involved *in vivo*. These types of MS measurements have typically been based on the protein's primary structure and protein measurement has now become routine in many clinical and research contexts. However though now routine, global standardisation is required in order to maintain high levels of accuracy, precision and comparability of measurements with traditional clinical assays. Over the past decade the metrological community, through the development and use of reference materials (RM) has made significant progress on the development of methods for quantification of the protein primary sequences which are traceable to the System of International Units (SI) [57, 58]. Yet, the application of these metrological methodologies on the standardisation of higher order certified protein RM is lacking, due to the essential need to consider the role of PTMs and three dimensional protein structures on any accurate measurements made of these higher order RMs.

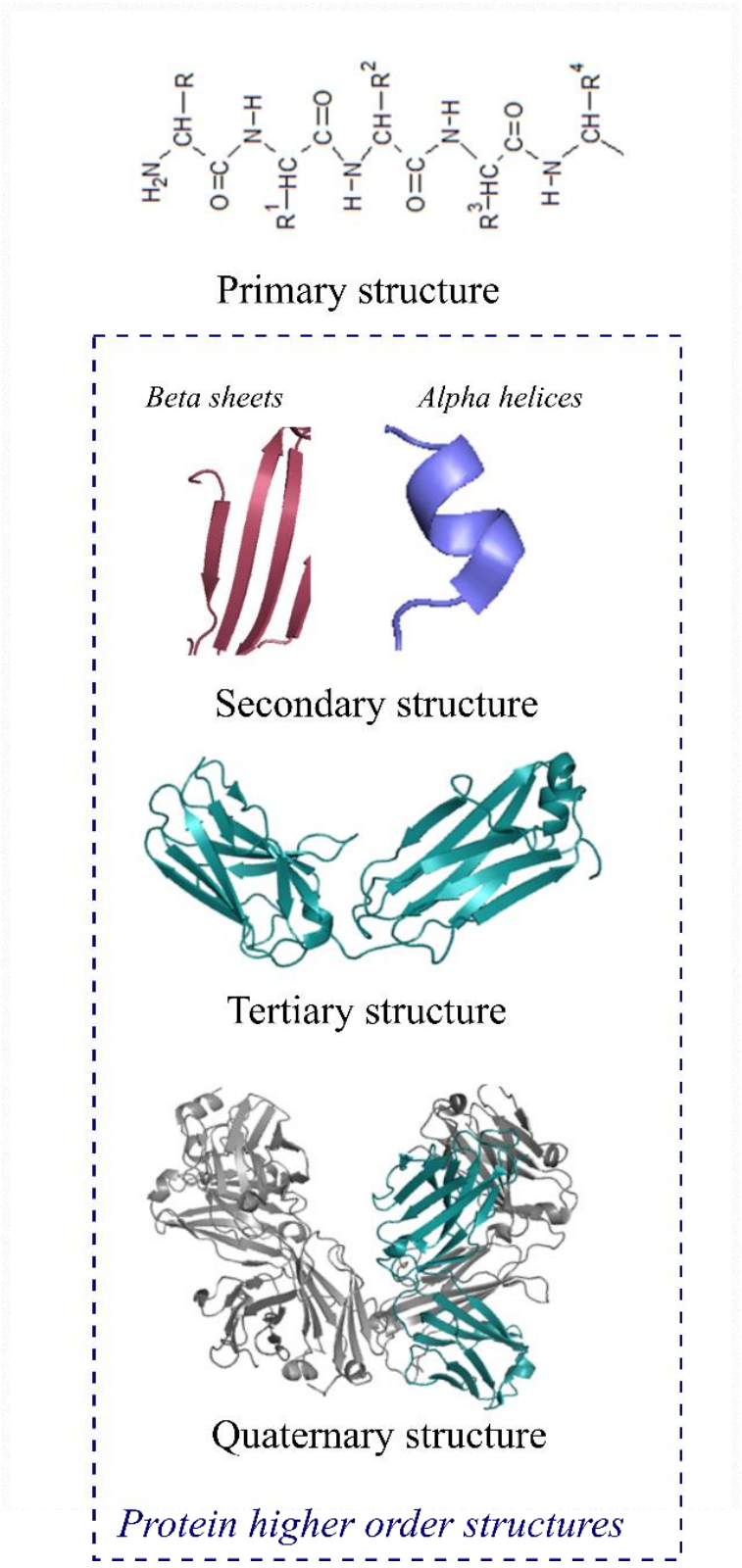


Figure 18. Schematic of the different levels of protein higher order structures.

1.2.1.1 Need for protein structural analysis and reference materials

This relationship between activity and protein HOS is important in the clinical and diagnostic areas, however, more recently, with the rapid expansion of drug therapeutics based on protein drugs (termed biotherapeutics), it has become particularly important in the pharmaceutical area [59].

In fact, HOS measurement has become a legislative requirement for the characterisation of biotherapeutics and biosimilars [60, 61]. Indeed regulation states that a series of physical, chemical and biological properties of a candidate drug, known as critical quality attributes (CQAs), must be monitored for and maintained within predefined limits, ranges or distributions to ensure product quality and established pharmacokinetics [60, 61]. CQAs include those properties or characteristics of a drug that may affect the identity, purity, biological activity and stability of a drug, and include such properties as glycosylation and protein HOS of a drug [62]. Protein HOS measurements provide important information on the stability, potency, efficacy, immunogenicity and bio-similarity of a bio-therapeutic. Indeed insights into the mechanisms of protein-protein interactions, [63, 64], protein-ligand interactions [65, 66] and protein aggregation[67], have all been provided by protein HOS measurements.

With these regulatory requirements comes a drive for the development of analytical tools capable of providing comprehensive HOS analysis and an understanding of the comparability of analytical measurements. Though there are a wide variety of techniques available for the analysis of protein HOS, little is known about the comparability of the data within like-for-like platforms and even less still about the comparability when comparing between different orthogonal platforms. In the clinical and diagnostic areas, it has been demonstrated that through the use of RM, and the understanding of measurement comparability, harmonisation of results can be achieved [68–70]. Indeed in these clinical and diagnostic areas, such standardisation is a recognised requirement. This standardisation is based on absolute quantities, therefore requiring SI traceability and harmonisation is based on equivalence of relative results. Yet, such international RM and model systems for harmonising and standardisation of protein HOS measurements have been lacking and as the area of protein HOS analysis expands so does the need for such materials.

1.2.1.2 Protein structural analysis techniques

For decades, nuclear magnetic resonance (NMR), x-ray crystallography (XRC) have stood as the gold standard analytical technologies for protein HOS measurements, facilitating atomic level high resolution measurements. For XRC, resolutions of < 1 Å, and 1.1-1.4 Å have been reported for small (10-15 kDa) [71] and large proteins (140 kDa) [72] respectively. More recently, experimental improvements in cryogenic electron microscopy (cryo-EM) have led to increased resolution ~ 1.8 Å and these measurements have been increasingly adopted in the area of structural biology [73].

There are several caveats to the analysis of protein HOS using these techniques including issues of crystallisation (~ XRC) and extensive sample preparation (~ cryo-EM) for dynamic proteins, lack of correlation between solid state structures and solution state structures (~ XRC/cryo-EM), poor signal to noise and upper mass limitations (~ NMR ~ 50 kDa). These techniques also suffer from heavy data analysis requirements and low sensitivity, resulting in a large sample requirement for high resolution analysis which in turn can also lead to aggregation issues for NMR in particular. In contrast these techniques which provide high resolution data, there are a wide range of biophysical analytical techniques that provide lower resolution analysis, allowing insights into overall global conformational changes and stabilities but with much higher throughput of analysis compared to high resolution measurements. These techniques include Fourier Transform infrared spectroscopy (FTIR), Raman spectroscopy, circular dichroism (CD), size exclusion chromatography (SEC) and differential scanning calorimetry (DSC). Each of these techniques can provide different structural information which can, when used in conjunction with each other, help to define HOS. These techniques however can often be limited in their sensitivity to small subtle structural fluctuations and lack the ability to localise the structural change.

Due to the regulatory requirements for HOS there has been a drive in the development of analytical techniques which provide higher resolution structural information whilst avoiding some of the limitations of the more established techniques of NMR and XRC. This has led to an increase in the research into the application of MS-based structural techniques, where sample requirements are much lower, there is often an increased tolerance to excipients and analysis can typically occur directly on solution

conformations and hence provide orthogonal data sets to traditional biophysical analysis.

1.2.1.3 MS for protein structural analysis: an overview

In short, MS for protein structural analysis can be split into two categories (Figure 19);

Native MS, where the protein's folded 'native' solution state conformation is maintained during gaseous analysis, and;

Protein footprinting, where the protein's folded solution state conformation is labelled in solution, and its subsequent analysis is performed in a denatured, often fragmented, state.

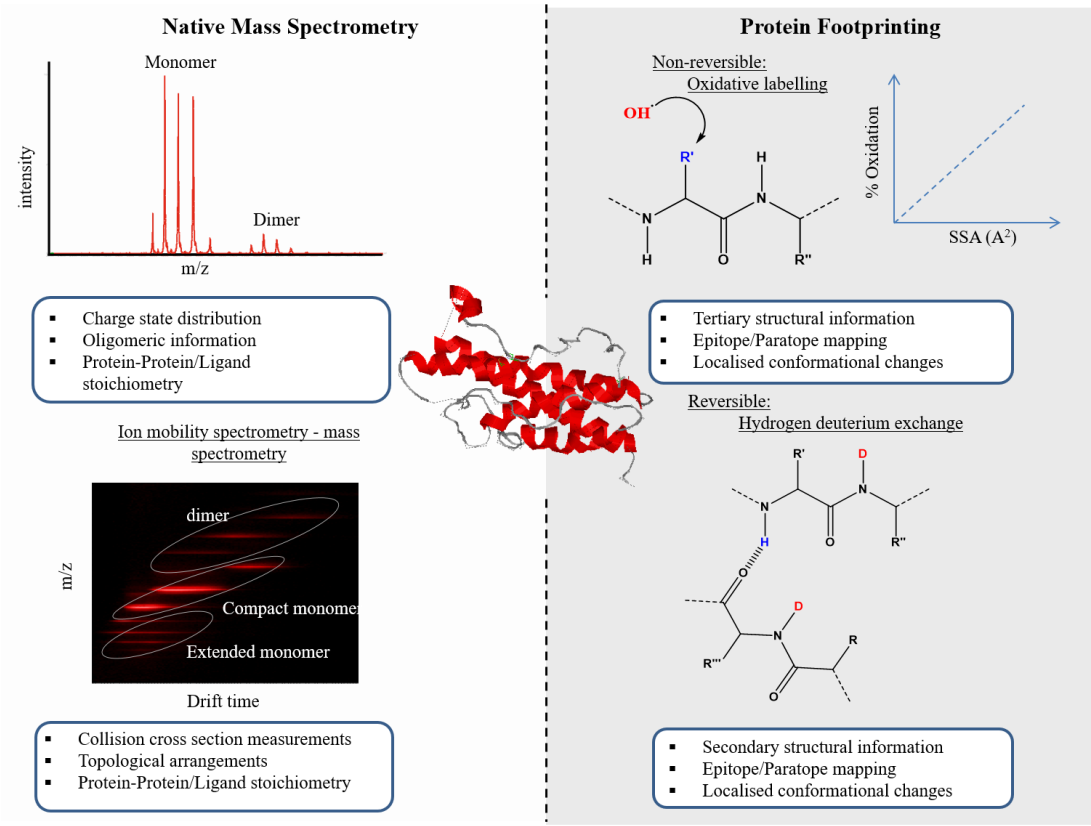


Figure 19. Schematic illustration of some of the range of structural information that can be achieved using a variety of MS techniques.

Classically native MS and protein footprinting can be thought of in terms of global and localised resolution for protein structural information respectively, though there are exceptions. Often providing complementary structural insights these methods can be applied in a standalone manner or in conjunction with one another. The rest of this

introduction will now discuss the theory and experimental considerations of these two methodologies in context of the data described in this thesis.

1.2.2 Native MS

1.2.2.1 History and overview

It was not long after the advent of ESI, in the late 1980s, that intact proteins were analysed by MS for the first time [74, 75]. As well as providing the exact intact mass measurement of the proteins primary sequence, it became apparent that structural insights of both protein and non-covalent protein complexes could be gained under controlled experimental conditions. Observations were made by Loo and co-workers [76], that stated that the charge state distribution (CSD) of a protein ion was greatly influenced by the solvent composition, which in turn related to the solution state structure of a protein during ionisation (discussed further in section 1.2.2.2). The mechanism for the production of a natively folded protein/complex during ESI is widely accepted as being that of the charge residue model (section 1.1.1)[77].

Native MS was expanded further by the development of nESI, with its higher ionisation efficiencies and lower capillary potential, facilitating reduced sample requirements and enhanced preservation of non-covalent complexes. Another driving force for native MS has been in MS instrumentation, such as the commercialisation of ToF based instrumentation in the 1990s, which were attractive due to their theoretical unlimited *m/z* range, high mass resolution and sensitivity. Increasingly instrumental modifications to Orbitrap [78] and FT-ICR type instrumentation have also increased their usage in this area.

Since then, native MS has been used to gain insights into virus assemblies (~ 18 MDa) [4, 79], aggregation [80], ligand-binding stoichiometry [63] and calculations of dissociation constants [81] (Figure 19).

It is important to note that when studying HOS by native MS, analysis occurs in the gas-phase and solution state conformations are inferred, hence great care must be taken so that protein solution conformations and any non-covalent interactions are retained in the gas phase, some of the experimental considerations that must be made in light of this will be discussed in the rest of this section.

1.2.2.2 Charge state distribution

As stated previously, it was observed that amount of charges or CSD accumulated by a protein during the ionisation process, correlated with the solution conformation, folded or unfolded, of the protein prior during ionisation [76]. Since then many additional studies [79, 82, 83] have provided information to support the idea that the number of protonation sites of a protein correlates with the solvent accessible surface area (SASA) of a protein [82].

For a more globular, folded protein, the extent of surface protonation is limited due to the burial of amino acidic side chain groups within the folded structure, leading to a narrow CSDs, low overall charge and higher *m/z* values (Figure 20). Equally the same protein, but in an unfolded conformation, will have more exposed amino acid side chains and hence a greater potential for surface protonation, resulting in higher charge states, greater CSD distribution and lower *m/z*. Increased coulombic repulsion as a result of increased surface charge is also thought to be a contributing factor towards a larger more unfolded protein ion at higher charge states. These ideas correlate well with the idea that a protein unfolds, its SASA increases and hence higher charge states are observed.

Though it provides structural insights, caution should be taken with interpreting CSD as can also been influenced by experimental parameters such as high instrumental voltages and solvent composition. Care should be taken to optimise these conditions to ensure the native conformation is retained.

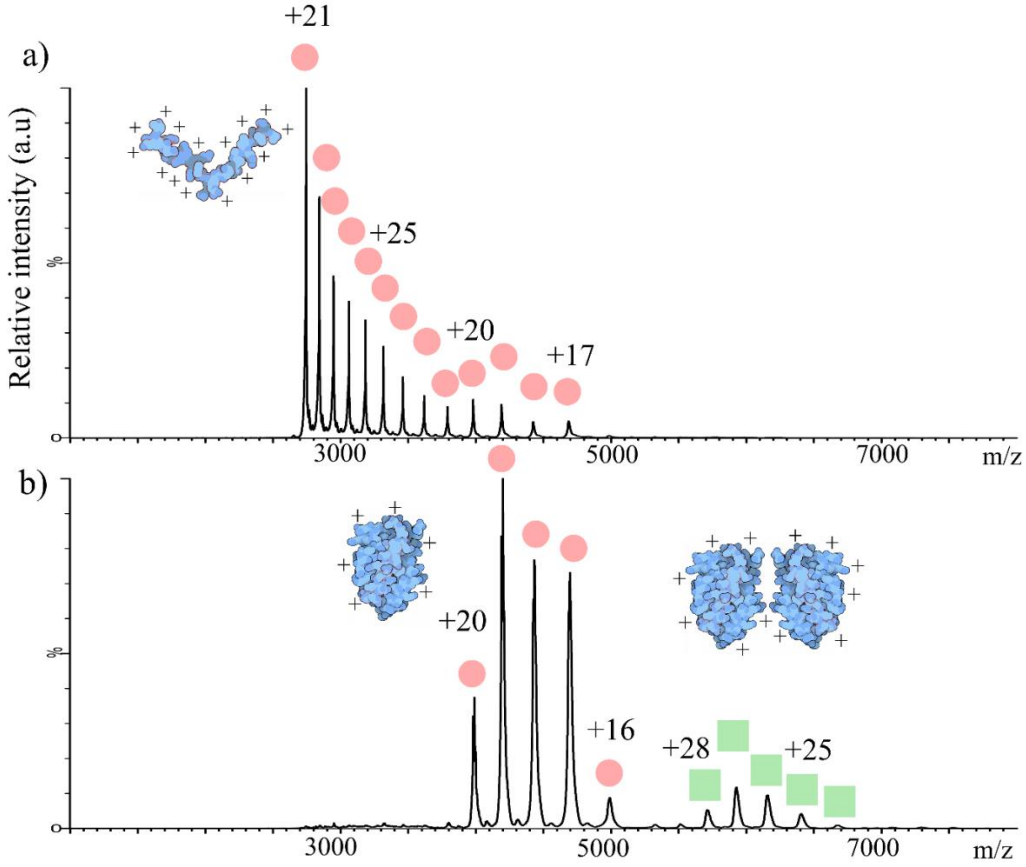


Figure 20. Illustration of the theoretical mass spectra of a protein in a denatured conformation and native conformation. Spectra represent a protein introduced from a solution containing a) 50% MeOH, giving a denatured, elongated protein conformation with higher charge, lower *m/z* and wide CSD, and b) buffering ammonium acetate, giving a native folded conformation, with lower charge, higher *m/z* and narrow CSD. For the spectra generated from native solution conditions both monomer (●)and dimeric (■) forms of the protein are observed. Protein illustrations are adapted from [84].

1.2.2.3 Transmission of higher *m/z* species

The theoretical unlimited mass range of some MS instrumentation is one of the most attractive features of the technique for structural biology, yet as research progressed to larger biomolecules it became apparent that ions ~ >100 kDa exhibited poor transmission [85]. These transmission losses occur at the small orifices between ion guides, due to the internal energy gains experienced by ions as they are accelerated from atmospheric pressure at the source into a vacuum. These energy gains can be dispersed both axially and radially and hence affect their transmission through these

small orifices. The internal energy gains are proportional to mass, meaning losses are more significant for larger molecular weight species and/or can result in the unfolding of higher *m/z* species. Using increased pressures in the source and ion guides of the instrument, can minimise these losses, by reduction of the analyte ion's internal energy through increased collisions with neutral gas molecules in a process termed collisional cooling [85]. Furthermore reduction in ion speed also allows for greater focusing of the ions and ultimately higher transmission. Therefore, modern MS instrumentation for native MS, such as the Synapt G2Si instrument used in this thesis, are modified to operate at these higher pressures to facilitate such collisional cooling and maintaining transmission.

1.2.2.4 Native IMS-MS

The coupling of IMS with native MS has been pivotal in understanding protein conformation that exists in the gas phase. Comparison of experimentally derived CCS values with theoretical CCS values from XRC/NMR, has allowed for many theories relating to the preservation of a solution phase structures into the gas phase to be validated [3, 5].

IMS-MS measurements can be used to understand solution state conformations (Figure 21 a), such as narrow CSDs being indicative of a single conformation, gas phase conformations such as understanding the influence of experimental conditions such as trapping voltages on the folding states (Figure 21 b and c).

Unlike many biophysical techniques for the analysis of proteins and protein complexes, IMS-MS is unique as it does not yield conformational averaged data, allowing for the true heterogeneity of a sample to be observed [3, 63]. Additionally the ms timescale of IMS-MS experiments, is much more applicable to the timescale of protein dynamics than those of NMR and XRC facilitating the analysis of dynamic nature of protein structures [86]. IMS-MS has been used not only to study heterogeneity but also protein-ligand interactions [65], protein-protein interactions[63], biosimilarity [49, 87, 88]) and aggregation assembly pathways [67] among others.

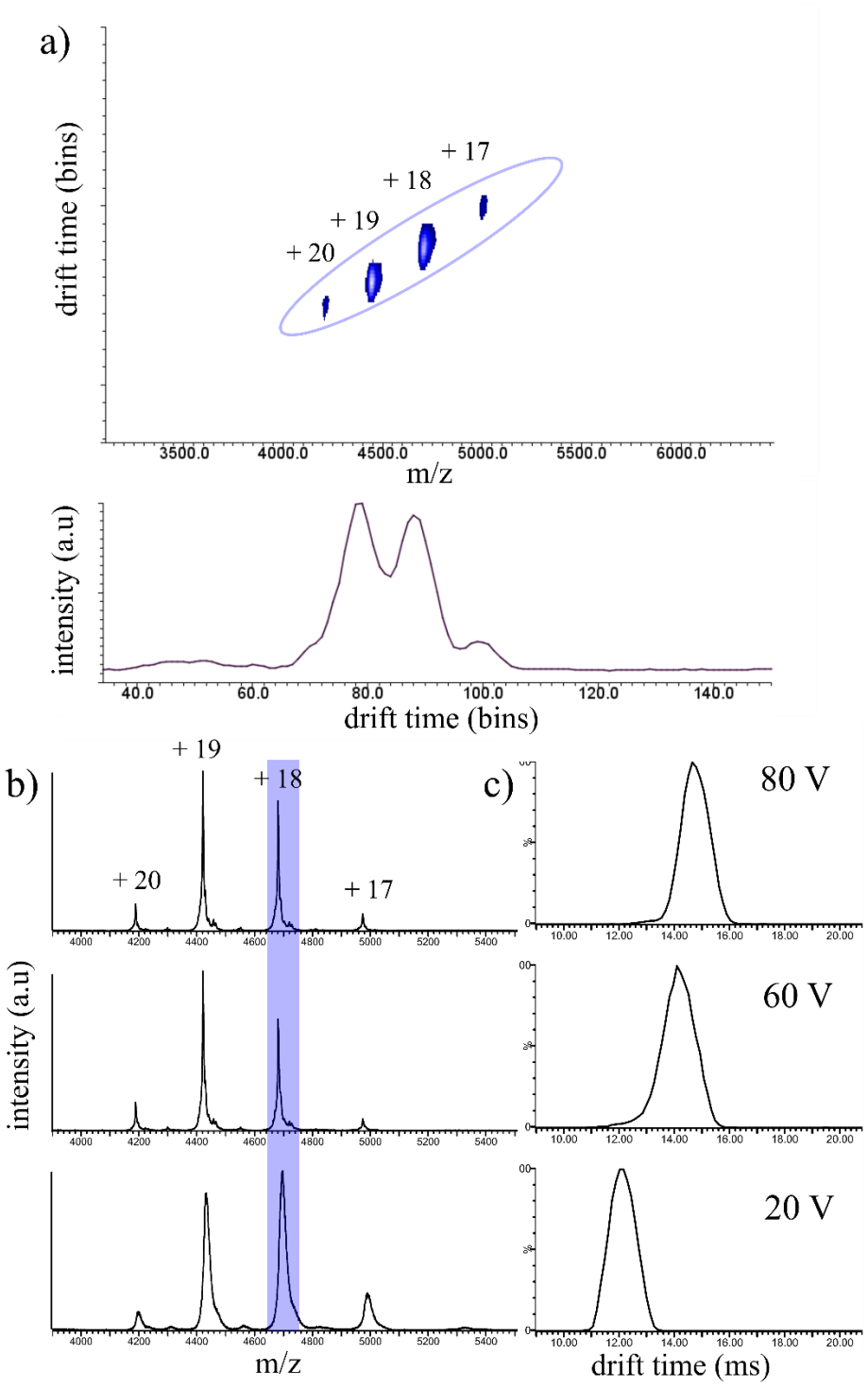


Figure 21. Structural insights of (holo-)transferrin from native IMS-MS measurements. a) Mobilogram spectra reveals a single monomeric conformation when using low energy internal trapping voltages (20V), b) increased trap voltages result in increased resolution of MS peaks however increased voltages also result in increased c) drift time indicating gas-phase unfolding of the protein.

1.2.2.4.1 Arrival time distribution (ATD)

As mentioned previously, IMS-MS measures the t_D of a population of ions, resulting in not a single t_D value, but an ATD which not only represents the t_D through the drift cell but any additional space travelled before hitting the detector. Often the centroid value of the ATD is used to determine a single CCS value, using the methods described in section 1.1.6, however the ATD profile itself can offer insights into protein conformation (Figure 22). If the relevant experimental controls are undertaken to ensure no gas-phase unfolding of the protein ions, wide ATDs can be indicative of the co-existence of several ensemble populations. Indeed the reporting and use of ATD peak width and shape as a metric for IMS-MS publication has been increasingly encouraged [49, 89].

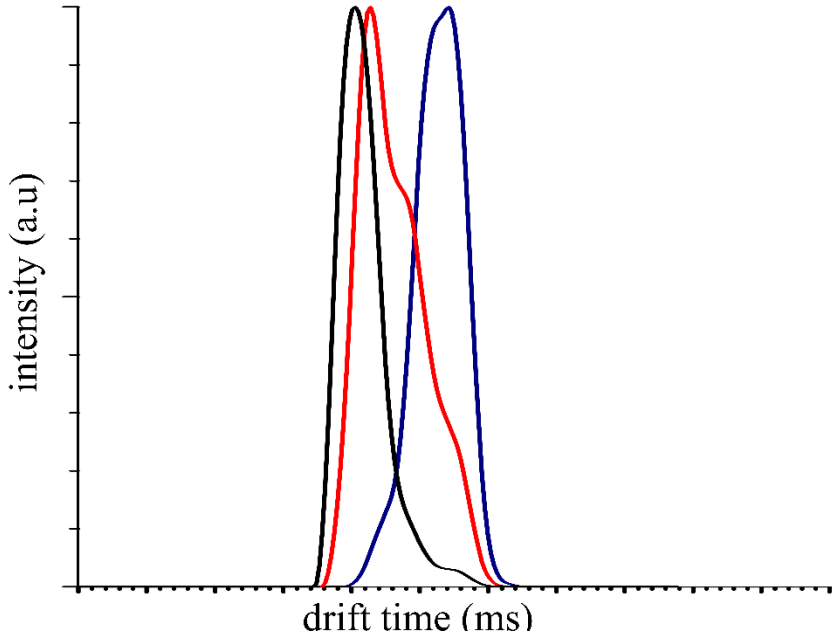


Figure 22. Illustration of a series of folding state as indicated by arrival time distributions (ATD) of the same *m/z* ion under differing experimental conditions. A single conformation is indicated by a Gaussian-like ATD (black trace), a ATD with an increased ATD indicates an increase in the CCS of the ion (blue trace) with a non-Gaussian-like shape and increased ATD width (red trace) indicative of multiple conformations of differing CCS existing for the ion.

1.2.2.4.2 Structural insights from CCS measurements

Classically, structural biology has relied heavily on XRC and NMR data to provide information on a protein's three dimensional structure; however, the ability of IMS to

provide both orthogonal data sets and insights into the potential heterogeneity of conformations, makes CCS an attractive measurement for comparison and validation of solution state conformations. Theoretical CCS values can be extrapolated from structures derived from XRC/NMR/cryo-EM measurements or indeed simulated structures when such experimentally derived structures are unavailable [3, 63] and structures discounted or validated by comparison with experimental measurements.

Several computational models are available for calculating theoretical CCS values, and some of the most established are summarised in Table 2, including some of their known advantages and biases. Good correlation between gas phase CCS measurements of small monomeric proteins with theoretical projections has been demonstrated [90, 91], however the systematic bias of the methods were clear emphasising the need to take care when interpreting CCS values obtained by IMS-MS.

These systematic biases are particularly important for CCS measurements extrapolated via calibration procedures from TWIMS measurements [52, 53]. Recently a concentration effect on ATD broadening was reported when deviating from published concentration values for calibrants [52], resulting in deviations in resulting measured CCS values [92].

Table 2. A selection of computational models available for the calculation of theoretical CCS measurements to be made from atomic co-ordinates obtained from NMR or XRC. These are all available from the open access programs MOBCAL [55, 56, 93, 94] each model provides a slightly different treatment of ion-buffer gas collisions.

Projection Algorithm	**PA**	**EHSS**	**TM**
Full title	Projection Algorithm [54]	Exact Hard Sphere Scattering Algorithm [55]	Trajectory Method [56]
Brief summary	Using fixed interaction radii for both buffer gas and ion average three dimensional CCS values are calculated from many two dimensional snapshots. *A modified update of this projection known as PSA (projection superposition algorithm) accounts for concave structures [94].	Atomic radii are modelled as overlapping hard spheres where the radii is equal to hard sphere collisional distances and scattering angles of the buffer gas upon collision.	Computation follows a similar approach to EHSS algorithm however it models interaction between buffer gas an ion using Lennard-Jones potentials.
Advantages	Least computationally demanding	Accounts for scattering, not very computationally demanding.	High accuracy model
Disadvantages	Fails for larger systems as long range buffer gas - ion interactions are not accounted for.	Scattering more significant for smaller proteins which can lead to over projection	Very computational demanding which can limit mass range of protein of interest
Applicable to	Small proteins	Large proteins	Small and large proteins

1.2.2.4.3 Collision induced unfolding (CIU) experiments

In addition to solely obtaining CCS measurements for proteins, collision induced unfolding (CIU) is an IMS-MS method which allows for additional insights into the stability and structure of proteins in the gas phase. CIU-IMS-MS is a collision activation method similar to that of tandem MS fragmentation, where an ion is first isolated and then activated post ionisation, however, instead of monitoring the fragmentation such as with tandem MS, in CIU-IMS-MS the unfolding pathway of an ion is monitored as function of collision voltage. These measurements can be indicative of protein structure classification [95], protein stability [96] and used for the characterisation of protein assemblies [3] and protein-ligand complexes[97, 98].

In CIU-IMS-MS, an isolated biomolecule is activated through energetic collisions with a background neutral gas (i.e. Argon), if caution is taken to keep the internal energy low enough to maintain covalent bonds within the structure, this increase in internal energy causes the biomolecule to unfold in the gas-phase. If the activation occurs before the drift cell, this unfolding process can be monitored using IMS-MS measurements, through increases in CCS or ATD measurement, as a function of activation energy, giving an insight into the stability of the biomolecule and/or specific domains of the structure (Figure 23). CIU type measurements can be thought of as analogous to melting point DSC measurements, though in the gas-phase, and can be used to generate fingerprinting data and allowing for high throughput direct comparison of protein stabilities and screening type experiments [99, 100].

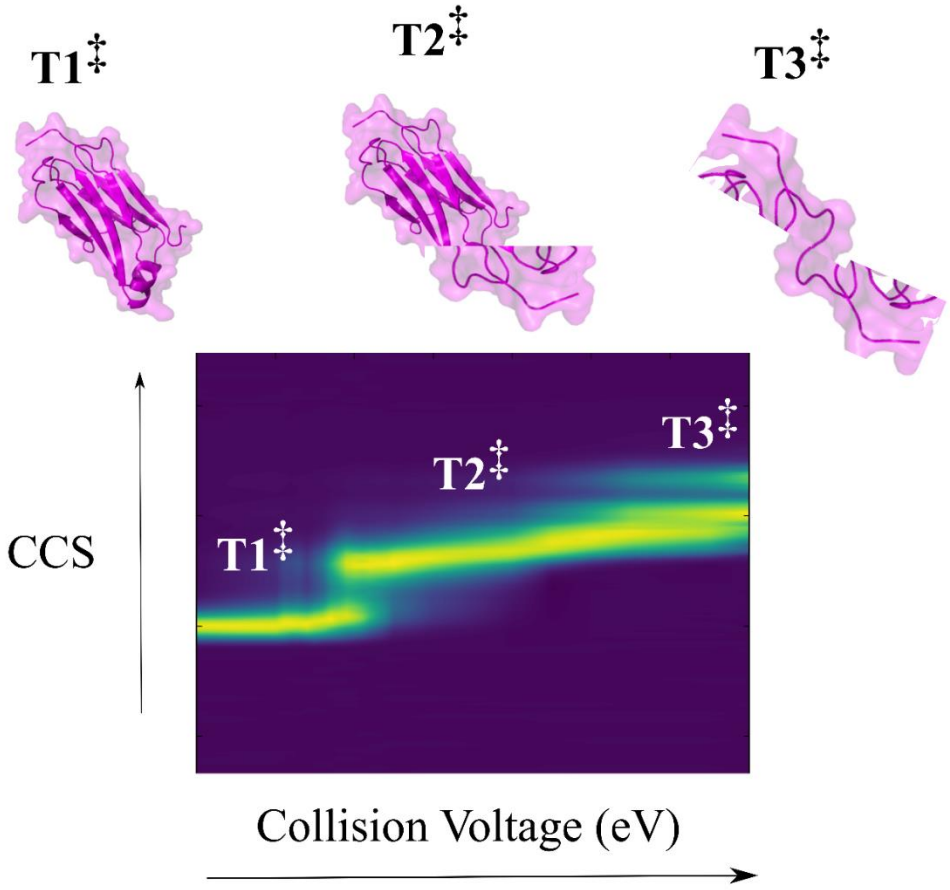


Figure 23. Schematic of collision induced unfolding (CIU) of an isolated biomolecule. Increasing the collision voltage, increased the internal energy of the native folded biomolecule (T1†), causing different regions and/or domains of the biomolecule to unfold giving rise to a partially folded state (T2†) and a fully-unfolded state (T3†), which are each indicated by their increase in CCS value.

1.2.3 Protein footprinting

In protein footprinting, chemical regents are used to label different features of a protein's primary sequence (Figure 19). The rate of these modifications are altered by the presence of HOS such as secondary [101], tertiary [102] and quaternary structures [103] or indeed ligand binding [104], thereby monitoring these labelling modification rates, can allow for HOS information to be inferred. Several different labelling strategies exist including hydrogen deuterium exchange (HDX) [105] hydroxyl-radical footprinting (HRPF)[106, 107], chemical labelling (using reagents such as carbenes and diethylpyrocarbonate) and cross-linking (XL) [108, 109].

Unlike native-MS techniques, where experimental considerations must be made to preserve a protein's solution state conformation into the gas-phase, protein footprinting techniques allow the structural behaviour of proteins to be labelled in solution prior to MS analysis, which subsequently occurs at either the top-down or bottom-up level. At the top-down level the accurate mass of the modified intact protein is monitored providing global conformational information. Bottom-up approaches rely on proteolytic digestion and subsequent LC-MS/MS analysis of labelled peptides to give high resolution localised information, where the labelling is located to either a region of the sequence or to a specific amino acid residue. This type of analysis can give insights into natively folded proteins or complexes [102, 103], define conformational changes [112] and mapping of inter/intra protein contacts [113].

This thesis will concentrate on just two types of footprinting techniques, one reversible labelling technique, HDX, probing the amide backbone structure of a protein and the other an irreversible labelling HRPF technique, fast photochemical oxidation of proteins (FPOP), probing the structure of amino acid side chain groups. These two strategies can be thought of as complementary providing orthogonal datasets for protein structure.

1.2.3.1 Hydrogen deuterium exchange (HDX)

1.2.3.1.1 History and overview

HDX was first developed by Linderstrøm-Lang and co-workers in the 1950s, when they observed exchange of labile protein hydrogens with tritium (^{3}H) when exposed to a tritium containing buffer [101, 114]. This exchange was measured, via density gradient tubes, as a function of increased mass of the intact protein. Using insulin as a model protein, Linderstrøm-Lang and co-workers first demonstrated reduced rate of amide hydrogen exchange as a result of the amide protons involvement in protein secondary structures [101, 114].

Measurement of hydrogen exchange by NMR, at the intact protein level, was later introduced in 1982, which allowed the monitoring of single-residue exchange rates gaining localised structural information for the first time [115, 116]. The introduction of an additional proteolysis step by Rosa and Richards [117], and combination with a high performance liquid chromatography (HPLC) separation step, facilitated localisation of the isotopic labels at the peptide level. With the advent of ESI a novel

workflow [118], was introduced which used deuterium (^{2}H, D), as an alternative labelling isotope, and the hyphenation of mass spectrometry for the analysis of labelled peptides. These experiments form the basis of HDX-MS experiments that we recognise today.

1.2.3.1.2 Principles of HDX

The fundamental basis of HDX is that when a protein is incubated in a deuterated buffer, labile hydrogens, those bonded to heteroatoms, (N, O and S), within the proteins sequence, readily undergo exchange with deuterons in solution. For protons present in amino acid side chains and N - /C-terminal groups, this rate of exchange is on ~ µs timescales, too fast for most current detection techniques. However the protons present in amide bonds of the polypeptide chain (Figure 24), exchange at rates spanning from ms - months [119] which are within the time scale of detection techniques. This rate of exchange is dependent on two factors:

The **intrinsic rate** of the amide proton exchange, of which there is a contribution from neighbouring residues [120] and experimental parameters such as pH [121], temperature [122] and salt effects[120, 123].

Protection of the amide proton**,** occurring as a result of the participation of the proton in hydrogen bonding (secondary structure) or as a result of reduced solvent accessibility due to protein folding (tertiary structure).

The intrinsic rate of exchange, k_{ex}, of an unprotected amide can be expressed as the sum of the rate of the acid, k_H, base, k_{OH}, and water, k_w -catalysed rates, as calculated in Equation 23 and illustrated in Figure 24.

Equation 23

$$k_{ex} = k_H [H^+] + k_{OH}[OH^-] + k_w$$

An equilibrium exists between these acid and base mechanisms, creating a pH minimum for HDX rates, which for polypeptides occurs in the pH 2.3-3 range (Figure 25). Here rates of exchange can be reduced by a magnitude of 10^8 from those observed at physiological pH [120], where the base catalysed exchange dominates. Temperature also affects HDX rates, with rates increasing around an order of magnitude with every increase of ~ 22 °C.

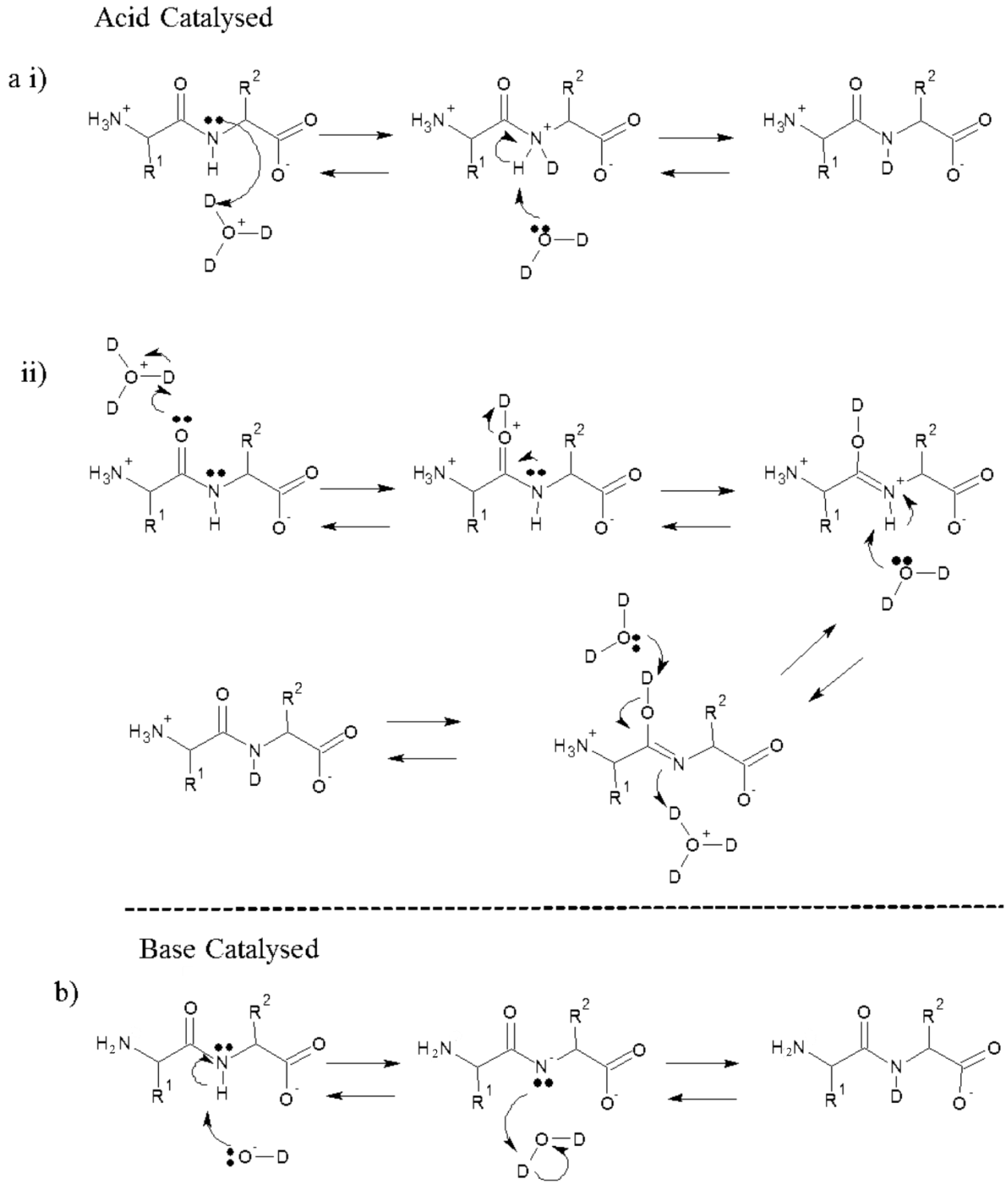


Figure 24. Acid (a) and base (b) catalysed mechanisms of HDX of amide hydrogens when incubated in deuterated buffer. Acid catalysis can proceed via N-protonation (i) or O-protonation (ii).

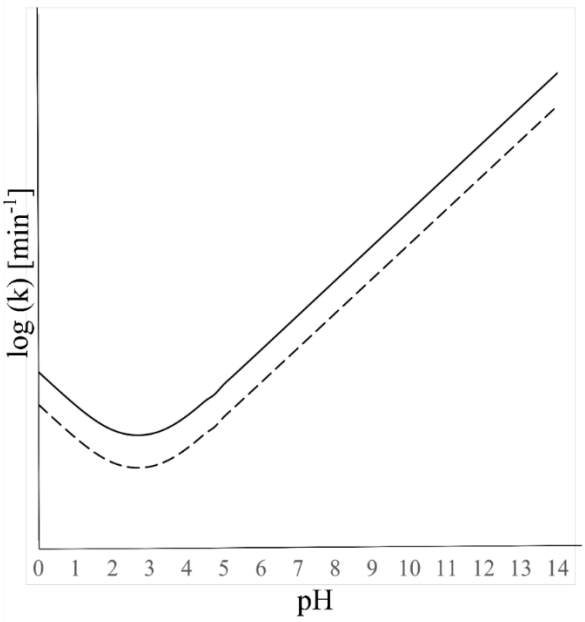


Figure 25. Dependence of intrinsic rate of exchange (k) of backbone amide protons on pH and temperature (solid line, 20 °C, dashed line 0 °C).

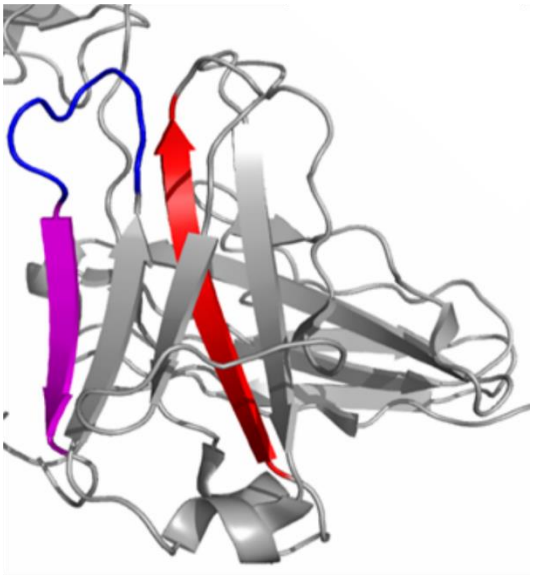


Figure 26. Representation of relative HDX rates due to the presence of protein higher order structure. HDX rates are fast for loop-like structures with no secondary structure and high solvent exposure (blue), rates are reduced due to the presence of hydrogen bonding such as β-sheets (pink) and/or burial within tertiary structures (red) leading to reduced solvent exposure. Illustration modified from PDB: 1HZH [124].

The strong pH and temperature dependence of exchange rates, requires rigorous control of these experimental parameters throughout the HDX experiment, complicating the experimental set up. Control of pH and temperature, however, provide an opportunity to take a snap-shot of the exchange present at a given time. This is typically achieved by "quenching" the labelling reaction by a reduction in pH within the pH 2.3-3.0 range and a reduction of temperature to ~ 0°C to give slow exchange conditions.

Finally, intrinsic rates of exchange can also be modified by both steric and inductive effects caused by neighbouring amino acid side chains in the absence of structure, by up to 10 fold [120]. These primary structure effects have been quantitatively evaluated by NMR, with the use of unstructured dipeptide standards, revealing the effect of close neighbouring amino acids to be additive, permitting prediction of intrinsic HDX rates as a function of primary sequence, pH and temperature[120].

Protection afforded by the folding of a proteins amino acid sequence into secondary and tertiary structures, also contributes to the observed exchange rate [125]. Those amides which are involved in hydrogen bonding or experience solvent shielding due to structural conformations display reduced exchange rates from intrinsic rates, with exchange rates potentially extended up to 8 orders of magnitude. The influence of protein structure on exchange rates, the presence of amide hydrogens presence across the entire length of the polypeptide chain (except at proline residues) and the technique's non-perturbative nature, combine to make HDX measurements a very sensitive structural probe (Figure 26).

HDX also allows insights into the dynamic nature of protein folding. In the Linderstøm-Lang model, for exchange to occur, a folded protein must repeatedly visit unfolded states in which the protected amides protons are free from hydrogen bonding/exposed to solvent and can undergo exchange[114, 119]. These unfolding fluctuations can be thought of as breathing like motions and are highly dynamic. For a stable folded protein in a single conformational state, the rate of refolding of the protein is greater than the rate of HDX, resulting in exchange kinetics that increase gradually with time, as amides are gradually exposed to deuteration, this instance is termed EX2 kinetics.

However if a protein exists in multiple conformational states and/or a region of the protein undergoes an unfolding event whilst in an open conformation during one of

these breathing motions, multiple amides are exposed to exchange and all the exposed amides exchange before refolding of the protein can occur. In this instance, exchange is observed as a bimodal distribution between folded/unlabelled and unfolded/labelled states of protein conformation and is termed EX1 kinetics. Most proteins at physiological conditions observe EX2 kinetics, though proteins such as intrinsically disordered proteins, that lack a stable conformational structure or which have been exposed to denaturants to artificially induce conformations observe EX1 kinetics [126].

1.2.3.1.3 MS for HDX measurement

The use of MS as a detection method for HDX experiments (HDX-MS), offers a number of advantages, over classical high resolution structural techniques of NMR and XRC. including sensitivity, protein solubility, speed of analysis, tolerance to excipients /impurities and accessible molecular mass range.

HDX-MS typically only requires nM-pM [127] protein amounts, in contrast to µM amounts required for HDX-NMR, which not only reduces sample requirements, particularly advantageous during research and development (R&D), but also avoids high concentration which may result in unwanted protein aggregation.

HDX-MS also has a higher tolerance to excipients compared with NMR and XRC, due to the opportunity of on-line desalting steps within HDX-MS workflows, allowing the technique to be directly applicable to formulated biotherapeutics and indeed study such effects of formulations on protein structure [128].

Additionally, there is a lack of mass limit of MS detection, although practically speaking the ability to separate, detect and identify the peptidic products of on-line digestion can become a limiting step for very large protein complexes [127]. Large proteins may require optimisation of chromatography steps to minimise impacts on data quality, see section 1.2.3.1.8 for more detail on this. Due to its applicability to large protein complexes, HDX-MS is used regularly to analyse large protein structures such as mAb ~ 150 kDa [129, 130] and viral capsid complexes in ~ MDa range [103, 131, 132]. These advantages of HDX-MS have resulted in an expansion of the application of the technique for protein structural analysis both in academia and industry.

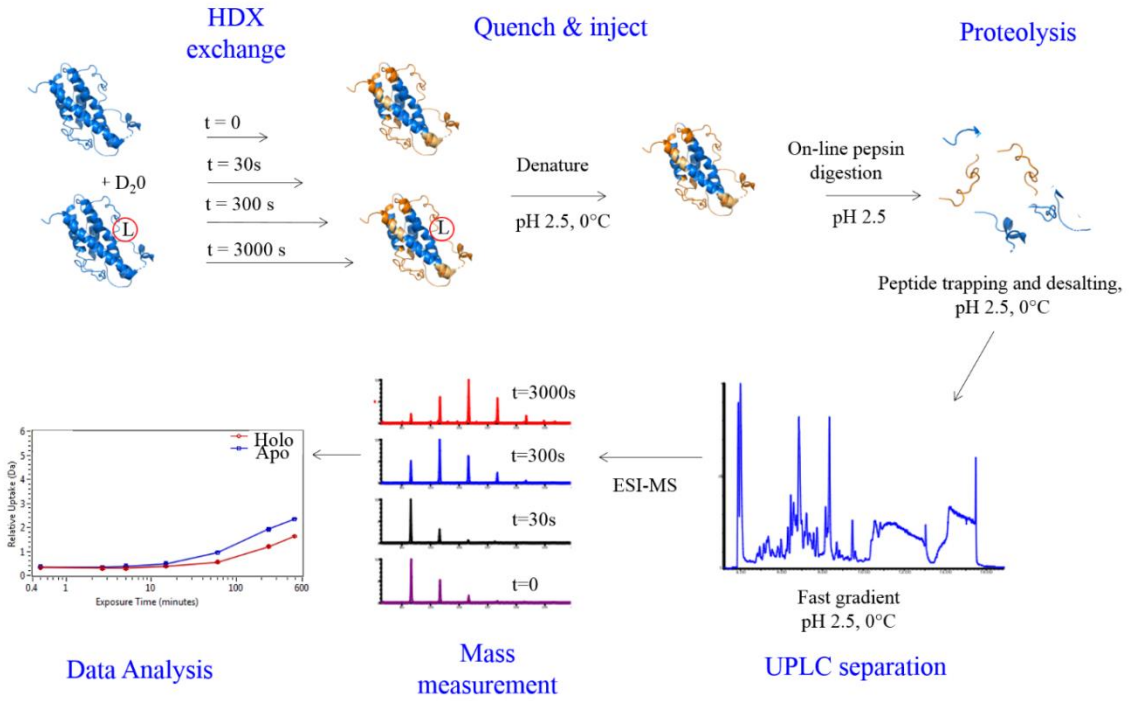


Figure 27. Schematic of a typical bottom up HDX-MS workflow for analysising protein dynamics upon ligand binding. Location of ligand binding is indicated by (L). The solubilised protein is diluted in an excess of deuterated buffer, where labile amide proton present in unstructured flexible regions exchange readily with deuterium atoms in the buffer (yellow). In contrast exchange occurs at much reduced rates for those amide protons involved in hydrogen bonding or protected by ligand binding or allosteric effects. Quenching of the labelling is induced (rapid decrease in pH, ~ 2.4-2.5, and temperature) losing the native conformation, however spatial information of secondary structure is maintained through the HDX labelling. The protein is then subjected to proteolysis using an acid stable protease and liquid chromatography (fast, 0 °C, often UPLC) mass spectrometry is applied to analyse the rate of deuterium exchange over time. Comparison of HDX data between the unbound (Apo) and bound (Holo) forms can be indicative of ligand binding effect

1.2.3.1.4 Back exchange

A major factor to consider during HDX-MS experiments is back exchange, which describes the process in which labelled labile deuterons, located in both the amide backbone and side chain, will readily exchange back with protons introduced during the quenching, digestion and LC steps, due to the reversible nature of HDX [125]. Back exchange can result in fast exchanging amide deuterons being incorrectly identified as un-exchanged hydrogens, therefore the process needs to be minimised in

order to preserve valuable structural information and the overall sensitivity of the technique. Correction factors have been suggested in order to account for back exchange[125], however these assume equivalent rates of exchange for all amino acids, an assumption which has been since shown to be false [133].

Minimising the amount of back exchange that occurs post-labelling, is typically achieved by reducing pH and temperature of HDX-MS workflows. These conditions represent a key analytical challenge and steps to overcome these challenges will be highlighted in the rest of this introduction.

1.2.3.1.5 HDX-MS workflows

HDX-MS experiments can adopt a wide variety of forms, labelling can occur in solution [117] or even in the solid state [134] and analysis can occur at the peptidic level (bottom-up), at the intact level (top-down) or even in cleaved domains (middle-down). A typical HDX-MS experiment involves the following steps; labelling, quenching, proteolysis (if bottom-up), desalting, chromatographic separation and MS analysis.

As it provides the highest resolution of measured exchange data, a bottom-up HDX-MS approach is the most common workflow, as illustrated Figure 27. In this example, two protein samples are compared, the unbound protein (Apo) and the ligand-bound protein (Holo). The samples, in parallel, are both incubated in a deuterated buffer for a range of defined time points (typically spanning from seconds to hours). Labelling is then quenched by reduction of the pH (~2.3-2.5) and temperature (0°C) to “slow exchange conditions”. Whilst maintaining the low pH to avoid back exchange, proteolytic digestion is then performed either using a solution based or on-line approach producing a range of peptides, which are then desalted and separated using a rapid chromatographic step before introduction into the mass spectrometer and consequently mass analysis. HDX-MS uptake plots are then generated by subtraction of the average (centroid) mass of the undeuterated peptide (control at t=0) and that of the deuterated peptide (for each incubation time) and plotting the difference as a function of time. At the peptide level (bottom-up) labelling sites can be localised to specific regions of the protein primary sequence and any differences in relative rates between the Apo and Holo forms of the protein can be attributed to changes in secondary structure or allosteric effects [135]. In addition, observation of reduced deuteration rates can identify ligand binding/epitope sites [111, 113].

Similarly the MS analysis can occur at the intact level (top-down approach) by removal of the proteolysis step in the workflow, the accurate mass of the intact protein is monitored providing lower resolution global conformational information. A comparison of the top-down and bottom-up workflows is illustrated in Figure 28.

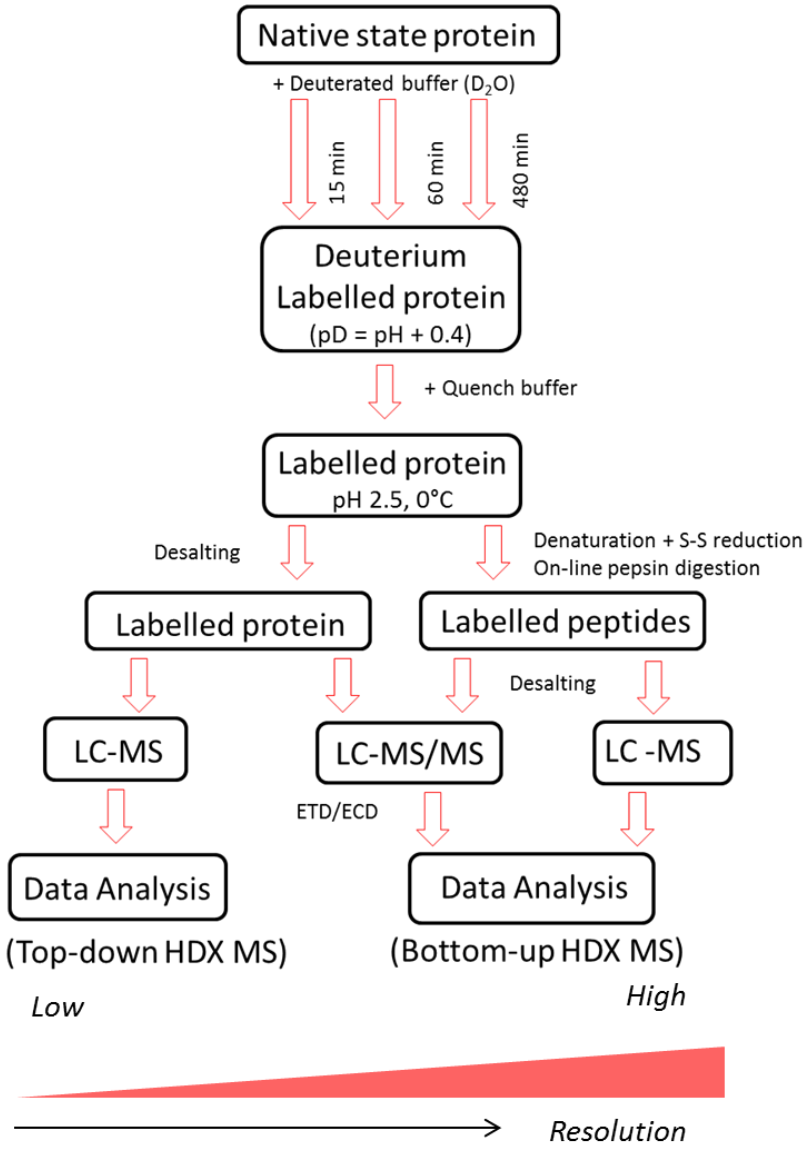


Figure 28. Schematic of top-down and bottom-up workflows for HDX-MS analysis following a continuous labelling experiment. Subsequent analysis of the HDX labelling can range from low resolution using intact analysis, medium resolution with the implementation of gas fragmentation strategies of the intact labelled protein and to high resolution analysis when performed at the peptide level.

1.2.3.1.6 Sample handling

Sample handling refers to the dilution, quenching and injection of deuterated protein solution onto the chromatographic system.

The dilution step, typically ~ 10-50 fold, of the protein sample in deuterated buffer initiating exchange can be performed for either a defined time interval (continuous, ~ 30sec to several hours) or for a short interval (pulsed, ~ tens of ms). Continuous labelling gives labelling incorporation as a function of time providing an insight into conformational dynamics in solution, whereas pulsed labelling allows access to short-lived conformations or structural intermediates [136]. Protein labelling is then quenched by reducing the rate of exchange by lowering the pH to 2.5 and the temperature to just above 0°C. It is often at this quenching stage that any denaturants or reducing agents are added to the sample to aid proteolytic digestion efficiency. Alternatively, if proteolytic digestion is not performed immediately following labelling, the samples can be quenched by flash-freezing in liquid nitrogen and stored at -80°C.

Performing these dilution and quenching steps manually as performed historically, was a bottleneck of the technique due to time constraints and questions of robustness/reproducibility relating to the operators technical ability. However the introduction of automated robotic sample handling has been instrumental in removing the requirement of the technical operator to be present, increasing experimental efficiency and leading to improvements in reproducibility due to the removal of operator errors and inconsistencies [112, 137–140].

1.2.3.1.7 Proteolysis

In bottom-up HDX measurements, the proteolysis step is key to establishing the resolution of measurements to specific regions of a protein's primary structure.

Traditional MS fragmentation strategies, such as CID, have prevented the localisation of HDX-MS rates at the single amino acid residue level, due to hydrogen scrambling (discussed further in section 1.2.3.1.9). Yet, through the generation of many peptides overlapping in primary sequence, known as redundant peptides, and subtractive analysis of their measurements the resolution of HDX-MS can be extended to just a few amino acid residues, or even to a single amide linkage [125]. These factors lead to the high dependency of HDX-MS resolution on effective non-specific digestion.

The “slow exchange” conditions (pH 2.5 and 0°C) required to quench labelling, however are generally not amenable to achieving high rates of digestion, especially not on the timescale required to minimise back exchange. Acid tolerant proteases including pepsin (EC. 3.4.23.1) [117], protease type XVIII from *Rhozpous* species (E.C. 3.4.23.6) [141], protease type XIII from *Aspergillus saitoi* species (E.C.23.18) [142] and Nepenthes fluid from *Nepenthaceae* species [143], to mention but a few, are exploited for their activity under these conditions, and their wide variety of enzymatic products. Traditional proteomic enzymes such as trypsin and Lys-C are inactive under these conditions.

Effective digestion during a HDX-MS experiment is reflected in the type of peptides produced. These peptides should be a mixture of small and long in length, reflect high sequence coverage of the protein’s primary sequence and also exhibit high redundancy [144]. Small peptides (~ 5-7 residues) are preferential, allowing for the accurate definition and localisation of exchange properties and structural interpretations that can be drawn to specific regions of the protein’s sequence. The relatively non-specific nature of pepsin in terms of amino acid cleavage site makes it an ideal choice for the production of such overlapping peptides [145], although the promiscuity of pepsin cleavage does complicate any subsequent data analysis steps significantly, requiring extensive computational analysis.

Digestion efficiency can be increased further still by the introduction of denaturing and reducing agents during the quenching stage, both of which can be optimised to the protein of interest [146], though care should be taken to desalt these additives prior to MS analysis due to suppression effects [147].

Often the digestion step is performed on-line using enzymes immobilised onto a column stationary phase [148] and digestion occurs as the quenched labelled solution is passed through the column. This approach has been shown to increase digestion efficiency [147], increase digestion reproducibility [127, 144] and prevent self-proteolysis of the enzyme, therefore eliminating self-pyrolysis products from downstream analysis. The on-line nature of the digestion also means that desalting of the peptidic products can also be performed on-line prior to the LC-MS stage with the use of a trap column further increasing the amenability to high through-put analysis. These advantages over in-solution digestion facilitate shorter digestion times, subsequently lower back exchange rates and reduced data complexity. For

these reasons most current bottom up HDX-MS workflows exploit the use of an on-line pepsin digestion columns [148], which can be produced in-house [149] but are also available commercially. However if using an on-line enzymatic digestion, care must be taken to minimise sample carry over [146, 150] as this can result in distortion of measured HDX rates, resulting inaccurate measurements and even incorrect kinetics identified [151, 152].

1.2.3.1.8 Chromatography

Generating a vast array of different peptidic products during the proteolysis step is essential for obtaining high resolution HDX-MS data, yet these peptidic products can significantly complicate the down-stream MS analysis. HDX-MS analysis using direct infusion of peptidic products has been carried out successfully for smaller proteins such as insulin ~ 6 kDa [153] yet for larger proteins a liquid chromatography separation step is required before MS analysis in order to reduce the spectral complexity and extend dynamic range of the experiment [64].

Aside from pH control (~ pH 2.5) there are two main strategies employed to minimize back-exchange; low temperature (~0^0C) and rapid chromatographic separation. When deuterated amides are exposed to protiated solvents under slow exchange conditions the half-life of exchange is reduced to the range $t_{1/2}$ ~ 30 to 120 min [125], thus requiring effective chromatographic separation in a matter of minutes to minimise deuterium losses. Achieving this with use of traditional reverse phase (RP)- HPLC has been challenging, often producing poor chromatographic separation, resulting in losses of MS sensitivity due to ionisation suppression effects, co-elution of peptides, increased spectral complexity and ultimately a lower confidence in peptide identification.

The introduction of ultra performance liquid chromatography (UPLC) systems which use sub 2 µm particle columns and high pressures (6000 to15000 psi) [127], has allowed for fast effective chromatographic separation under these challenging conditions in a matter of minutes (~ 1.5 to 6 min) [154], minimising back exchange. Typical back exchange rates observed for bottom-up HDX-MS approaches are often reported in the range of 10-30% [154–156]. The superior chromatographic separation achieved with UPLC over traditional HPLC methods has led to significant gains in

sensitivity reducing sample requirements 10-20 fold [127] and opening the door to the HDX-MS analysis of much larger and/or protein complexes [130, 131].

Further to gains observed by UPLC, use of IMS-based separation, prior to the MS analysis step (Figure 29) , has been illustrated to increase peak capacity [157, 158], remove issues of spectral overlap [159], increase peptide coverage and data quality of HDX-MS measurements, without the addition of analysis time or contribution to back-exchange.

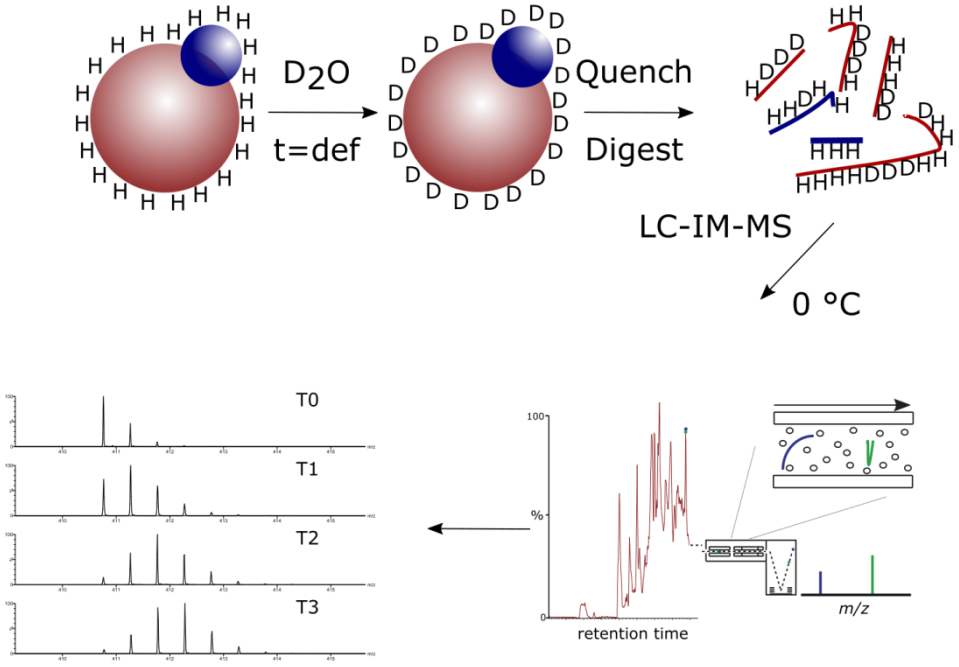


Figure 29. Schematic of the integration of IMS separation within HDX-MS workflow, figure taken from [159].

1.2.3.1.9 MS measurement and identification

Prior to any labelling experiments, the initial stage of HDX-MS bottom up workflow, involves repeat injections/digest of an undeuterated protein sample and MS/MS for peptide identification. This process is termed peptide mapping, and forms the basis of any subsequent analysis of HDX labelled products.

As with most proteomic workflows, peptide identification is most typically achieved using CID fragmentation and resulting spectral data is then exported to processing software such as PEAKS [160], MASCOT [161] or MS-vendor specific software such as PLGS (Waters, Milford, USA), where peptides are identified based on analysis of

fragment ions and database comparison. The identified peptides of the peptide map are then exported as a list of *m/z* and retention times onto which HDX-MS data can be aligned based on mass envelopes and retention times.

The fast elution times and co-elution of peptides observed using UPLC conditions required for HDX-MS measurements, can result in poor MS/MS using data dependent analysis (DDA) based measurements, due to slow scanning speeds and ultimately a loss of sensitivity towards low abundanct peptides.

However, MS^E is a scanning mode, available for the Synapt G2Si instrument used in this thesis, which allows for faster MS/MS scanning speeds (Figure 30) [162]. When in this mode the instrument switches between low and higher (or elevated as denotes the ^{E}) ramped collision energies, within the collision cell. The low energy scan, acts as a precursor scan and the high energy scan, causes fragmentation and act as the resultant product ion scan. By setting the high energy scan at a voltage range (~ kV) rather than a specific single voltage, means peptide fragmentation occurs irrespective of their intensity or *m/z*, facilitating very fast scanning speeds. However, as no precursor ion selection has occurred there is no instrumental correlation between product and precursor ion. Instead this correlation is established post acquisition based on the overlap of chromatographic elution profile between precursor and product ion using PLGS software. This type of ion scan is used for peptide mapping experiments described in HDX-MS experiments performed in this thesis.

Though useful at the identification stage, due to its high efficiency, traditional CID fragmentation cannot be used to localise deuterium labelling to specific amide linkages. The high internal energy of CID, leads to the gas phase interchange of labile protons/deuterons at exchangeable sites and loss of amide specific site information in a process called hydrogen scrambling [163], meaning the deuterium can only be located to the specific sequence of the peptide.

Site specific labelling information can be achieved using alternative fragmentation techniques such as ETD [164] and ECD [165], as their reduced internal energies do not result in hydrogen scrambling. However both fragmentation techniques suffer from low fragmentation efficiencies, lack of software support and increased practicality issues [166] limiting their utilisation into widespread HDX-MS workflows.

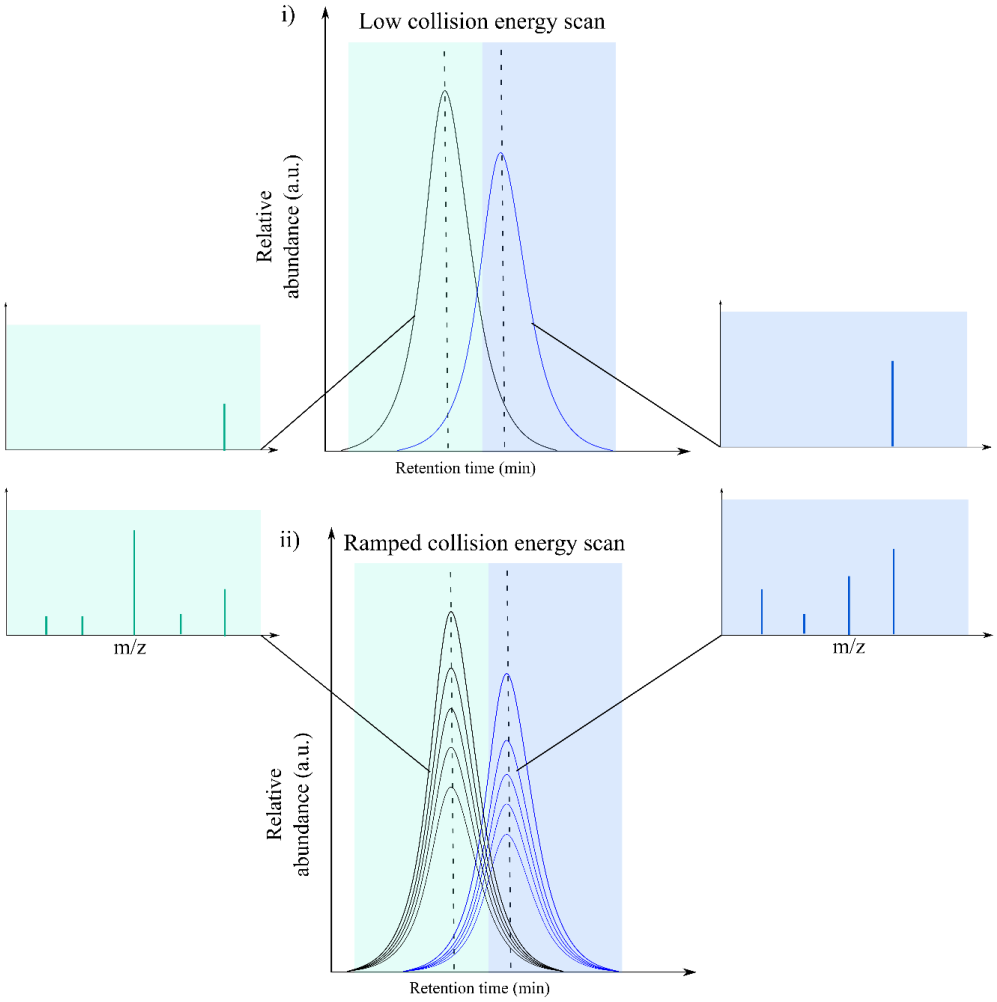


Figure 30. Schematic of MS^E scanning modes. A full product ion scan is continuously acquired whilst switching between i) low collision energy mode resulting in a precursor ion measurement and ii) a ramped higher collision energy mode producing fragment ion. Alignment between precursor and fragment ions is achieved using retention time post-acquisition, allowing for spectra from co-eluting analytes to be separated.

1.2.3.1.10 Data analysis

Due to the abundance of heavy isotopes such as ^{13}C and ^{14}N in organic ions, peptide ions are not represented by a single *m/z* peak, but instead consist of a series of peaks, or envelopes (Figure 31) , representing the range of masses (across all isotopes). The centroid mass, or the average mass across all isotopes present, is an intensity weighted mass measurement based on the *m/z* (m_i) for the given spectral peak (i) within the

envelope, the intensity of the spectra peak (I_i) and total intensity of all the ions present in the envelope give the centroid mass, calculated from;

Equation 24

$$centroid\ mass = \frac{\sum m_i I_i}{\sum I_i}$$

Due to the Δ mass difference of ~ 1 Da between hydrogen and deuterium isotopes, the uptake of deuterium increases the average mass of a peptide, shifting the isotopic envelope to higher *m/z*. HDX-MS exchange rates are therefore measured by the average mass of deuterated peptides as a function of deuteration time. This is typically calculated by subtraction of the average mass of the undeuterated peptide at time, t=ref from the mass of a the peptide exposed to the deuterated solvent at t=exch [125, 167].

Due to the long list of potential variables within a HDX-MS experiment including time, pH, temperature, protein concentration, mobile phases and chromatography, comparison of protein states (non-perturbed protein state vs perturbed protein state) by HDX-MS, is often performed relatively, termed differential HDX-MS [112, 137, 168, 169]. This essentially negates the impact of day-to-day variation of the variables listed above on HDX-MS measurement. However, when performing a straightforward two state protein comparison it is possible to produce hundred/thousands of individual measurements when performing just a few replicates at each time point [112, 170].

Performing this type of data analysis manually is extremely labour-intensive and for a long time a lack of software available for HDX-MS data analysis was a limiting factor in industrial usage. In addition to this, the complexity of data can often lead to practical protein size constraints. Counter to this, however, in the last decade or so, there has been an increase in the software available for semi-automated analysis, such as HDX Workbench [171, 172], MS Tools [173], DynamX (Waters, Milford, USA), HDExaminer (Sierra Analytics, CA, USA), HX-Express2 [174] and EXMS [175] to highlight a few. This automated software still requires specialist analyst knowledge. The review by Iacob and Engen [176] gives a comprehensive list of available software which is not within the scope of this report. The combination of good peak picking algorithms and high quality data enables high accuracy and precise measurements to be obtained in a higher throughput manner.

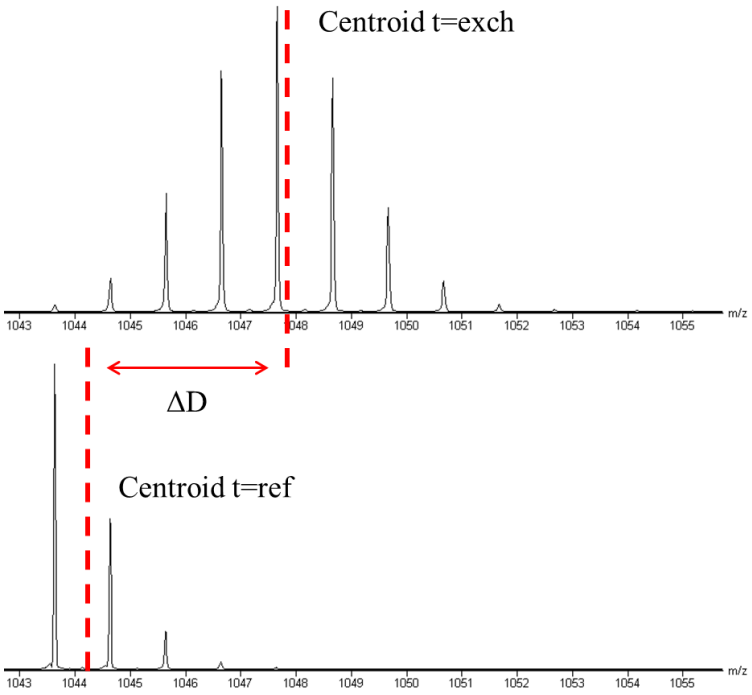


Figure 31. Centroid mass measurement of spectral envelopes gives the average centre of mass for analytes across all isotopes present. By calculating the area within the spectral peaks using *m/z* and intensity measurements the centre of mass can be calculated for each spectral envelope. Deuterium uptake is then calculated by measuring the difference in centroid mass between the undeuterated peptide mass (t = ref) and the deuterated peptide mass (t = exch).

1.2.3.2 Fast photochemical oxidation of proteins (FPOP)

An alternative footprinting technology to HDX-MS, oxidative labelling offers an alternative analytical method for measuring protein HOS. The theory, workflows and experimental considerations of the technique will now be discussed in detail.

Oxidation of proteins can occur both *in vivo*, during aging or various disease states, and *in vitro* during sample preparation and processing causing product degradation. It has long since been a common amino acidic modification studied in proteomic workflows. Furthermore oxidative labelling of proteins and protein complexes, in combination with MS analysis of the products, has been exploited as a means for proving protein structure [102] and protein-ligand interactions [177], via hydroxyl radical protein footprinting (HRPF) techniques (Figure 19).

1.2.3.2.1. History and overview

Oxidative footprinting was first developed in 1986 by Tullius and Dombroski [178], as a footprinting technique for DNA, whereby hydroxyl radicals generated using Fe^{2+} EDTA-Fenton chemistry, were used for studying the interactions between proteins bound to DNA. Cleavage of nucleic acids by hydroxyl radicals was observed to occur independently of base sequence. This cleavage, however, was mitigated by the interaction between that particular nucleic acid with a bound protein, allowing for the localisation of protein interaction sites.

The first reporting of protein footprinting however was not until 1988, by Sheshberadaran and Paye [179], who used hydroxyl radicals to non-specifically cleave the amide backbone of proteins, when studying the interactions between of antigen-mAb complexes, and used SDS-PAGE in order to separate out the proteolysis fragments. The coupling of the technique with LC-MS/MS based detection was performed by Chance and co-workers [180], dramatically increasing the resolution and sensitivity of HRPF.

A variety of hydroxyl radical sources including water radiolysis [181], and more recently laser photolysis [182], have since been applied, facilitating the protein labelling on ns - s timescale. Pioneered by Gross and co-workers [182, 183], fast photochemical oxidation of proteins (FPOP) utilises laser production of hydroxyl radicals on a ns - µs timescale, measuring oxidative footprinting of amino acid side chains as a function of solvent accessibility and protein HOS. The irreversible nature of the HRPF, in contrast to HDX, allows for extensive sample handling post-labelling, simplifying analysis and making the whole process readily amenable to standard bottom-up MS approaches.

Since its development, FPOP has been applied, but not limited to, the study of protein folding [184, 185], membrane proteins [186], structure elucidation [187] and epitope mapping [188], plus the analysis of isolated proteins and protein complexes or even *in cell* and *in vivo* [189].

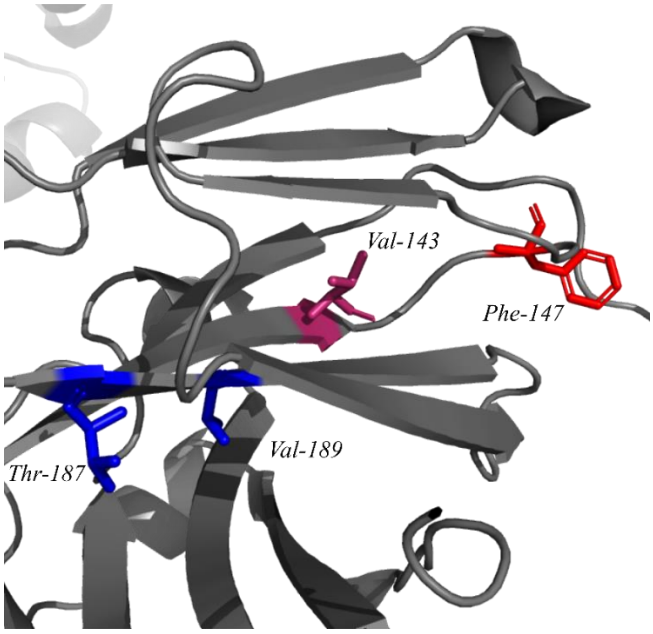


Figure 32. Representation of potential oxidation sites for hydroxyl radical footprinting as function of solvent accessibility, intrinsic reactivity and higher order structural effects . Amino acid side chains buried within tertiary structure with poor solvent exposure (blue) are unlikely to undergo oxidation compared to their solvent exposed counterparts. Aromatic side chains such a phenylalanine (red) when devoid of structure have higher intrinsic reactivity towards oxidation than other amino acid side chains, such as valine (pink) which will affect the oxidation rates observed. Illustration modified from PDB: 1HZH [124].

1.2.3.2.2 Principles of oxidative labelling

Oxidative labelling is an attractive chemical labelling technique, due to the high reactivity and relatively non-specific nature of hydroxyl radical labelling [190]. Though exposure to hydroxyl radicals can result in protein backbone cleavage, the reaction of amino acidic side chains is in the order of 10-1000 faster than this cleavage, so by tuning hydroxyl radical generation appropriately, amino acidic oxidation is the main product of HRPF. By monitoring oxidation rates as a function of the corresponding oxidative products, structural insights can be gained. Though oxidative labelling is a direct indication of a residue's solvent exposure, lack of labelling can also indirectly give structural insights. To understand this fully, it should be understood that HRPF oxidation of amino acids is dependent on three factors, solvent accessibility, intrinsic rates and structural effects (Figure 32):

Solvent accessibility, hydroxyl radicals share a similar van-der-Waals radius to that of water, meaning a quantitative relationship exists between SASA and the extent of oxidative labelling of amino acids [191, 192]. In the context of protein HOS, residues that are solvent exposed, undergo oxidation, in contrast, lack of oxidation or non-reactive side chains, are indicative of areas that are buried or involved in protein-protein or protein-ligand interactions and hence protected from modification [193].

Intrinsic reactivity, with 14 out 20 amino acids undergoing oxidation when exposed to hydroxyl radicals, HRPF is often thought of as non-specific, offering structural probes over the vast majority of a protein primary sequence. However, in reality, the intrinsic reactivity of the amino acids is not equivalent, with the rate constants for free amino acids in the range of 10^7 - 10^{10} M^{-1} s^{-1} [191]. Table 3, describes rate constants of free amino acids with hydroxyl radicals, and their chemical reactivity is as follows: Cys > Met > Trp > Tyr > Phe > Cysteine > His > Leu ~ Ile > Arg ~ Lys~ Val > Ser ~ Thr ~ Pro > Gln ~ Glu > Asp ~ Asn >Ala > Gly. In typical HRPF those amino acids with reactivity ≤ Gln, essentially remain unoxidised. This difference in intrinsic reactivity is an important factor to consider when inferring structural elucidations of HRPF data, and care should be taken not to over interpret data.

Primary and higher order structural effects, oxidation rates of amino acid residues can be influenced by the protein's primary sequence, due to electron donating or electron withdrawing group effects [194]. In addition, local tertiary structures surrounding an amino acid residue can influence oxidation rates due to charge-charge interactions, charge-dipole interactions, dipole-dipole interactions, and/or hydrophobic effects [195].

Table 3. Rate constants for the reaction of amino acids with hydroxyl radicals, including the oxidation products observed, table modified with data reported in Xu et al. [191].

Amino Acid substrate	**rate (M^{-1} s^{-1})**	**pH**	**Oxidative products**
Cys	$3.5\ 10^{10}$	7.0	Cys sulfonic acid (+48), sulfinic acid (+32), hydroxy (-16)
Trp	$1.3\ 10^{10}$	6.5-8.5	Trp hydroxy- (+16, +32, +48, etc.), pyrrol ring-open (+32, etc.)
Tyr	$1.3\ 10^{10}$	7.0	Tyr hydroxy- (+16, +32, etc.)
Met	$8.5\ 10^{9}$	6-7 4.5	Met sulfoxide (+16), sulfone (+32), aldehyde (-32)
Phe	$6.9\ 10^{9}$	7.0-8.0	Phe hydroxy- (+16, +32, etc.)
His	$4.8\ 10^{9}$	7.5	His oxo- (+16), ring-open (-22, -10, +5)
Arg	$3.5\ 10^{9}$	6.5-7.5	Arg deguanidination (-43), hydroxy- (+16), carbonyl (+14)
Cystine	$2.1\ 10^{9}$	6.5	cysteine sulfonic acid, sulfinic acid
Ile	$1.8\ 10^{9}$	6.6	Ile hydroxy- (+16), carbonyl (+14)
Leu	$1.7\ 10^{9}$	~6	Leu hydroxy- (+16), carbonyl (+14)
Val	$8.5\ 10^{8}$	6.9	Val hydroxy- (+16), carbonyl (+14)
Pro	$6.5\ 10^{8}$	6.8	Pro hydroxy- (+16), carbonyl (+14)
Gln	$5.4\ 10^{8}$	6.0	Gln hydroxy- (+16), carbonyl (+14)
Thr	$5.1\ 10^{8}$	6.6	Thr hydroxy- (+16), carbonyl (-2- or +16-H_2O)
Lys	$3.5\ 10^{8}$	6.6	Lys hydroxy- (+16), carbonyl (+14)
Ser	$3.2\ 10^{8}$	~6	Scr hydroxy- (+16), carbonyl (-2- or +16-H_2O)
Glu	$2.3\ 10^{8}$	6.5	Glu decarboxylation (-30), hydroxy- (+16), carbonyl (+14)
Ala	$7.7\ 10^{7}$	5.8	Ala hydroxy- (+16)
Asp	$7.5\ 10^{7}$	6.9	Asp decarboxylation (-30), hydroxy- (+16)
Asn	$4.9\ 10^{7}$	6.6	Asn hydroxy- (+16)
Gly	$1.7\ 10^{7}$	5.9	-

1.2.3.2.3 Oxidative products

Oxidative labelling can be identified and localised by characteristic mass shifts, which in turn can be relatively quantified by LC-MS/MS workflows (described in section 1.2.3.2.8). However a vast array of oxidative products exist, with the 14 amino acids that are oxidised, having the potential to yield > 50 different oxidative products. An extensive summary of the reaction mechanisms of all the different oxidation pathways is given in the review by Xu and Chance [191] which have been summarised into characteristic modifications/mass shifts in Table 3.

Most commonly, for aliphatic amino acids (Ala, Val, Leu, Ile, Pro), oxidation results in either the addition of a hydroxyl (+15.99 Da) or carbonyl (+13.98 Da) group to the amino acid side chain. For aromatic amino acids (Phe, Tyr, Trp) addition of one or multiple hydroxyl (+15.99 Da) groups to the aromatic ring can occur in ortho-, para- and meta- ring positions of the aromatic ring. Asp and Glu undergo oxidative decarboxylation (-30 Da), basic amino acid side groups give rise to a series of unique oxidative products. Sulphur containing residues, Cys and Met, can undergo multiple additions of hydroxyl groups (~ +15.99 Da, +31.98 Da) and due to high reaction rates, caution should be taken to ensure secondary oxidation has not occurred.

1.2.3.2.4 Oxidative radical sources

Hydroxyl radicals can be produced from a variety of sources including Fenton-Fenton reactions via metal catalysis [179], radiolysis of water by either γ or x-rays including synchrotron sources [185, 192] and laser photolysis of H_2O_2 [182]. However, only laser photolysis of H_2O_2 and x-ray radiolysis of water produces radicals on a fast enough times scale (ms-µs) for unbiased protein measurements. This times scale is faster than unfolding of both tertiary and secondary structures [196], allowing access to transient structures permitting possible exploration of dynamic conformations [197]. Figure 33, gives the reaction scheme for hydroxyl radical production from laser photolysis. In contrast to the expensive and limited access to sources of synchrotron radiation, FPOP requires only a modest experimental set up consisting of a UV 285nm KrF excimer laser, a syringe pump and a flow system [182], as illustrated in Figure 34. The experimental workflows will now be discussed in more detail.

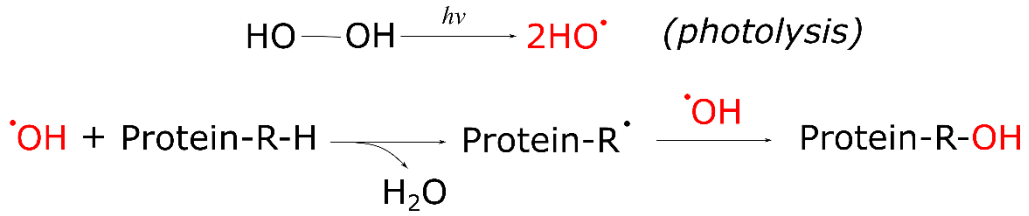


Figure 33. Reaction scheme for the modification of amino acid side chain groups with hydroxyl radicals generated by laser photolysis.

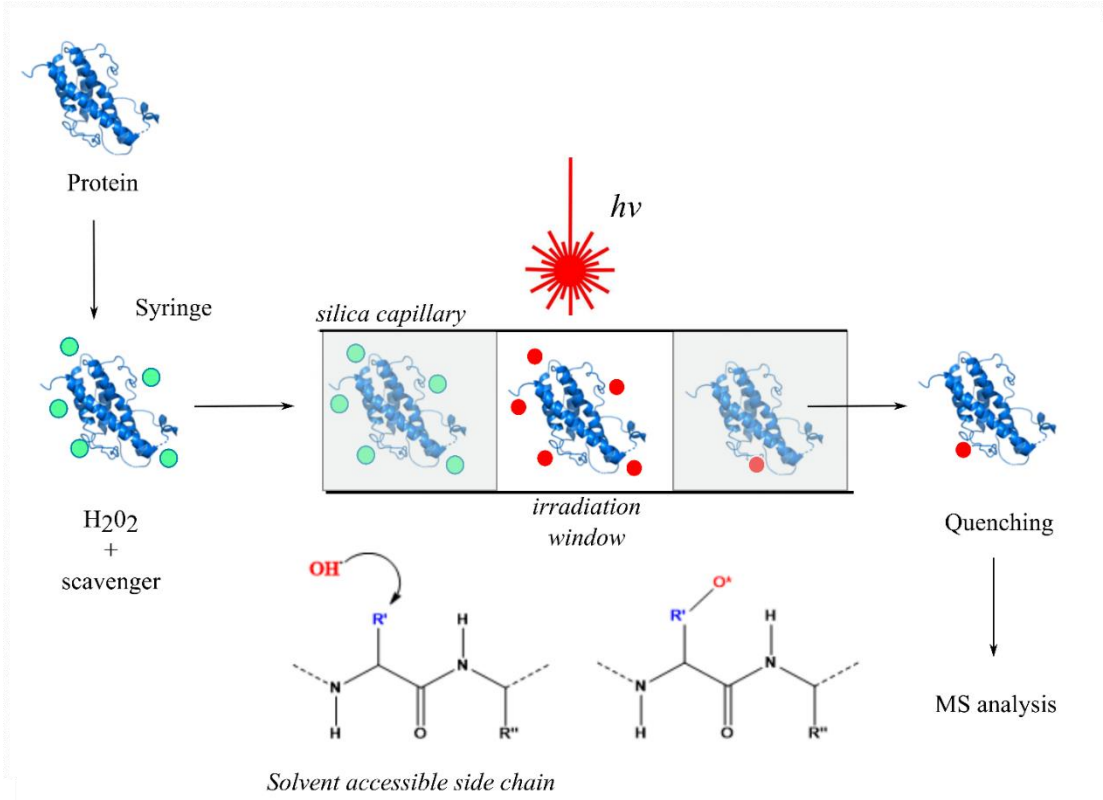


Figure 34. Schematic of hydroxyl radical oxidation of proteins using a capillary flow FPOP labelling workflow. The protein sample is incubated in low levels of H_2O_2 then flowed through a capillary system, where the sample is exposed to a single exposure to a KrF laser irradiation through a window on the capillary, producing hydroxyl radicals from H_2O_2. Solvent accessible side chains will be susceptible to hydroxyl radical attack and oxidation. The sample is then quenched with the addition of the enzyme catalase, and then can be frozen/stored prior to further sample treatment and oxidative MS analysis.

1.2.3.2.5 FPOP labelling

Conventional FPOP set-ups use a KrF laser, to produce ns pulses of 248 nm UV light, a wavelength sufficient to generate high yields of hydroxyl radicals yet is absorbed minimally by proteins in aqueous buffers. The laser is directed to a narrow window a fused silica capillary via a series of lenses or optics which, focusing the beam to a tiny spot ~ 2 mm x 1mm, covering the height of a conventional capillary (100-400 µm outer diameter). Figure 34 gives a schematic of the FPOP workflow described. The sample (protein + H_2O_2 + scavenger) is transported via a sample flow system, consisting of a syringe pump, delivering a constant flow of the sample (~ µL/min) and a capillary of fused silica tubing. This tubing is then held into careful alignment with the laser beam via a mounting system and a transparent window formed at the point of contact using an ethanol burner. During the FPOP experiment, the system is then irradiated once through this window, in small segments as a result of the laser pulse, meaning a small volume of sample in between laser pulses, known as the exclusion volume remains irradiated [107]. This exclusion volume, typically ~ 30% , can be calculated by capillary dimensions, laser frequency and flow rates to ensure of single shot exposure of sample.

Two specific features of the FPOP set-up ensure that the short lifetime and exposure of protein to hydroxyl radicals occurs in the fast sub-ms scale, these are use of scavenger reagents and ensuring single irradiation exposure, allowing the FPOP footprinting to capture rapid structural dynamics. The generation of hydroxyl radicals through a narrow window, ensures the packet of protein passing through the window is only exposed once on a ns time scale and removes secondary oxidisation events. Secondly the use of scavenger solutions, typically in the forms of free amino acids, allows the tuning of hydroxyl radical life time on the 1 ms -100 ns times scale[198].

1.2.3.2.6 Sample requirements and handling

In order to abstract meaningful structural conclusions from oxidatively labelled proteins, as with all MS structural methodologies, it is critical that oxidation must occur while the protein is in its native conformation that is being probed. Controlling the extent of oxidation is also critical for data quality and assurance, too low and it risks being below the MS limit of detection and quantification, however to high and it risks saturation of the modification which can affect the quantification of the unmodified forms. Key steps to control these levels of oxidation, other than the

experimental set ups described include sample preparation, the FPOP experimental process and the sample collection steps.

A typical FPOP sample contains µM levels of protein (owing to the high sensitivity of MS analysis) plus any ligands of interest, any sort of buffering requirements for the protein of interest and a tuned H_2O_2/scavenger ratio. It is worth noting that additives within buffers can diminish oxidation induced by the FPOP, particularly free amino acids often present in protein formulations or reducing substances such as dithiotreitol (DTT), all of which can have quenching effects on hydroxyl radicals produced. Where possible additives such as these should be kept to a minimum. Levels of H_2O_2 should be kept ~ 15-20 mM range to ensure sufficient oxidation [199] and radical lifetime is controlled using an optimised scavenger solution, which competes with the protein for oxidation with the hydroxyl radical. Glutamine, histidine and methionine are commonly used as scavengers for FPOP experiments [182, 200]. Measurements by time-resolved [181] and lifetime profiles [182] UV spectrometry, suggest that radical lifetimes are on the 0.1-1 µsec timescale in the presence of radical scavengers.

Typically during FPOP experiments the H_2O_2/scavenger solution is added to the protein sample immediately prior to irradiation, minimising any oxidation which may occur from low levels of free radicals already present. Alternative set-ups such as one in which multiple sample pumps and a micro mixer-Tee are used to mix the protein sample and H_2O_2/scavenger solutions on-line immediately prior to the irradiation window have been demonstrated. These workflows have been demonstrated to be useful when excessive oxidation levels are a concern [201] and indeed has been shown to increase the reproducibility of oxidation levels [184]. Laser-off controls, in which the protein sample goes through the FPOP procedure but with no irradiation, should also be run in parallel so as to correct for any background oxidation that may occur during the labelling procedure or during post-labelling handling and storage.

The final stage of the flow path is the sample collection tube, such as an Eppendorf, which is filled with a quenching solution typically containing the enzyme catalase and mM levels of free methionine which remove any unreacted hydroxyl radicals or other radical species which can continue to oxidise the protein post FPOP. Following completion of the sample flow, the irradiated sample and quenching solution are immediately vortexed and samples can now be stored under appropriate conditions for further processing and analysis.

1.2.3.2.7 MS measurements and identification

As with all protein footprinting FPOP mediated oxidation can be measured either at intact protein or the peptide level, providing global and localised structural information respectively (Figure 35). The wide variety of oxidative products (~ + 13.98, +15.99, +32, -30) can then be assigned based on deconvoluted mass shifts in MS spectra.

Intact analysis is the most simple form of analysis, typically including some form of desalting or RP-LC step before ESI-MS analysis. The denaturing conditions in RP-LC however can lead to loss of sensitivity towards low abundant oxidative species, due to excessive charging of the protein ions leading to loss of signal-to-noise. Oxidation extent can be assessed using the base peak of the protein spectra or by using deconvolution software to abstract all the masses of the protein species present. Presence of Poisson distribution of the unmodified and oxidated species is indicative of a single protein conformation and the absence of secondary oxidation events [196, 202].

One of the advantages of covalent footprinting is the irreversible nature of the labelling, meaning samples can be stored and re-analysed in multiple analyses. Facilitating more established and sensitive bottom-up proteomic workflows which enable peptide and amino acid residue localisation of modification. Sequence coverage across the whole protein can be maximised with the use of single or combined proteases (e.g. Trypsin, Lys-C, Asp-N, chymotrypsin) to digest FPOP samples, as well as effective chromatographic separation of proteolytic products. Extensive fragmentation of modified and unmodified peptides can be ensured with use of MS/MS experiments (eg. CID, ETD, HCD), of which subsequent datasets can be analysed using a range of proteomic software packages (e.g PEAKS studio [160] and Mascot [161]) to identify peptides and localisation of modification sites.

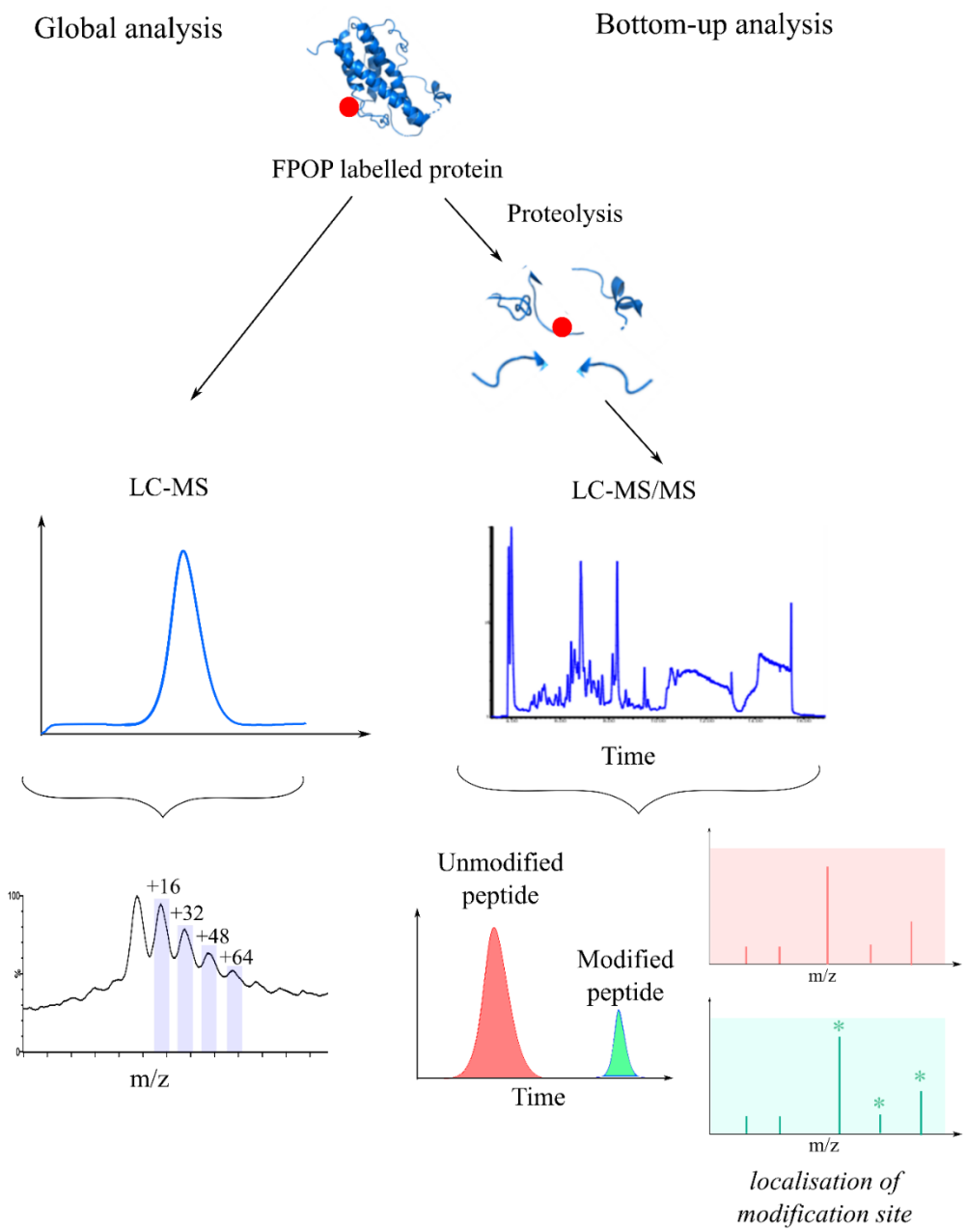


Figure 35. MS analysis workflows of FPOP labelled protein samples. Analysis can be conducted at the global level to measure the extent of modification or bottom-up workflows to localise and measure the extent of modification to specific amino acid side chains.

1.2.3.2.8 MS quantification

The final stage in the FPOP workflow is that of relative quantification, in which the extent of modification is compared to that of the unmodified protein or peptide, and modification levels compared between samples (e.g a ligand bound and unbound

FPOP labelled proteins). Quantification is performed relatively, expressed as a % modified, using the extracted ion chromatogram (XIC) of the exact masses of modified and unmodified peptide/protein forms, normalising the fraction of the modified peptide against the unoxidised fraction, such as:

Equation 25

$$\%\, modified = \frac{I(c_i)_{mod}}{I(c_i)_{mod} + I(c_i)_{unmod}}$$

Where I is the intensity or peak area of the XIC for the modified and unmodified species. Quantification is considered relative due to differences in retention time between modified and unmodified peptides (oxidised peptides often eluting prior to the unmodified peak [102]), which lead to differences in ionisation efficiencies due to differing chromatographic compositions. Should multiple forms of modified peptides arise due to multiple sites of oxidation at differing amino acid side chains, the relative extent of modification can also be expressed at an amino acidic level.

Modification levels can then be compared across sample types, and use of statistical tools such as Student's t-test applied to identify statistically significant differences in oxidation. Relative increases or decreases in the amount (or location) of oxidation observed can be interpreted with respect to the SASA of the amino acid residue relative to a reference protein state, though intrinsic reactivity and primary structural effects should also be considered when interpreting data (see section 1.2.3.2.2). These data can be considered in parallel with high resolution datasets such as XRC, NMR and HDX.

1.3 Aims and objectives of work

High resolution measurement of protein HOS has become increasingly important within the biopharmaceutical space from a regulatory perspective. Due to its relatively high sensitivity, resolution and applicability to all levels of protein HOS, MS based methods have been increasingly utilised as an analytical tool in this space. Yet, as these powerful tools move from academia to industry the quality, reproducibility, precision and interpretation of data generated is being challenged. Protocols and quality control (QC) materials are needed to validate such methodologies and to better

understand comparability of both within-platform analytical measurements and across orthogonal platforms. Furthermore, data needs to be produced to discriminate between structural changes and measurement variability, to increase confidence in the MS measurements and facilitate comparability studies.

International reference standards have existed in the clinical and diagnostic area for decades, for both small molecule and protein measurements, facilitating comparability studies and harmonisation of diagnostic assays. However such materials for protein HOS measurement have been lacking and as such understanding the comparability and differing sensitivity of platforms for HOS measurement has been challenging. This thesis aims to support the metrological gap in these measurements by developing reference protocols and model systems to increase confidence in MS based HOS measurements and validate new analytical platforms for HOS measurements.

This thesis will seek to apply analytical concepts of repeatability, reproducibility and sources of measurement variability to assess three MS based technologies for protein HOS characterisation. These include two orthogonal protein footprinting technologies, HDX-MS and FPOP-MS, to provide high resolution structural insights and use of higher throughput IMS-MS technology, that would be more applicable to a regulatory QC environment.

The specific objectives of this study are;

- To develop protocols for the generation of modified protein HOS states, known as protein model systems, of two commercially available protein RM, for the assessment of the sensitivity of differing analytical platforms towards structural changes. These modified protein states will be pharmaceutically relevant including ligand states (Chapter 2) and glycan variants (Chapter 3).
- To use these RMs to develop and establish the reproducibility and robustness of MS based methods, HDX-MS, IMS-MS (Chapter 2) and FPOP-MS (Chapter 4) for protein HOS measurement.
- To use these protein RMs to systematically establish and minimise sources of measurement variability of MS methods for protein HOS measurement so as to optimise their sensitivity towards small structural changes (Chapter 2 and 4).
- To use HDX-MS and IMS-MS methods developed to validate the model protein systems (Chapters 2-3) and characterise the structural changes occurring in the

modified protein states. These data in turn will be used to validate more novel analysis based on FPOP-MS measurements (Chapter 4) and understand the comparability of MS based HOS measurements.

- To generate data that act as reference datasets of these model systems, that can be used to benchmark and optimise more novel analytical technologies, such as described for FPOP-MS (Chapter 4).

Chapter 2

Development of a reference protocol for higher order structure characterisation of small recombinant protein biotherapeutics

Chapter 2

Development of a reference protocol for higher order structure characterisation of small recombinant protein biotherapeutics

2.1 Introduction

2.1.1 Need for reference standards and model systems for HOS measurements

The availability of reference standards and model systems to assess and validate analytical platforms used for measuring changes in HOS is lacking. Further still, regulatory guidelines on the most appropriate analytical platform to use when trying to define if a protein has undergone a structural change are also wanting. This study seeks to address this space by the development of model systems, protocols and generating data sets to under pin comparability of measurements across different analytical platforms and define their sensitivity towards structural changes.

An ideal candidate protein for conducting this cross platform characterisation study should be stable, bio-pharmaceutically relevant or simulate as such, have a defined HOS and be easily sourced with a batch-to-batch consistency. A model system will then be developed where the candidate protein exists in a reference HOS state and in a "stressed" state, where the protein has undergone some sort of controlled change in HOS. These changes should be small, localised and not associated with the total unfolding of the reference protein state, so as to truly simulate the type of structural heterogeneity that may arise during manufacturing processing. Better still, ideally the magnitude of the HOS change should be controllable, facilitating the use of the model system to assess the sensitivity of an analytical platform towards structural changes (i.e. the ability to detect both large and small structural changes).

The WHO material 98/574 is an international RM , consisting of recombinant human growth hormone (rhGH), that fulfils the criteria of candidate protein and additionally, the well characterised nature of rhGH structure (HDX-MS [203–205], IMS-MS [206], XRC [207, 208], NMR [209], CD [210] and SEC [211]) facilitates its use for validation of methods developed here.

2.1.2 Human growth hormone (hGH)

Human growth hormone (hGH) is a class-I helical cytokine protein produced in the pituitary gland [212], involved in modulation of growth, muscle differentiation, bone and cartilage cells and lactogenic signalling [208]. A variety of health implications can arise from hGH deficiencies including increased risk of cardiovascular disease, growth abnormalities, increased insulin resistance plus decreased muscle and bone mass. Since 1984, a recombinant form of hGH (rhGH, DNA-derived), has been used as a biotherapeutic treatment for hGH deficient adolescents and it is an example of a well characterised protein therapeutic [204, 208, 209].

rhGH consists of 191 amino acid residues, corresponding to a molecular mass of ~ 22 kDa. Predominantly α-helical in structure, its tertiary structure is based on a four-α-helix bundle motif observed with class-I cytokines. This bundle of helices, designated A-D (Figure 36) is tightly packed, with the first helices, A and B, running in parallel with respect to one another followed by the second two helices, C and D, running in anti-parallel. These adjacent helices are stabilised by two disulphide bridges and this combined with predominantly hydrophobic amino acid residues present in these four helices gives rhGH structure a very stable hydrophobic core [207]. rhGH also displays regions lacking in defined structure, located in the linker regions between the helices. This mixture of regions of very defined HOS and more flexible HOS make rhGH an attractive candidate protein to test the capabilities of structural analytical platforms.

2.1.3 Model system: rhGH and zinc

The presence of zinc was identified as a reagent for generating a potential "stressed" protein state for the rhGH model system based on the literature [208, 213, 214]. The presence of Zn^{2+} has been documented to promote the binding affinity of hGH to the transmembrane growth hormone receptors (hGHR) and pro-lactin receptors (hPRLR)[214], by up to 8,000 times, a key interaction in modulation in cell growth and development. This Zn^{2+}-binding site has been identified by XRC to occur at His18 and Glu174 residues, with the His21 side chain contributing towards the binding through the co-ordination and stabilisation of Glu174 residue [208]. Binding stoichiometry of hGH-Zn^{2+} complex is reported to be 1:1, and the Kd of the complex estimated at ~ 1 μM [213]. Additionally, at high concentrations the presence of Zn^{2+} has been shown to induce dimerization of hGH and the hGH-dimer acts as a possible storage form due to the high levels of Zn^{2+} found in hGH storage granules located in

the pituitary glands [215]. The use of Zn^{2+} has also been explored as a precipitating agent in formulation processing of protein biotherapeutics with minor tertiary structural changes identified [210].

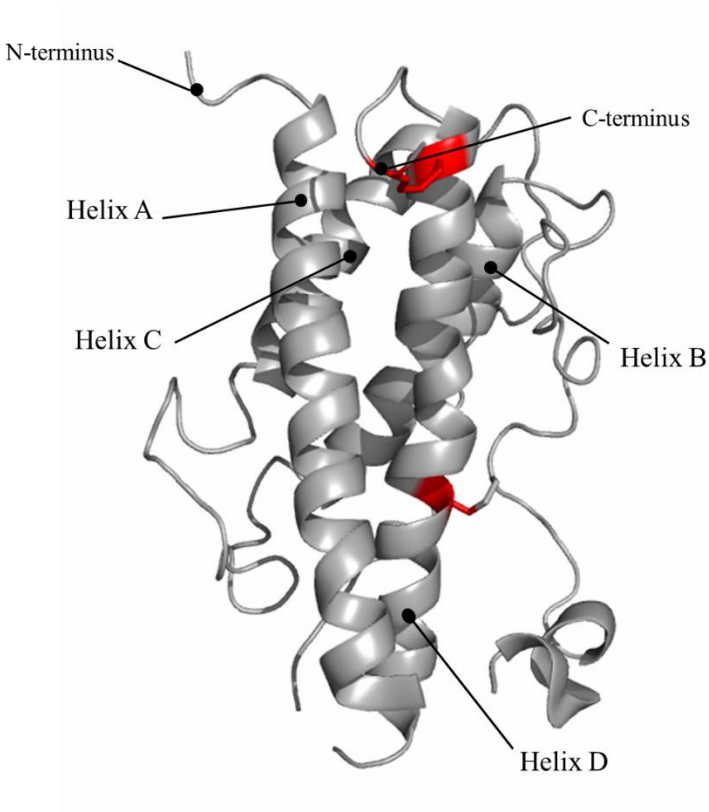


Figure 36. Higher order structure of hGH based on the crystal structure of PDB 1HGU [216].

2.1.4 Defining the sensitivity of HDX-MS measurements

When moving towards the standardization of an analytical technique, standard operating procedures (SOP), QCs, repeatability and reproducibility studies should be utilised to understand and minimise measurement uncertainty, and hence facilitate both inter and cross-platform comparison of measurements.

2.1.4.1 Comparability and reproducibility of HDX-MS measurements

Recently, several commercially available automated on-line HDX-MS platforms have become available which allow for automated sample handling, dilution, injection, on-line digestion and chromatographic separation [127]. There is, however, much work to be done in this area with respect to the standardization of HDX-MS measurements. Indeed a report published by the International Bureau of Weights and Measures

(BIPM) recognized the potential of HDX-MS measurements for the characterisation of protein HOS but also the need for standardisation of the methods employed [217]. The use of robotic sample handling systems has been effective in creating data sets large enough to access the variability of HDX-MS. Studies from Chalmers[137], Burkitt [138] and Houde et al. [112] have assessed platforms over a defined time period, all of which concluding that robotic set ups aid the reproducibility of the system and that inter-day variability was higher than intra-day variability, though careful control of variables could result in the variability of percentage deuterium to be as low as 1.0-2.2% [138].

2.1.4.2 Establishing significance thresholds of HDX-MS measurements

Pharmaceutical regulators require manufacturers to demonstrate the manufacturing consistency across batches and process changes, termed comparability, of a drug [218]. Traditionally, comparability studies have been performed using biophysical techniques such as CD, LC-UV, DSC, isothermal titration calorimetry (ITC) and analytical ultracentrifugation (AUC), which provide only low resolution global folding profiles of biopharmaceuticals and are known to often demonstrate low sensitivity to small structural changes. The high resolution nature of HDX-MS measurements make the technique attractive not only in pharmaceutical R&D but also in comparability studies. It is worth noting that though HDX-MS measurements provide higher resolution, there is a substantial cost in terms of analysis time and throughput compared to traditional biophysical techniques. The commercialisation and gains in reproducibility from the development of automated set-ups has led to increased usage of HDX-MS across academia and industry, particularly the biopharmaceutical sector. Yet, measurement reproducibility and/or uncertainty must first be established to define significance thresholds before the technique is widely exploited for comparability studies.

Statistical approaches that help define significant structural changes are required to support the use of multiple HDX-MS datasets in this comparability area, and indeed this has been a small but growing field of research, Table 4 summarises some of this published work. Though much of this research, accounts statistically for inter-day variability differences between two samples, there is no account for intra-day variability nor for measurement variability, reducing the robustness of the approaches.

One of the most comprehensive approaches, described by Houde et al. [112], reported the use of HDX-MS measurements in a comparability study of several different preparations of interferon-β-1a (IFN), by performing a series of lot-to-lot comparisons, known as a vial-to-vial study, of seemingly identical samples of IFN. The resulting differential HDX-MS measurements were used to establish several measurement thresholds above which would define a significant structural change in a quantitative way whilst considering the variability of HDX-MS measurements. A drawback of this approach is that a single global threshold value was established across all the peptides observed for IFN. Potentially this may mask any small significant structural change that is experienced by an individual but reproducible (in terms of comparability) peptide, decreasing the sensitivity of the approaches towards small but real structural changes. Recently this approach has been extended further by Hageman et al. [168], who illustrated the false positives and inaccuracy of using a single global threshold limit and how sensitivity and accuracy of structural changes could be increased by use of a hybrid statistical approach which assessed the data first on a global threshold limited then on an individual threshold basis using a Welch's t-test, an unequal variances t-test. This approach was shown to reduce false positives, in this study, however is only applicable to discrete HDX-MS labelling times, removing any HDX trends that may have a time dependency.

Table 4. Summary of some of published statistical approaches for defining the significance of differential HDX-MS measurements.

Publication	Software incorporation	Statistical approach	Brief description
Chalmers et al. [137]	HDX-Workbench	Two-tailed t-test, Tukey multiple comparison procedure	Devoid of intra-day variability estimation, multiple time points or time-dependency of HDX
Houde et al.[112]	N/A	Single global significance value based on standard error of the mean (SEM)	Multiple vial-to-vial data sets generated, global significance values generated for individual or multiple summed time points of HDX. Single global threshold value not peptide specific.
Rey et al. [219]	MassSpec Studio	Two tailed t-tests	Devoid of intra-day variability estimation, multiple time points or time-dependency of HDX
Hourdel et al. [220]	MEM-HDX	Linear mixed- effects model	Devoid of intra-day variability estimation, replicates considered random effects, time dependency and replicate variability considered, p-value for ΔD and change in dynamics
Hageman et al. [168]	N/A	Hybrid significance testing approach, global significance based on SEM, individual significance generated using Welch's t-tests	Intra-day variability estimation generated from vial-to-vial matrix, global and individual significance testing of single time points, no testing of multiple HDX time points.

2.1.5 Native IMS-MS and comparability studies

Though providing high sensitivity and high resolution insights into HOS structure, HDX-MS comparability studies have several caveats. These include often lengthy and complex experimental/data analysis approaches and measurements provide only an average measurement of the conformations present in solution, masking differences in heterogeniety. Native IMS-MS measurements, however, offer an alternative approach for HOS comparison with a much reduced experimental/data analysis time, requiring much less sample and the ability to measure and monitor the presence of multiple conformational species that could be present in a sample.

Ferguson et al. [49] systematically assessed the suitability of IMS-MS and CIU IMS-MS to compare and differentiate between mAb-based biotherapeutics, evaluating the reproducibility of several different metrics and applying these to establish lot-to-lot variability. These metrics included species drift time (t_D), a peak intensity ratio (PIR), peak width at 50% maximal intensity (FWHM) and peak width at 10% maximal intensity (FWTH). t_D measurements were established to be the most reproducible metrics (% relative standard deviation, RSD < 1.44%). The utility of IMS-MS and CIU-IMS-MS measurements for distinguishing between two mAb based drugs was demonstrated for systems with large global changes in structure. However, the study questioned the applicability of the technique to compare across biosimilar mAbs from two differing manufacturers. The study also compared $^{TW}CCS_{N2}$ values calculated using external calibration required for non-drift tube IMS-MS and concluded that considering t_D metrics was more sensitive for structural change detection than comparison of $^{TW}CCS_{N2}$ values alone.

Analytical methods used to support regulation of therapeutic pharmaceuticals must be demonstrated to be robust and reproducible [60]. This Chapter reports the use of a rhGH RM to optimise and critically assess repeatability and reproducibility of HDX-MS and IMS-MS measurements. Considering the reproducibility of these measurements, these methodologies were then applied to characterise the proposed model system of rhGH in the absence and presence of zinc at two rhGH:Zn concentrations,1:2 and 1:10, defining any significant structural changes observed whilst commenting on the sensitivity of each analytical platform towards structural changes.

2.2 Materials and Methods

2.2.1. Materials

All chemicals and standards were purchased from Sigma Aldrich [Gillingham, UK] and were of research grade purity unless otherwise stated. Ultra-pure water (18.2 MΩ·cm) was used. Formic acid and TCEP (tris(2-carboxyethyl)phosphine) were purchased from Thermo Fisher Scientific [Loughborough, UK]. Acetonitrile was Optigrade HPLC Special Grade [LGC Standards, Teddington, UK]. rhGH RM (WHO98/574) was purchased from NIBSC [Potters Bar, UK].

2.2.2. Protein sample preparation

A 36 µM rhGH stock was prepared by solubilisation of the lyophilised RM in water and vortexing for 1 hour at room temperature. Following this the stock was aliquoted and stored at -20 °C to be removed for single use.

2.2.1.1 Sample preparation for HDX-MS and SEC measurements

A 0.9 mM zinc acetate stock solution was also freshly prepared in water and further diluted in 10mM KPBS (pH 7.4) to obtain a series of solutions; solution C (180 µM zinc acetate), solution D (90 µM zinc acetate) and solution E (18 µM zinc acetate). Prior to analysis, 36 µM stock aliquots of rhGH were defrosted and further diluted to 7.2 µM (0.16 mg/mL) with 10 mM KPBS or solution D or E to obtain 1:10 and 1:2 respectively. Samples were left to equilibrate at room temperature for 1h before analysis.

2.2.1.2 Sample preparation for IMS-MS measurements

For native IMS-MS analysis, zinc acetate solutions were prepared from dilutions of a 0.9 mM zinc acetate stock solution with a 20 mM ammonium acetate solution. A second series of solutions were prepared; solution F (90 µM zinc acetate) and solution G (18 µM zinc acetate).

Prior to buffer exchange and desalting steps, aliquots of 36 µM rhGH stock were removed from the freezer and allowed to defrost and equilibrate for 45min. Desalting and buffer exchange of the stock sample was performed using Bio-spin micro P6 pre-packed size exclusion columns (Bio-rad, California, U.S.A) as per the manufacturer's

instructions. In short, excess packing buffer was removed using an 5810R centrifuge (Eppendorf, Stevenage, U.K) at 1000 g for 2min. Buffer exchange was then performed by loading 500 µL 20 mM ammonium acetate buffer and centrifuging for 2 min at 1000 g, repeating this step 4 times to ensure 99.9% buffer exchange. 75 µL of rhGH stock was then loaded directly onto the centre of the column and centrifuged for 4min, collecting the eluent. To maximise desalting the buffer exchange process was repeated as described and the sample loaded for a second time.

The desalted and buffer exchanged rhGH was then diluted to 7.2 µM using 20 mM ammonium acetate or zinc acetate solutions F and G to obtain rhGH:Zn 1:10 and 1:2 respectively. A final sample was prepared by dilution of the desalted rhGH stock using a solution of 180 µM of magnesium acetate to give rhGH in the presence of magnesium (rhGH:Mg, 1:10). Samples were left to equilibrate at room temperature for 1h before analysis.

Calibrant solutions of 10 µM Beta-lactoglobulin (B-LAC) in 200 mM ammonium acetate were also prepared as described by Bush et al [52].

2.2.1.3 Preparation of deuterated and undeuterated peptide mixture

A peptide mixture containing MRFA [74 µM], bombesin [24 µM], angiotensin II [angio II, 191 µM], and glu-1-fibrinopeptide B [glufib, 64 µM], insulin chain B [ICB, 77 µM] was prepared from stock solutions. This mixture was then diluted 1200-fold in either deuterated 10mM phosphate buffer, (pH 7.0) or phosphate buffer, (pH 7.4) to act as deuterated and undeuterated peptide mixtures respectively. Both samples were shaken at 25 °C overnight to ensure complete exchange occurred for the deuterated sample and to take into account potential stability issues.

2.2.1.4 Preparation of deuterated and undeuterated α1-Antitrypsin

A 18.5 µM deuterated α1-Antitrypsin stock was prepared by solubilisation of lyophilised α1-Antitrypsin in a deuterated buffer containing pre-deuterated 2M urea, 10 mM potassium phosphate (pH 7.0) solution. To facilitate complete deuteration the protein stock was then heated at 30 °C overnight. Prior to analysis protein stock was then to 3.7 µM using the pre-deuterated buffer. An undeuterated α1-Antitrypsin control stock was also prepared in parallel using undeurated buffers.

2.2.3. HDX-MS

2.2.3.1 Instrumentation

Sample handling and mixing steps were performed using a first generation LEAP PAL system set up for HDX analysis [LEAP Technologies, Carrboro, USA]. Samples were injected onto a refrigerated nanoACQUITY UPLC System with HDX technology [Waters, Milford, USA] for on-line pepsin digestion and chromatographic separation. MS experiments were performed on a Synapt G2Si QToF instrument [Waters, Milford, USA].

2.2.3.2 HDX-MS experiments

For the generation of peptide maps and time zero exchange experiments, 15 µL of rhGH (0.16 mg/mL) was diluted 10-fold in 10 mM KPBS (pH 7.4). For exchange experiments, deuterium labelling was performed by diluting 15 µL of rhGH 10-fold in 10 mM KPBS (pH 7.0) prepared in D_2O. Exchange times were 0.5, 5, 60, 240 and 480 min, with the exchange occurring at room temperature (21±2 °C). Triplicate analyses were carried out for each incubation. Following the appropriate incubation period, exchange was quenched by a 2-fold dilution with 50 µL of 300 mM TCEP, 2 M Guanidine Hydrochloride (Gnd.HCl) in 100 mM KPBS (pH 2.5) at 4°C. Following a 0.5 min quench delay, a 50 µL sample loop was overfilled with 95 µL of quenched sample.

On-line digestion was performed using an Enzymate BEH pepsin column (5 µm, 2.1 x 30 mm) [Waters, Milford, USA] at 25 °C and a flow rate was 80 µL/min (mobile phase 0.05% v/v formic acid). Proteolytic peptides were trapped on an ACE C18 guard cartridge (5 µm, 2.1 mm i.d.) and sequentially chromatographically separated on an ACE Excel Super C18, (2 µm, 2.1 x 100 mm i.d, both Hichrom Reading, U.K), both trap and analytical column were held at 0 °C. Digestion pressure was held at ~ 4000 psi using a PEEK restrictor placed post-trap column. Chromatographic separation was carried out at 100 µL/min by application of a 7 min linear gradient from 92% A / 8% B to 65% A / 35% B. Mobile phases consisted of aqueous, 0.1% v/v formic acid (A) and CH_3CN, 0.1% v/v formic (B).

MS experiments were performed in positive ESI mode using an 80 °C source temperature, 2.8 kV capillary voltage, 80 V cone voltage and a desolvation temperature of 250 °C. For peptide map generation, data were acquired in MS^E

continuum mode over a range of 50 – 2000 amu, with a scan time of 0.2 s, inter scan time of 0.1s and a transfer ramp of 15-40 V. Exchange experiments acquired in MS mode under the same conditions but with no collision energy applied. Mass calibration was performed using a 100 fmol solution of Glu-fibrogen, (50% MeOH, 0.1% formic) with a lockmass of 785.8426 *m/z*, mass correction was applied post acquisition.

2.2.4 Data analysis

2.2.4.1 Peptide identification

ProteinLynx Global Server software v 3.02 [PLGS, Waters, Milford, USA] was used to generate peak lists by inputting MS^E data and searching against a database consisting of both FASTA sequences of porcine Pepsin A (P00791, Uniprot) and Somatotropin (P01241, Uniprot), see Table A-24, in the appendix for parameters. The PLGS outputs were imported in DynamX v3.0 [Waters, Milford, USA] to generate peptide maps, file thresholds (peptides observed in x files out of y, n=x/y) of n=3/3 and n=4/5 were used for digestion optimisation experiments and HDX-MS peptide maps respectively. Additional processing parameters included 0.01, minimum products per amino acid and 20 ppm, maximum MHP+ error.

2.2.4.2 Deuterium uptake measurements

Peptide deuterium levels were calculated as described by Burkitt et al. [138], based on the difference between the calculated centroid mass measurement of the isotopic distribution of the undeuterated control sample and the deuterated sample at t=x. No correction was made for back exchange during deuterium uptake calculations. Measurements are expressed as deuterium uptake (in Da) or as percentage relative deuterium uptake (rDU), calculated by the division of deuterium uptake (in Da) by the theoretical maximum number of exchangeable backbone amide protons in the peptide (using the formula, N-2-P, where N is the number of amino acid residues in the sequence and P is the number of proline residues not including the first two N-terminal residues [120] and expressed as a percentage). Differential calculations were performed using deuterium uptake (in Da) measurements.

2.2.4.3 Qualitative assessment of peptide map

To make a qualitative assessment of the efficiency of digestion under different conditions, a parameter called digestion factor (DF) [221], shown in Equation 26 was used, this factor incorporates a weighted average of the sequence coverage achieved

(y_1), the percentage of potential cleavage sites accessed (y_2) and the redundancy (y_3), a measure of overlapping peptide sequences..

Equation 26

$$DF = 8y_1 + 5y_2 + 0.6y_3$$

The sequence coverage (y_1) is the proportion of the primary sequence which is present in the peptides identified (S) relative to the number of amino acids present in the sequence (N).

Equation 27

$$y_1 = \frac{S}{N}$$

The cleavage rate (y_2) shows the proportion of peptide bonds cleaved (C) relative to N.

Equation 28

$$y_2 = \frac{C}{N}$$

The redundancy rate (y_3) is the sum of the number of times that an individual amino acid is seen across multiple peptides (R) relative to N.

Equation 29

$$y_3 = \frac{\sum R}{N}$$

The digestion factor calculation allows for the three components contributing to efficient digestion to be evaluated simultaneously, whist still allowing some definition in the contribution of these factors, for instance the factor of highest importance, sequence coverage, y_1, is weighted more than redundancy, y_3.

2.2.5 Size exclusion chromatography

SEC-UV experiments were performed using an 1100 series LC system with a variable wavelength detector [both Agilent technologies, California, U.S.A]. Separation was achieved across a Yarra SEC-X150 column (1.8 µm, 4.6 x 150 mm i.d, Phenomenex,

Cheshire, U.K) using an isocratic gradient of a 150 mM PBS, 50 mM NaCl (pH 7.4) mobile phase, at a flow rate of 0.18 mL/min for 15 min.

Triplicate injections of 27.5 µL for each rhGH sample (0.16 mg/mL) were carried out and measurements acquired at ʎ=210 nm. A bovine serum albumin solution (BSA, 5 µL, 1 mg/mL) was also injected to act as a reference point against rhGH aggregates. Integration of chromatographic peaks was performed using ChemStation software. Results were reported as relative percentage dimer (%) calculated as the (area of dimeric peak)/(area of dimeric peak + area of monomeric peak).

2.2.6 Native IMS-MS

2.2.6.1 Instrumentation and experimental settings

nESI-IMS-MS experiments were performed on a Synapt G2Si Q-ToF instrument equipped with a nanoESI source and infused using gold-plated borosilicate capillaries prepared *in-house*. Experimental settings were optimised as follows: source temperature 30°C, sampling cone 5 V, source offset 2 V, capillary voltage 1.4 kV, trap CE 5 V, trap gas flow 3 mL/min, helium cell has flow 120 mL/min, IMS gas flow 60 mL/min, trap DC entrance 3 V, trap DC bias 30 V, trap DC 0 V, trap DC exit 0 V, IMS DC entrance 20V, helium Cell DC 50 V, helium exit -20 V, IMS bias 3 V, IMS DC exit 0 V, transfer DC Entrance 4 V, transfer DC exit 15 V, IMS wave velocity 1100 m/s, IMS wave height 40 V, transfer wave velocity 89 m/s, transfer wave height 1 V, stepwave1 RF offset 300, stepwave2 RF wave offset 350. Unless otherwise stated all other settings were set to default values. Data were acquired in resolution mode in the 500-8000 *m/z* with a 1 sec scan time for 2 minutes. No cone or purge gas was used. Parameters were systematically optimised by infusion of rhGH in IMS-MS mode whilst monitoring t_D to ensure effective transition of gas ions whilst avoiding unfolding of the gas ions within the instrument.

Mass calibration was performed by infusion of a 1 mg/mL caesium iodide solution in the range of 500-5000 *m/z*. Data were processed using MassLynx v4.1 and Driftscope v2.5 software [both Waters, Milford, USA].

2.2.6.2 CIU-IMS-MS experiments

CIU IMS-MS experiments were conducted using the experimental conditions stated above but with the quadrupole selection +8 charge state (2765.5 *m/z*) of the rhGH control and that of the equivalent singly-bound zinc species (2773.5 *m/z*) for rhGH:Zn

1:2 and 1:10 samples. The trap collision voltage was increased in 2 V increments between the range of 6 – 36 V and data acquired for 1 min at a 2 s scan rate for each voltage. To avoid any systematic errors which may occur during the course of the experiments/spray, IMS-MS spectra experiments were performed in a randomised order across the voltage range.

2.2.6.3 Data Analysis

Mass spectra, chromatograms and extracted mobiligrams were analysed using MassLynx v 4.1. Mobility data were also analysed using DriftScope v. 2.5. CIU IMS-MS plots were generated using PULSAR (Protein Unfolding for Ligand Stabilisation and Ranking) [98].

2.2.6.4 CCS measurements and other equations

Using the methods described by [53] and calibrant reference measurements published by Bush et al [52], B-LAC monomer and dimer populations were used as calibrants to obtain $^{TW}CCS_{N2}$ measurements. Briefly using the optimised IMS-MS methods, drift time (t_D) measurements were recorded for monomeric (+7, +8) and dimeric (+11, +12, +13) charge states of B-LAC. Using Microsoft Excel a linear regression model was fitted to plots of Ω' ($^{TW}CCS_{N2}$ nm^2) vs t_D' (ms), using equations described in (1.1.6.2 ~ Equation 20). The subsequent calibration line generated was used to adjust measured TWIMS t_D of rhGH species.

An estimation of measurement uncertainty was calculated for each $^{TW}CCS_{N2}$ measurement generated, which encompassed measurement error arising from mass uncertainties and uncertainty arising from the observed deviation from a linear regression calibration line.

Equation 30

$$u(\Omega) = \Omega \sqrt{\left(\frac{u(e)}{e}\right)^2 + \left(\frac{u(m_{\text{reduced}})}{m_{\text{reduced}}}\right)^2 + \left(\frac{u(\Omega')}{\Omega'}\right)^2}$$

To calculate the theoretical maximum charge, z max, that can be held by a protein of mass, m, whilst retaining its native-like conformation, the following equation was used [77]:

Equation 31

$$z = 0.077 \times m^{½}$$

2.2.7 rhGH structural models

The model used to represent the structure of free rhGH was based on the crystal structure of the free wild type hGH (PDB ID:1HGU [216]) solved to a resolution of 2.5 Å. The model used to represent the structure of Zn^{2+}-bound rhGH is constructed using 1:2 hGH:hGHPLR complex (PBD ID: 1BP3, [208]), in which a Zn^{2+} ion is present and the structure is solved to resolution of 2.9 Å. Graphics representing this three dimensional data have been generated using PyMOL (The PyMOL Molecular Graphics System, Version 1.74 Schrödinger, LLC.) as have all further XRC documented in this thesis.

2.3 Results and discussion

2.3.1 HDX-MS method optimisation

The key analytical stages involved in production of robust, reliable and reproducible HDX-MS measurements were systematically optimised and assessed and described herein.

2.3.1.1 Proteolytic digestion

In a bottom-up HDX-MS experiment, structural measurements are defined by the pool of peptides observed, termed the peptide map, onto which HDX-MS measurements are mapped, making optimisation of the proteolytic digestion stage a key experimental stage. Peptide map generation involves the production of a series of repeat on-line digestions of undeuterated sample, each of which is then desalted and separated by UPLC and identified accurately using MS/MS fragmentation.

High sequence coverage of the protein of interest is critical to generate peptides useful for monitoring structural changes by HDX-MS. Furthermore generating a large number of overlapping peptides (redundant peptides) is important to increase amino acidic resolution for the accurate identification and localisation of induced structural changes.

A range of proteolytic digestion conditions was identified as influencing proteolytic digestion and systematically assessed to maximize sequence coverage and

redundancy. These included: composition of the quenching solution (reducing and denaturing reagents present), temperature, and pepsin column flow rate (effective digestion time). Using the on-line digestion set-up described in the Materials and Methods 2.2.3.2, these factors were assessed and optimised for peptide map generation for HDX-MS experiments. Samples of rhGH (7.2 µM) prepared in 10 mM KPBS (pH 7.4) were diluted 1:10 in KPBS, mixed, quenched and injected on the HDX-MS system under a range of digestion conditions. To ensure a robust and repeatable peptide map only those peptides observed in 4/5 repeat injections were considered in the final peptide map. Peptide maps are qualitatively assessed considering the digestion factor (DF) assessment criteria as described in 2.2.4.3, and summarised in Table 5.

Table 5. Effects of varying digestion conditions on rhGH peptide map. All quench solutions contained 100 mM KPBS (pH 2.5) and any additional presence of reducing and/or denaturing reagents are stated. Resulting peptide maps were generated using those peptides observed in n≥4/5 digests and described considering sequence coverage (y1), cleavage rate (y2) redundancy (y3) and overall digestion factor (DF) score [221].

Condition number	**Flow rate (µL/min)**	**Temp (° C)**	**Quench composition**	**Coverage (y_1)**	**Cleavage (y_2)**	**Redundancy (y_3)**	**DF** ***($DF = 8y_1 + 5y_2 + 0.6y_3$)***
i	100	20	-	64.4%	23%	172%	7.3
ii	100	20	300 mM TCEP, 2M Gnd.HCl	90.6%	36%	279%	10.7
iii	80	10	300 mM TCEP, 2M Gnd.HCl	82.2%	28%	179%	9.1
iv	80	25	300 mM TCEP, 2M Gnd.HCl	96.3%	42%	388%	12.1
v	150	25	300 mM TCEP, 2M Gnd.HCl	94.2%	39%	369%	11.8

2.3.1.1.1 Effect of denaturing reagents

Comparison of resultant peptide maps produced in the absence and presence of denaturing and reducing agents (Table 5, conditions i and ii, and Figure 37), demonstrates the increased sequence coverage and redundancy due to the presence of the reagents. Digestion was increased, particularly around the two cysteine bridges in rhGH, positioned at Cys53-C165 and Cys182-189 [208], upon the addition of denaturing and reducing agents, as was coverage for the Phe10-Glu30 region and redundancy in area of Cys55-Leu75. It appears that the presence of these reagents in the quenching solution has led to the opening of the central α-helical bundle structure of rhGH, with the reduction of these bridges allowing for greater pepsin accessibility to these areas, resulting in significant increases in digestion of these covalently linked areas, as demonstrated. Secondary structures in this region are known to include a turn and α-helices; therefore, the reduction of cysteine bridges combined with the denaturing of the α-helices has facilitated better digestion.

2.3.1.1.2 Effect of on-line pepsin flow rate

During the digestion stage of the chromatography, controlling the flow rate of the pepsin column affects the contact time between protein and immobilised enzyme, or effective digestion time. Thereby controlling the pepsin flow rate allows efficacy of the digestion based on digestion time to be evaluated. Slower flow rates illustrate longer effective digestion times, with longer protein to pepsin interactions, compared to faster flow rates which have shorter digestion periods and less opportunity for interaction between the protein and enzyme. Figure 38 (produced from conditions iv and v, Table 5) illustrates a combined peptide map generated from rhGH peptides observed using two flow rates, 80 µL/min and 150 µL/min, giving effective digestion times of 3.8 and 2 min, respectively. Overall the digestions are similar in terms of total digestion factors (DF), with DF values of 12.13 (y1 – 96.3%, y2 – 42%, y3 – 388%) 11.75 (y1 – 94.2%, y2 – 39%, y3 – 369%) for 80 and 150 µL/min flow rates respectively. Many peptides were observed mutually under both flow rates covering some 93% of the sequence. Average peptide length for each flow rate is similar with values of 12.67 and 12.80 for 80 and 150 µL/min flow rates respectively. However, considering the standard deviation (SD) across peptide length for each flow rate reveals a value of 8.0 for 80 µL/min compared to 5.7 at 150 µL/min illustrating a

larger variety in peptide length at the faster digestion times. This is favourable for HDX-MS as it could lead to more spatial resolution of any HDX-MS data obtained under these conditions.

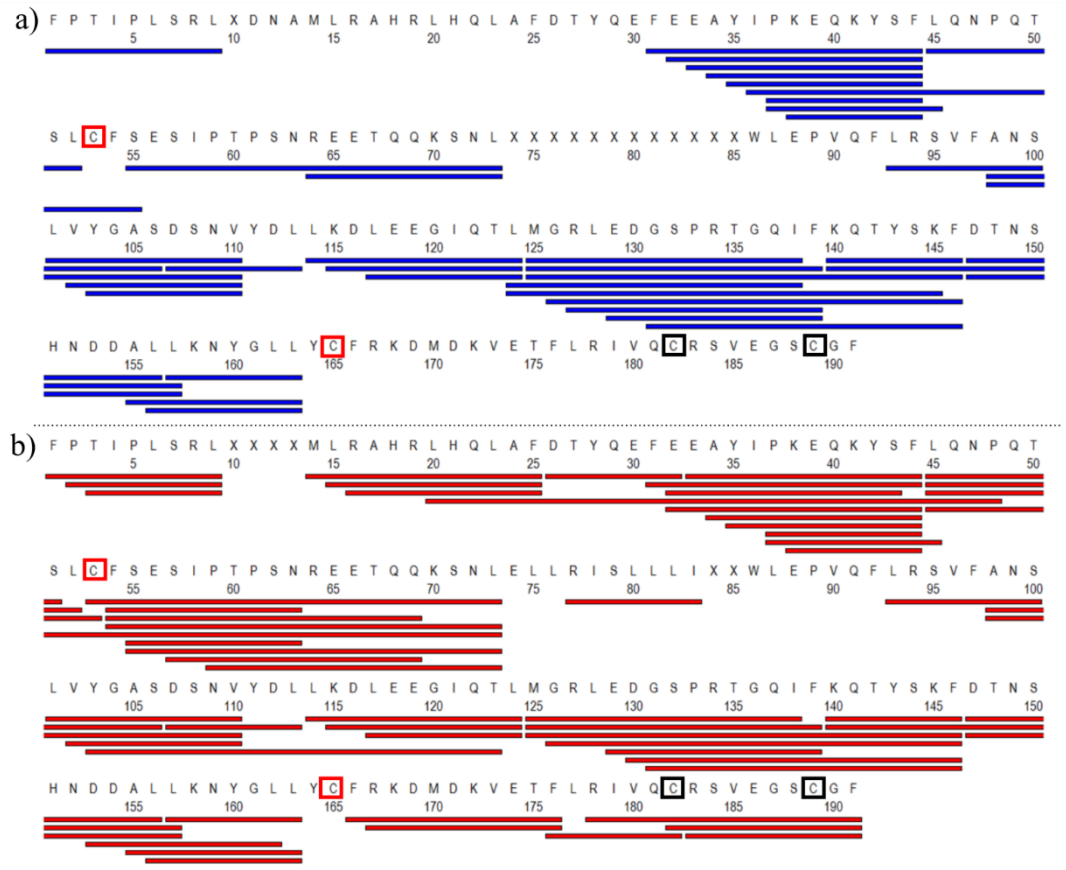


Figure 37. Peptide maps of rhGH generated from pepsin digestion in the a) absence and b) presence of 300 mM TCEP and 2 M Gnd.HCl. rhGH samples were digested using a flow rate of 100 ul/min and a temperature of 20 °C. Only peptides identified in n=4/5 repeats are illustrated. Cysteine residues between Cys53-Cys165 (red) and Cys182-Cys189 (black) are indicated.

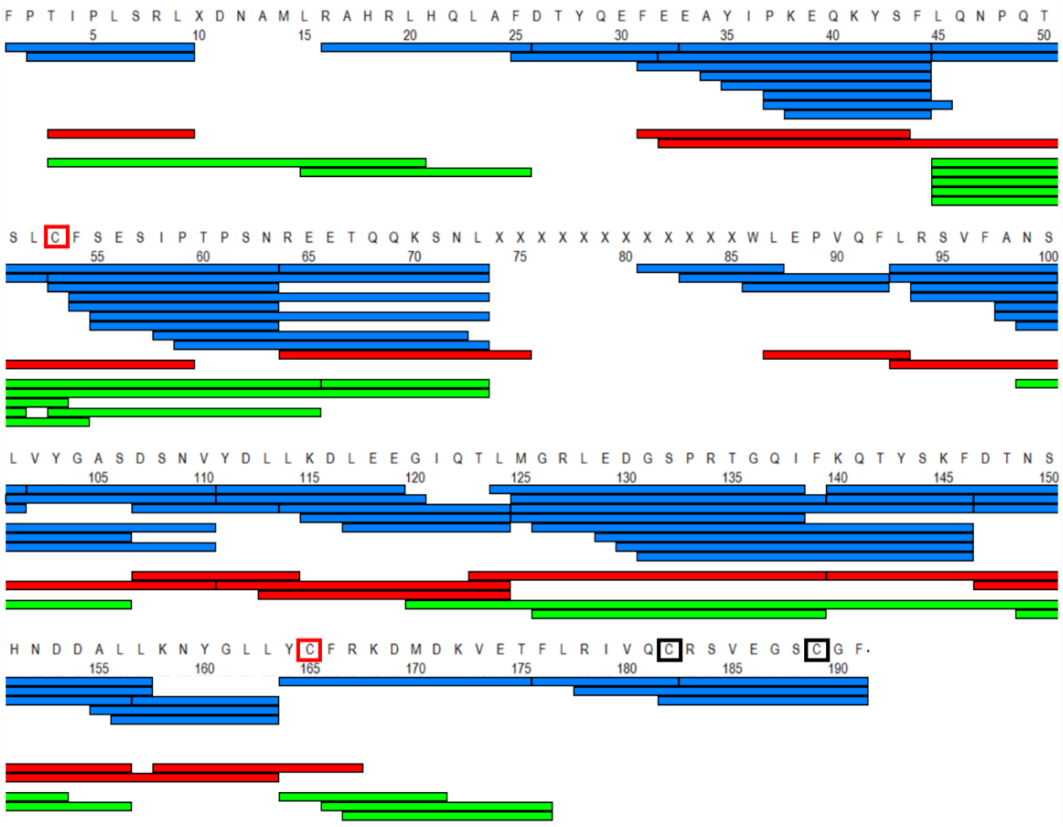


Figure 38. rhGH peptide maps generated with varying on-line pepsin flow rates.Flow rates used were 80 µL/min and 150 µL/min, with the presence of 300 mM TCEP and 2 M Gnd.HCl at a temperature of 10 °C. Peptides observed, under both flow rates (blue), exclusively at 80 µL/min (green) and exclusively at 150 µL/min (red) are illustrated. Only peptides identified in n=4/5 repeats are illustrated. This illustration was generated using MSTools [173] .

2.3.1.1.3 Effect of temperature

As with many proteases, it has been demonstrated that increased temperatures (> 0°C) increase the digestion efficiency of pepsin [222]. A temperature controlled compartment, housing the on-line pepsin column, built within the HDX-MS platform allowed digestion temperature to be tightly controlled. Using the optimised digestion additives and flow rates, rhGH digestions were performed at both 10 °C and 25 °C. (Table 5, conditions iii and iv, Figure 39). Increase in digestion efficiency across all the digestion factor components, was observed at higher temperature, but redundancy in particular was effected. A 15 °C increase resulted in peptide redundancy increasing from 179% to 388%, further increasing the resolution of any HDX-MS data produced under these digestion conditions. Though sequence coverage is relatively high ~ 90% with the optimised conditions, there is still 10% of the primary sequence, residues 10-

15 and 74-93, missing from the peptide maps produced and hence no HDX-MS data will be measured in these regions.

To avoid excessive contributions towards back exchange from increased temperatures and long digestion times, a compromise was met under these optimised conditions (25 °C and 80 µL/min) which attempted to maximise peptide map quality yet allow for the conservation of labelling. Though back exchange is generally minimised by the short time spent on the pepsin column and pre-cooling of solvents, keeping temperature to a minimum yet achieving good digestion is an important consideration for peptide optimisation and worth noting.

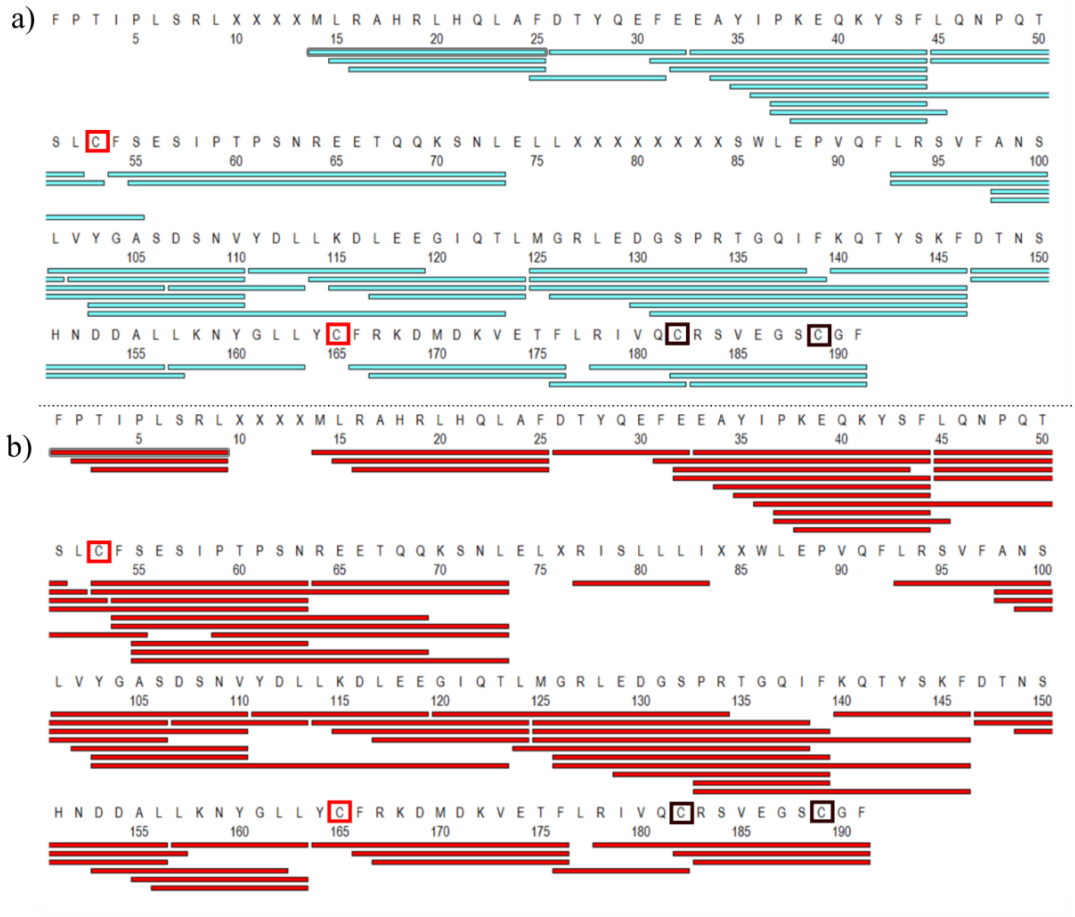


Figure 39. Peptide maps of rhGH generated from pepsin digestion with varying on-line pepsin temperatures. rhGH samples were digested using a flow rate of 80 µL/min, with the presence of 300 mM TCEP and 2 M Gnd.HCl and a temperature of a) 10 °C and b) 25°C Only peptides identified in n=4/5 repeats are illustrated. Disulphide bridges between Cys53-Cys165 (red) and Cys182-Cys189 (black) are indicated.

These data illustrate that in the case of the on-line pepsin digestion of rhGH, elevated temperatures, the presence of denaturing/reducing agents, and moderate pepsin flow rates are all favourable to producing a high quality peptide map for rhGH HDX-MS data. The optimal pepsin digestion conditions were found to be 300 mM TCEP and 2M Gnd.HCl, 25 °C, 80 µL/min flow rate and were applied to the HDX-MS experiments performed herein.

2.3.1.2 Back exchange assessment

As aforementioned, controlling the back exchange within the HDX-MS experiment, is key to maximising the sensitivity with respect to structural changes and reducing measurement variability. A mixture of undeuterated and fully deuterated peptide standards (MRFA, bombesin, angiotensin II, glu-1-fibrinopeptide B and insulin chain B) were analysed to assess the contribution to the total back exchange of (a) the trap and LC column, (b) the pepsin column, (c) the dilution/mixing steps (d) the MS source.

For the back-exchange assessment of the entire system, the mixtures of deuterated and non-deuterated peptides were quenched and analysed as described in section 2.2.3.2. Total back exchange (%) for the whole HDX-MS system for individual peptide standards are displayed in Table 6. To isolate different system components, the peptide sample was also injected directly on to the LC system with and without the pepsin column in place. Contribution of each component towards back exchange was then calculated by comparison of the deuterium uptake across these experiments.

The most significant amount of back exchange was found to occur across the trap and LC column (contributing ~ 76% of the total back exchange), presumably due to the overall proportion of time that the deuterated peptides are exposed to protons present in the mobile phase. As the temperature and flow rates of the pepsin column are critical factors in the optimization of the proteolytic digestion method, their effects on back exchange were also investigated by comparing the deuterium uptake of a number of peptides observed from a fully deuterated off-line digested sample of α1-Antitrypsin. It was found that the effect of flow rate on the back exchange was proportional to the amount of time the peptides spent on the pepsin column, whereas a drop of temperature from 25 to 20 °C during pepsin digestion resulted in no significant difference in back exchange (decrease of 0.1 ± 1.5%). An increase to 30 °C resulted in much greater back exchange (increase of 5.1 ± 1.5% across the peptides)

reiterating the need for careful consideration of digestion parameters when developing a HDX-MS method.

Table 6. Details of peptides used for back exchange experiments. Total percentage back exchange (across entire system including sample handling, pepsin and UPLC stages) observed for each peptide standard assumes complete deuteration of the peptides.

peptide	primary sequence	*m/z*	max exchange backbone (Da)	retention time (min)	back exchange
MRFA	MRFA	524.5	2	6.2	24 ± 2.3%
angio II	DRVYIHPF	1046.5	5	7.3	29 ± 1.8%
Glufib	EGVNDNEEGFFSAR	785.8	12	7.6	33 ± 1.7%
bombesin	PQRLGNQTAVGHLM	810.4	11	8.0	17 ± 2.3%
ICB	FVNQHLCGSHLVEALYLVCGERGFFYTPKA	1165.6	27	9.0	33 ± 1.3%

2.3.1.3 HDX-MS measurements and rhGH structure prediction

Prior to making any structural comparisons from rhGH HDX-MS measurements, the optimised system and methods were validated by comparison of rhGH HDX-MS measurements with reported hGH protein structures [203, 205, 209, 216].

Figure 40 displays the percentage relative deuterium uptake (rDU) rates observed for the rhGH control, covering 93% of rhGH sequence. Rates are displayed in relation to the known secondary and tertiary structures (Figure 40 a and b respectively), as described in the XRC structure of free wild type hGH (PDB ID: 1HGU, [216]). Slower rDU rates were observed for all four α- helices present within the structure (Figure 40 a), residues 9-33 (helix A), 72-89 (helix B), 114-128 (helix C) and 155-186 (helix D), as expected due to the protection afforded by the hydrogen bonding in these structures. The more flexible linker and loop regions of the structure, such as residues 50-77 and 130-140, exhibit much faster HDX rates, with their lack of secondary structure

facilitating their participation in exchange. Figure 40 b maps the observed rDU, represented by a heat colour scale, as slow exchanging peptides displayed as cool (blue) colours and fast exchanging as warm (red) colours. Considering this, the effects of tertiary structure on rDU were apparent for the residues 15-25, 77-87, 114-117 and 165-177, positioned within the central cavity of the antiparallel four α-helical bundle of rhGH, known to be a central hydrophobic region of the protein. Even after 8 hours incubation minimal exchange occurs, due to decreased solvent accessibility as a result of the tight packing of the α-helical structures. Overall HDX dynamics observed here correlate well with previous studies, both HDX-MS [203–205] and NMR [209], which describe a folded, stable structure of rhGH at neutral pH values. These data demonstrate the utility of HDX-MS measurements alone, as valuable tool for protein structure prediction.

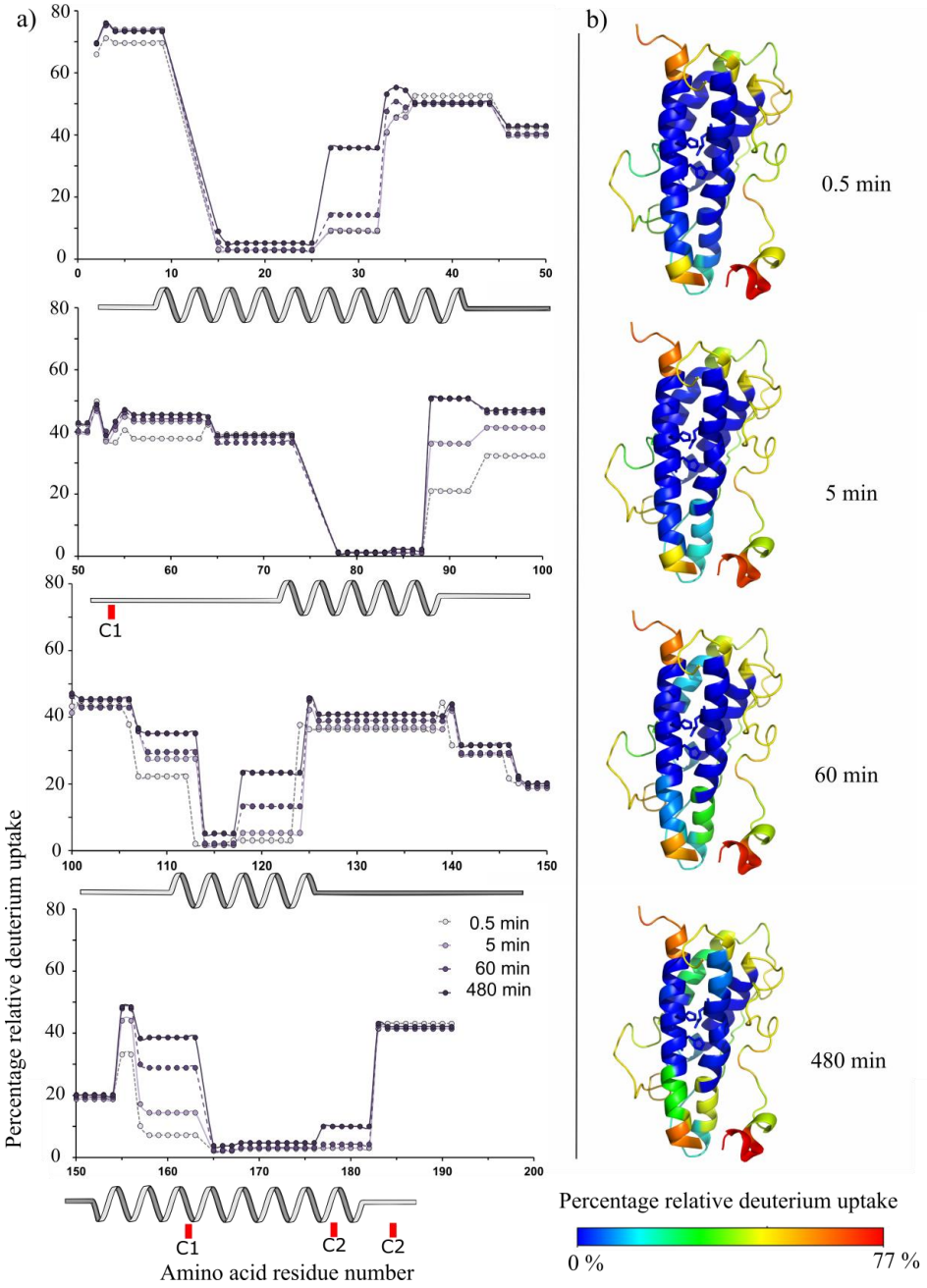


Figure 40. Relative deuterium uptake of the rhGH peptides analysed after 0.5, 5, 60, 480 min incubation periods plotted against a) the rhGH secondary structure and b) the rhGH tertiary structure as measured by XRC [216]. Cysteine bridges C1 and C2, are indicated. HDX-MS measurements are expressed as percentage relative deuterium uptake (rDU) values. Amino acidic rDU values are plotted according to the shortest localised peptide in the corresponding region. Where a residue is covered by multiple peptides the rDU for that given residue is defined by the shortest peptide out of those peptides covering that residue.

2.3.1.4 Reproducibility and intermediate precision assessment

To assess the repeatability and intermediate measurement precision (IMP) of the optimised HDX-MS methods and system developed here, triplicate within-day HDX-MS measurements of rhGH control samples (~ 0.5 min deuterium incubation) were made across a one year period. For each experimental dataset, rhGH samples, buffers, mobile phases and quench solutions were all prepared fresh. HDX-MS measurements across the year period encompass the use of three pepsin columns, four batches of rhGH reference material and three analysts that conducted both experimental and data analysis stages.

Considering the 41 peptides observed in all datasets, the repeatability and IMP of the rDU were calculated as the mean RSD of the rDU for each peptide within a day and over one year period respectively. The repeatability (n=3) was 0.6% and the IMP (n=33) was 2.1%. Table 7 displays the rDU, repeatability and IMP for each peptide. Comparison of these values with other published reproducibility data from different laboratories, albeit produced using difference protein standards, Table 8, demonstrates the robustness of the optimised platform and methods developed here [137, 138, 170, 223].

In addition, this in-house dataset was probed for the main sources of measurement variability. To visualise any correlation between rDU and precision, a beeswarm plot was generated, Figure 41, using the RSD for each peptide and rDU measurements. Peptides with higher RSD showed lower rDU for both repeatability and IMP, reflecting the effects of small variations in instrumental conditions, different protein conformational ensembles in different rhGH samples, temperature, pH and mobile phase fluctuations on HDX-MS measurements. Mobile phases were freshly prepared in a controlled environment (uncertainty, pH meter ± 0.01, balance ± 0.00030 g) and temperatures were controlled throughout the deuterium exchange experiments (temperature uncertainty, PAL ± 1 °C, laboratory ± 2 °C). Those factors were therefore considered to be unlikely causes of large measurement variability.

The contribution of chromatographic reproducibility and pepsin column integrity to back exchange [123] were then further evaluated as potential components affecting HDX-MS measurement variability. Peptide retention time (RT) reproducibility provides an indication of the varying exposure times of eluting peptides in repeat injections to protons in the mobile phase. If large enough this variability will

consequently affect back exchange rates and ultimate absolute HDX-MS measurements. RT RSDs within the one year period, for the 41 peptides considered were merely > 1% for 93% of rhGH peptides, suggesting that chromatographic conditions such as flow rate, mobile phases, column lifetime variability and chromatographic temperature were not the major components affecting measurement variability. These data also validate the usage of a 0.5 min window set in DynamX processing parameters, matching deuterated and undeuterated peptides in the peptide map based on RT, further increasing reliability of deuterated peptide identification and reducing need for manual data curation, particularly attractive when considering the industrialisation of the HDX-MS methodology.

On the contrary, fluctuations in back exchange were observed during the life time of the pepsin column, with sudden increases in back exchange for rapidly exchanging peptides after numerous injections. During the life time of each of the three pepsin columns used in this study, mean RSDs for the rDU across 41 peptides were 15%, 13% and 14% respectively, similar to the mean RSD of the 41 peptides across the year (13%) and all three pepsin columns. This confirms that the major source of variability on rDU of the measurements carried out here was within the pepsin column life time. Evaluation of the peptide maps generated across the lifetime of the column, showed average values of the rhGH sequence coverage and cleavage rate over the one year period were 95 ± 1% and 37 ± 3% respectively, indicating that any possible effects on deuterium uptake due to column degradation occurred prior to loss of proteolytic digestion. In light of this, a QC injection of rhGH incubated in deuterated solvent for t=30s was added to the beginning of every HDX-MS run, when large fluctuations > 10% of back exchange were observed for this sample the pepsin column was replaced for new.

Table 7. Measurement repeatability (Rep) and intermediate measurement precision (IMP) of the 41 rhGH peptides detected in all n=33 experiments carried out to calculate IMP over one year period. rDU is the average of the relative deuterium uptake percentage and repeatability and IMS expressed as the rDU standard deviation, for within-day n=3 and n=33 respectively.

Peptide residues	**rDU (%)**	**Rep (%)**	**IMP (%)**
[rhGH] 1-9	60.6	0.2	7.8
[rhGH] 2-9	63.0	2.0	9.5
[rhGH] 3-9	60.7	0.2	8.4
[rhGH] 15-25	2.3	0.2	0.6
[rhGH] 16-25	2.5	0.2	0.4
[rhGH] 26-32	8.5	0.1	0.7
[rhGH] 31-44	36.0	0.1	1.4
[rhGH] 32-44	41.0	0.6	1.2
[rhGH] 34-44	47.3	0.3	1.6
[rhGH] 35-44	53.5	0.9	1.5
[rhGH] 45-51	41.9	0.4	4.0
[rhGH] 45-52	48.8	0.5	2.9
[rhGH] 53-63	36.4	1.4	2.8
[rhGH] 53-73	42.2	0.9	2.9
[rhGH] 54-63	41.9	0.8	1.8
[rhGH] 54-73	42.4	1.4	3.1
[rhGH] 55-63	38.1	0.9	2.6
[rhGH] 55-73	43.5	0.7	3.0
[rhGH] 64-73	39.3	0.6	2.1
[rhGH] 93-100	32.0	0.7	1.8
[rhGH] 98-106	48.7	0.7	1.3
[rhGH] 98-110	43.8	0.4	2.1
[rhGH] 101-110	41.5	0.2	2.3

[rhGH] 102-110	38.1	0.1	2.2
[rhGH] 107-113	22.7	0.5	0.7
[rhGH] 114-124	1.6	0.2	0.6
[rhGH] 115-124	2.9	0.3	1.0
[rhGH] 117-124	2.6	0.1	0.8
[rhGH] 125-138	35.0	0.3	3.2
[rhGH] 125-139	36.1	0.5	2.4
[rhGH] 125-146	42.6	1.8	2.8
[rhGH] 140-146	34.7	2.1	3.2
[rhGH] 147-156	20.3	0.8	1.4
[rhGH] 147-157	20.8	0.5	1.1
[rhGH] 155-163	15.3	0.5	1.2
[rhGH] 157-163	6.2	0.5	1.2
[rhGH] 164-175	1.6	0.3	0.9
[rhGH] 167-176	3.4	0.5	0.6
[rhGH] 176-182	2.0	0.3	0.8
[rhGH] 182-191	40.7	0.5	1.8
[rhGH] 183-191	42.3	0.5	2.0

Table 8. Comparison of rhGH HDX-MS measurement Rep and IMP with published HDX-MS data [137, 138, 170, 223]. Proteins measured are abbreviated to the following: Bovine carbonic anhydrase (BCA), Equine cytochrome c (Cyt. C), Bovine ubiquitin (B. Ubiq) and Vitamin D receptor (VDR). Rep and IMP are repeatability and intermediate measurement procedure respectively expressed as the relative standard deviation of relative deuterium uptake for n replicates. N pep is the number of peptides considered in the study. Days is the study duration. N/A means that the information was not available

Protein	**Rep (n)**	**IMP (n)**	**N pep**	**Analysts**	**Days**
rhGH	0.6 (3)	2.3 (33)	41	3	360
BCA [138]	0.4 (4)	1.6 (26)	5	1	60
Cyt C [170]	1.4 (3)	N/A	52	N/A	1
B.Ubiq [223]	0.9 (20)	2 (60)	39	N/A	37
VDR [137]	N/A	3.9 (127)	34	N/A	240

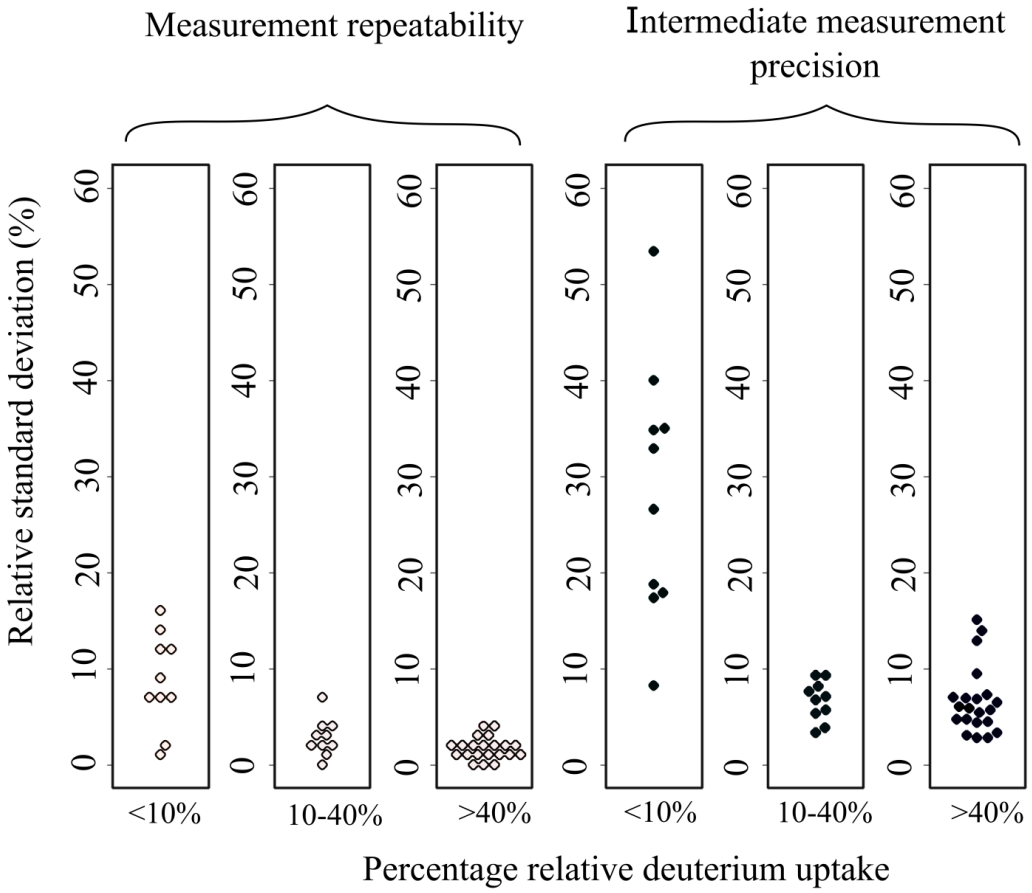


Figure 41. IMP of HDX-MS measurements for 41 rhGH peptides following 0.5 min incubation, visualised using a beeswarm plot. Each data point represents the relative standard deviation (RSD) percentage of individual peptide HDX-MS measurements and points are grouped accordingly to the relative deuterium uptake of each peptide.

2.3.1.5 Establishing significance thresholds from differential vial-to-vial study

A statistical differential analysis approach was developed to define rhGH structural changes induced by zinc and establish thresholds to discriminate significant structural changes from measurement variability [224]. This approach builds on the differential analysis method developed by Houde et al.[112] with a modification made to avoid a perceived erroneous calculation of the measurement error and to identify individual thresholds for each individual peptide in contrast to establishing a single global confidence limit, an approach which has been demonstrated to give both under and over estimations of error [168].

To enable comparability studies and define thresholds, first a vial-to-vial comparability study was conducted using the rhGH control [112, 168]. For this comparability study, a rhGH control sample was diluted in buffer, allowed to equilibrate at room temperature, the sample split into two vials and analysed concurrently by HDX-MS using the optimised method described. Replicate vial-to-vial comparability studies were conducted across three repeat relative experiments performed on three different days (n=9). Peptides were given a number, i, based on the computed position in the rhGH amino acid sequence of the peptide's mid-point [112].

The differential analysis approach developed was first applied to the vial-to-vial comparison, to establish measurement thresholds, above and below which structural changes can be confidently defined, and then applied to the HDX-MS measurements to determine structural differences between the rhGH control (ref) and the rhGH:Zn 1:2 and 1:10 samples (exp).

HDX-MS measurements of different rhGH control vials were obtained for each "n" rhGH control and only those peptides common to all datasets were considered. The data for each peptide were treated as follows:

For an experiment carried out in a single experiment on one day, the absolute differential deuterium uptake at a time point, t, is

Equation 32

$$D_t = \left| M_{\mathrm{exp},t} - M_{\mathrm{ref},t} \right|$$

where $M_{exp,t}$ and $M_{ref,t}$ represent the mean of triplicate measurements of the deuterium uptake at time, t, for the experimental and reference samples respectively. Mean values are used because there is no correlation between samples at the repeatability level. The precision of D_t depends on the level of replication, and the figures provided are for triplicate measurements. If the experiment is repeated over a number of days, this will result in a set of values of D_t for each peptide. The correlation introduced by day-to-day variation is implicitly taken into account by calculating D_t separately for each day.

For differential analysis, $M_{ref,t}$ represents the reference sample (e.g. rhGH control) and $M_{exp,t}$ represents the perturbed sample (e.g. rhGH:Zn 1:2). In the context of the vial-

to-vial comparability, reference and experimental states were equal to rhGH control vial 1 and vial 2 respectively and are interchangeable between vials.

The purpose of using absolute values is to ensure that if both destabilising and stabilising dynamics (variability in the context of a comparability study) occurred across the course experiment for an individual peptide, this would not be concealed in final differential thresholds.

The absolute difference values at each incubation time point, t, are added together to produce a differential sum;

Equation 33

$$S = \sum_{t} D_t$$

This is done separately for each day, producing a set of values of S, the variability of which represents the intermediate precision of the differential sum.
A threshold value, S_L, can then be defined using the mean and standard deviation of S, across the multiple datasets produced, and used to define a limit above which structural changes can be differentiated from measurement variability:

Equation 34

$$S_L = \bar{S} + s(S)$$

Considering only those peptides, identified in all n data sets, including a final rhGH control vs each of rhGH:Zn (1:2 and 1:10) differential study (n = 5 datasets, 10 samples), measurement thresholds, represented by the dotted lines in Figure 42 a and b, were established for 53 (96.3% sequence coverage, 211% redundancy) rhGH peptide. For the differential analysis of rhGH vs rhGH:Zn, Equation 32 and Equation 33, were applied to each dataset. All raw differential measurements are detailed in Table A-25, located in the Appendices.

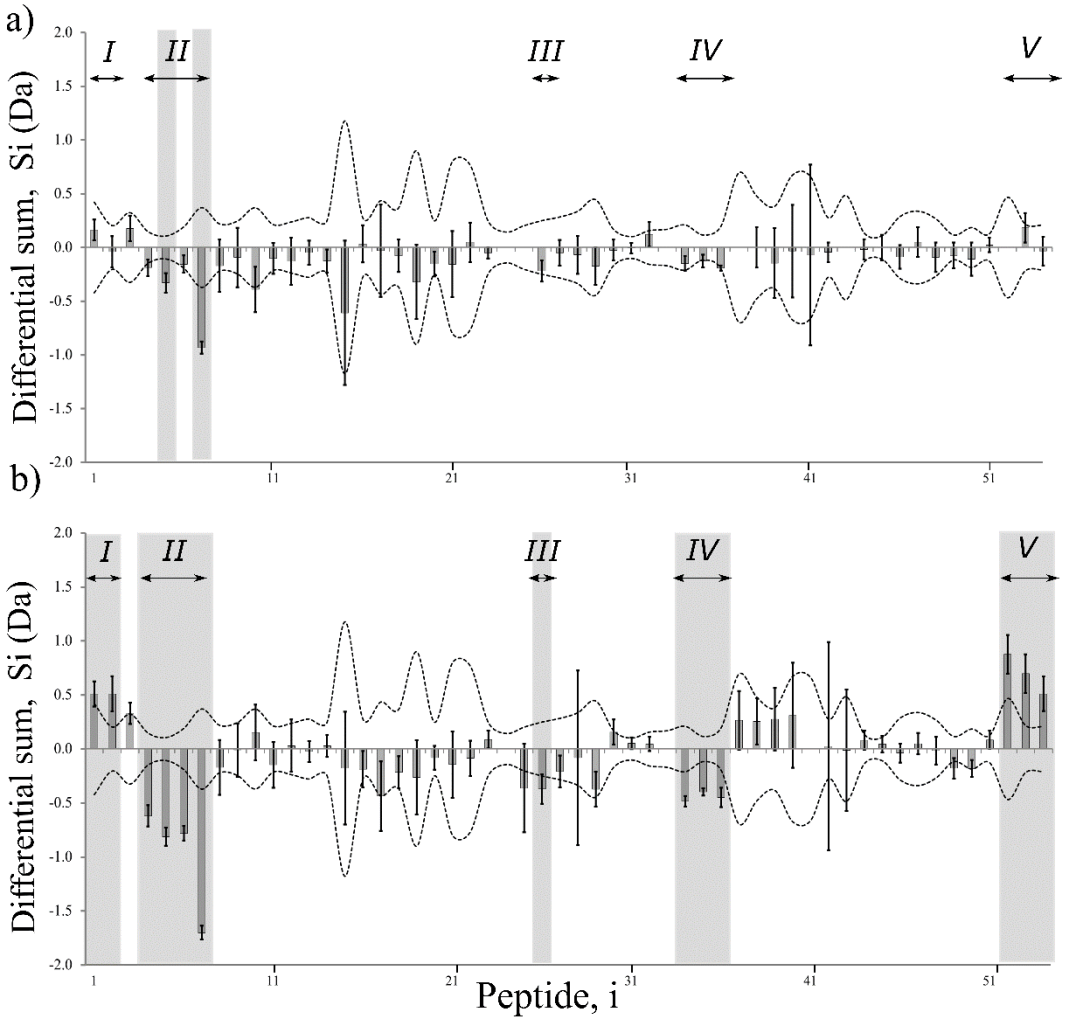


Figure 42. Differential HDX-MS measurements (dark grey bars) between rhGH samples. Dotted lines represent the positive and negative individual peptide threshold limits (S_L) as calculated from vial-to-vial study of rhGH control vs rhGH control a) rhGH control and rhGH in the presence of 2 molar equivalents of Zn^{2+} (1:2) and b) rhGH control and rhGH in the presence of 10 molar equivalents of Zn^{2+} (1:10). Those peptides that are considered to demonstrate a significant structural change by exceeding the calculated threshold limits are highlighted in shaded grey and labelled as regions I, II, III, IV, V.

2.3.1.6 HDX-MS measurements to identify HOS changes induced by zinc

Relative HDX-MS experiments were carried out to compare the structure of rhGH in the absence and presence of zinc, and differential analysis performed (applying Equation 32 and Equation 33). Differential values were compared with thresholds, S_L,

calculated from the rhGH control vial-to-vial, allowing structural changes from measurement variability.

Figure 42, shows the differential sums for each peptide calculated for the comparison a) rhGH and rhGH:Zn 1:2, and b) rhGH and rhGH:Zn 1:10 . Table 9 summarises those peptides identified as showing significant changes due to the presence of zinc at two and ten molar equivalents.

Table 9. rhGH peptides identified as showing significant changes due to the presence of zinc, by HDX-MS differential analysis. Changes are expressed as HDX increases (+) and decreases (-) relative to rhGH control, with no significant change is indicated by blank input.

Identifier			**Relative HDX change**	
Peptide, i	**Peptide ID**	**Sequence**	**Zn 1:2**	**Zn 1:10**
1	1-9-	FPTIPLSRL		-
2	2-9-	PTIPLSRL		-
4	14-25-	MLRAHRLHQLAF		+
5	15-25-	LRAHRLHQLAF	+	+
6	16-25-	RAHRLHQLAF		+
7	26-32-	DTYQEFE	+	+
26	93-100-	LRSVFANS		+
34	114-124-	LKDLEEGIQTL		+
35	115-124-	KDLEEGIQTL		+
36	117-124-	LEEGIQTL		+
51	178-191-	RIVQCRSVEGSCGF		-
52	182-191-	CRSVEGSCGF		-
53	183-191-	RSVEGSCGF		-

Two peptides exceeded the threshold limits localised for rhGH:Zn 1:2 sample comparison, indicating structural changes, corresponding to the residues 15-25 and 26-32 (region II in Figure 42). Both of these peptides are located in the helix A of rhGH with peptide 15-25 including the residues His18 and His21, that are two of the three residues reported to be involved in the Zn^{2+} ion binding to rhGH [208, 213]. Increased rates of deuteration for both these peptides suggest that Zn^{2+} binding disrupts the hydrogen bonding of helix A, possibly shifting the orientation of the

hydrophobic α-helix region outwards and increasing the incidence of deuterium exchange. Although the two peptides demonstrate a continuation of the same primary sequence, the magnitude of increased deuteration rates is highest for the peptide corresponding to residues 26-32, ~ -0.9 Da, suggesting that long range structural changes, positioned away from the zinc binding site, are greater than those correlating to zinc binding interactions alone. This further promotes an idea of the orientation of the hydrophobic α-helix region outwards or disruption of the hydrogen bonding between helix A and D.

By increasing the concentration of zinc to obtain a ratio rhGH:Zn equal to 1:10, thirteen peptides exceeded the threshold limits defined for rhGH. The structural differences identified for the rhGH 1:2 were confirmed and further differences were identified (Figure 42 b, Figure 43 c and Table 9). Particularly the structural changes observed in region II for the rhGH:Zn 1:2 sample were confirmed and two additional peptides at residues 14-25 and 16-25 also exhibit structural changes.

In addition, a decrease in deuterium uptake in regions I and V was observed at earlier incubation periods (0.5 and 5 min) suggesting further long range structural changes. For region I, considering overlapping peptides and loss of two N-terminal amides during back exchange [125], data suggest that the structural changes were located in residues 5-9 and for region V, a larger magnitude of change is observed for peptide 178-191. Considering the magnitude of decreased deuterium rates for the three overlapping peptides 178-191 (~ 0.88 Da) , 181-191(~ 0.70 Da) and 182-191(~ 0.50 Da), this would tend to suggest the structural changes to be occurring across both the N- and C-terminal regions of the 180-191 sequence.

A possible cause for these decreased HDX rates at the N and C-terminal could be an outward movement of helix A upon zinc binding orientating the two termini closer together and therefore decreasing their solvent accessibility. Region III corresponding to residues 98-106, located in the loop region between helix B and helix C, also exhibits a very small, ~ 0.2 Da, increase in deuteration uptake at the early incubation points. Finally, region IV shows increased deuteration rates in the presence of zinc. Two overlapping peptides in this region, 115-124 and 117-124, show an increase in uptake from residue 115 to 117, localising the structural changes observed to residues 119-124 of helix C. All regions identified as showing structural changes due to the

presence of zinc appeared to exhibit EX2 kinetics in both the absence and presence of zinc suggesting stable folded conformations.

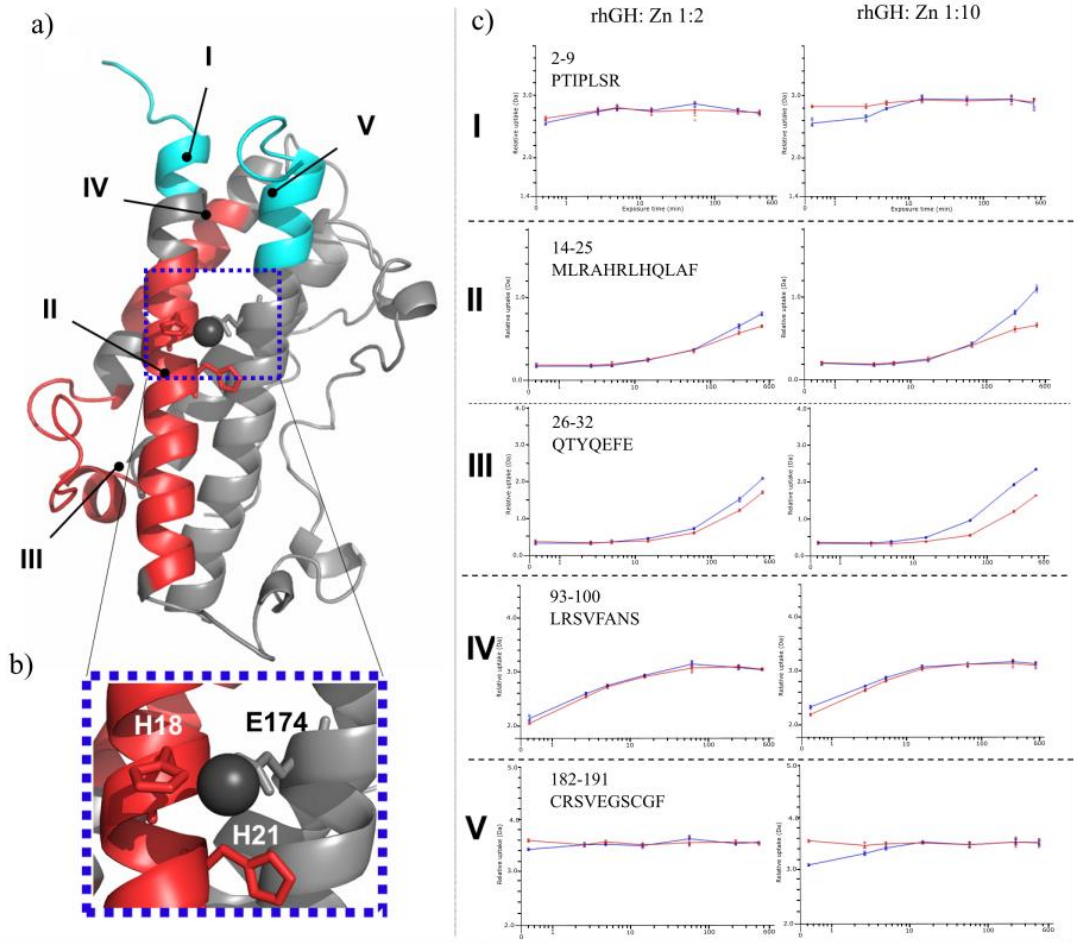


Figure 43. Structural changes of rhGH induced by zinc. a) XRC of rhGH in the presence of zinc (adapted from PDB 1BP3 [208]). Regions of significant structural change due to the presence of zinc as identified by HDX-MS differential analysis are highlighted in terms of deuterium uptake in presence of zinc, as reduced uptake (blue) and increased uptake (red). Labelled regions include the following residues: region I [1-9], II [14-32], III [98-106], IV [114-124], V [178-191]. b) Zinc binding residues His18, His21 and Glu174 are illustrated by stick representations. c) Uptake plots of deuterium uptake (Da) against incubation time (min) are displayed for five peptides located in those regions of structural change.

The benefits of defining individual thresholds for each peptide were demonstrated by comparing the results here discussed with the results obtained by using the approach described by Houde et al. [112] and the DynamX software approach utilised by many

users for data analysis, respectively. Applying the approach described by Houde et al. [112] a single unique threshold for all peptides was calculated as the standard deviation of the mean differential sum value across all peptides, corresponding to ± 0.98 Da. When this threshold was applied no structural changes were observed for the rhGH:Zn 1:2 comparison and only one difference was observed in the rhGH:Zn 1:10 comparison (peptide 26-32). Hence this approach would lead to an over estimation of measurement variability for many peptides and therefore reduce the sensitivity of the differential approach to detect small but statistically significant structural changes.

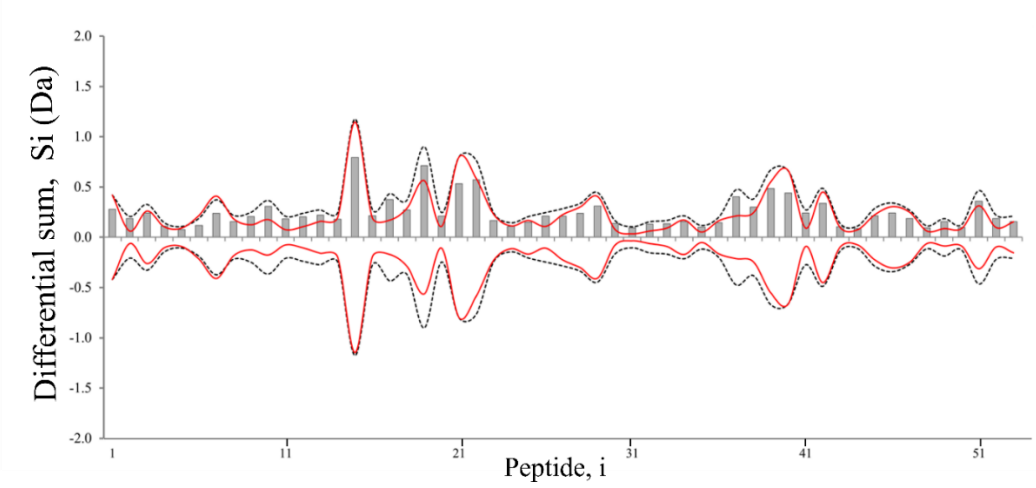


Figure 44. Plot of mean differential HDX-MS measurements (dark grey bars) from a rhGH vial to vial study. Individual peptide threshold limits are plotted as calculated using a 3σ criterion approach (red line) and in-house approach (dotted line) to differential vial-to-vial measurements

A generic approach to establishing differential data analysis measurement threshold limits can be taken using the software DynamX based on a "3σ", where σ represents SD of measurements. The thresholds limits calculated by applying the 3σ approach and the threshold approach developed here, are represented in Figure 44 . Those regions identified as showing a structural change demonstrated high similarity when comparing the 3σ approach with the statistical method here developed, with general trends of measurement variability consistent across the two methodologies. However for a number of peptides, the 3σ criterion underestimated the measurement variability, probably due to the fact that through the DynamX approach, it is only possible to compare data generated from an intra-day comparison. Furthermore the proprietry calculations performed by DynamX limits its use in combining data from different analyses.

2.3.2 Size exclusion chromatography

The presence of Zn^{2+} has been indicated previously as inducing dimerisation of hGH [213]. To rule out any overinterpretation of HDX-MS measurements as a result of such aggregation, SEC-UV experiments were carried out to monitor for the potential formation of rhGH aggregates in the presence of increasing concentrations of zinc (1:2, 1:10 and 1:20; rhGH:Zn). A bovine serum albumin (BSA) sample was also run in parallel, to act as a standard for differentiating between different sized species, using BSA monomers and dimers observed.

Three distinct peaks are observed in all the SEC-UV chromatograms for rhGH and 1:2 and 1:10 samples (overlaid in Figure 45), and were attributed to a rhGH dimer(~ 7.9 min), a rhGH monomer (~ 8.5min) and excipient present in the samples (10.2 min). The small dimeric peak eluting prior to the main monomeric form of rhGH was observed in all samples, and quantified as corresponding to ~ 0.5% of total rhGH, regardless of the absence and presence of zinc (Figure 46). The consistent percentage of dimers in the samples, even at the elevated 1:20 ratio, indicates that increasing amount of zinc under these operative conditions doesn't induce formation of aggregates that may bias the HDX-MS results.

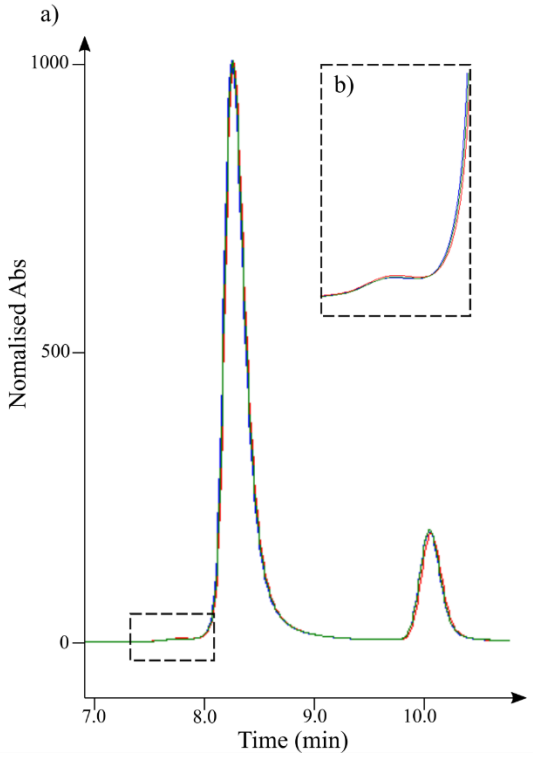


Figure 45. a) Overlay of SEC-UV chromatograms of rhGH containing samples. rhGH control, rhGH:Zn 1:2 and rhGH:Zn 1:10 (blue, red and green traces respectively) rhGH dimer, rhGH monomer and sample excipient peaks were observed across all samples, with retention times of 7.9, 8.5 and 10.2 min respectively. b) Zoom of dimeric peak.

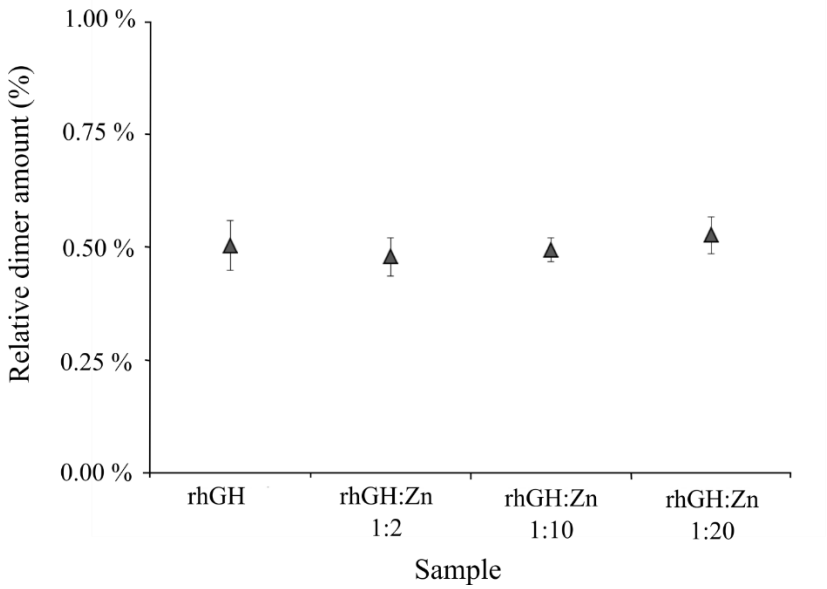


Figure 46. Relative percentage of dimer present in rhGH samples as calculated by SEC-UV.

2.3.3 Native IMS-MS

Unlike HDX-MS measurements which are representative of solution average of conformations present in a sample, one of the key attractions of native-IMS-MS is the technique's ability to measure multiple isobaric copopulated protein populations that differ in physical size or shape. Native IMS-MS measurements were applied to the rhGH model system, providing a complementary dataset to HDX-MS measurements and assessing the applicability of the technique to detect the small structural changes observed for the candidate model system.

2.3.3.1 Sample buffer optimisation

Several experimental conditions and instrumental parameters were optimised for IMS-MS measurements of rhGH. Two differing buffer conditions, 20 mM and 200 mM ammonium acetate, were assessed, representing relatively low and high levels of buffering concentrations. Native IMS-MS spectra obtained under these conditions are presented in Figure 47.

Both spectra exhibit similar CSDs at both high and low buffering conditions, suggesting stable and "native-like" conformations are retained [4, 76, 82, 83]. Increased peak broadening is observed under higher buffering conditions, observed particularly at the +8 base peak, with the 20 mM ammonium acetate buffer giving a more optimal and cleaner spectrum with less peak broadening. Higher concentrations of buffer ions offer greater protection to the protein during ionisation, yet cause beak broadening due to protein binding and ultimately decreased signal to noise of the spectra. It is therefore advantageous to optimise these buffering conditions to ensure that native like structures are maintained whilst simultaneously maximising signal to noise and spectra quality.

Additionally, the lower molarity buffering conditions of 20 mM ammonium acetate are more analogous to the 10 mM KPBS used in HDX-MS experiments making drawing parallel conclusions across platforms more representative. For all further IMS-MS experiments, samples were introduced in 20 mM ammonium acetate buffer.

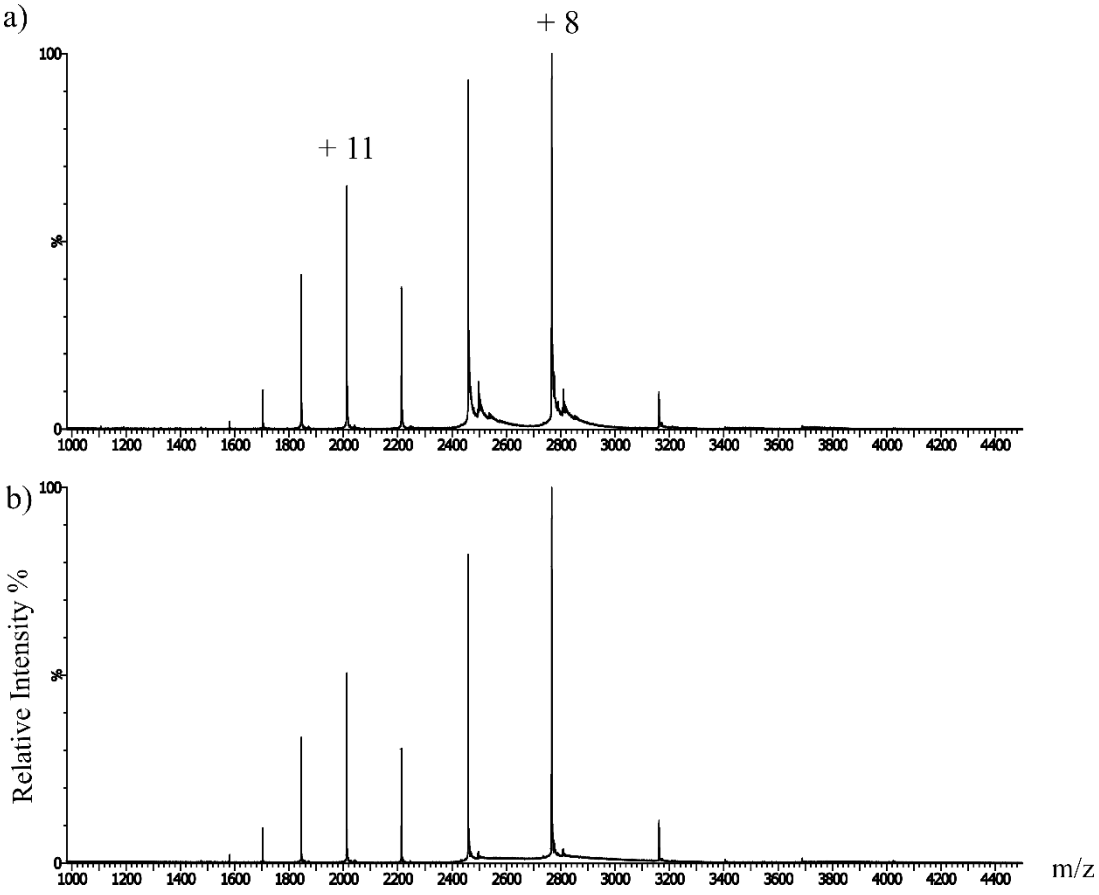


Figure 47. Native IMS-MS spectra of rhGH in a) 200 mM and b) 20 mM ammonium acetate.

2.3.3.2 Native IMS-MS measurements and rhGH structure prediction

The native IMS-MS spectra of the rhGH control (Figure 48) revealed multiple (~3) Gaussian CSDs, indicating the presence of multiple conformations of rhGH. The most abundant distribution is centred around the +8 charge state, with another centred around the +11 charge state, suggesting the presence of a compact and extended monomeric conformation for each respectively. A final, much smaller dimer distribution is also observed centred around the higher ~3400 *m/z* range.

Analysis of the mobilogram for the rhGH control (Figure 49), confirms the presence of a small dimer conformation (+12 to +14, Figure 49 i) plus compact monomer (+7 to +11, Figure 49 ii) and extended monomer (+10 to +13, Figure 49 iii) conformations at lower and higher charge states respectively. Additionally, a final very small distribution of a further extended conformation (> +14) was observed in the t_D measurements that was not observed in native IMS-MS spectra, however the signal-to-noise of the mobility measurements meant that data quality is poor and extracting

t_Ds unreliable. Overall the relatively narrow CSDs observed are consistent with a globular-like folded protein conformations [77].

These observations are also in accordance with previous published IMS-MS measurements of rhGH [206] and solution based studies [225] which identify the existence of > 1 folded monomeric populations of rhGH, that differ in tertiary structure whilst maintaining a stable secondary structure [226]. The observation of an extended monomeric conformation for charge states in excess of +11, is in accordance with the calculation (Equation 31) that the maximum charge that can be theoretically carried by a protein of rhGH's size, ~ 22 kDa, as equal to 12 [77].

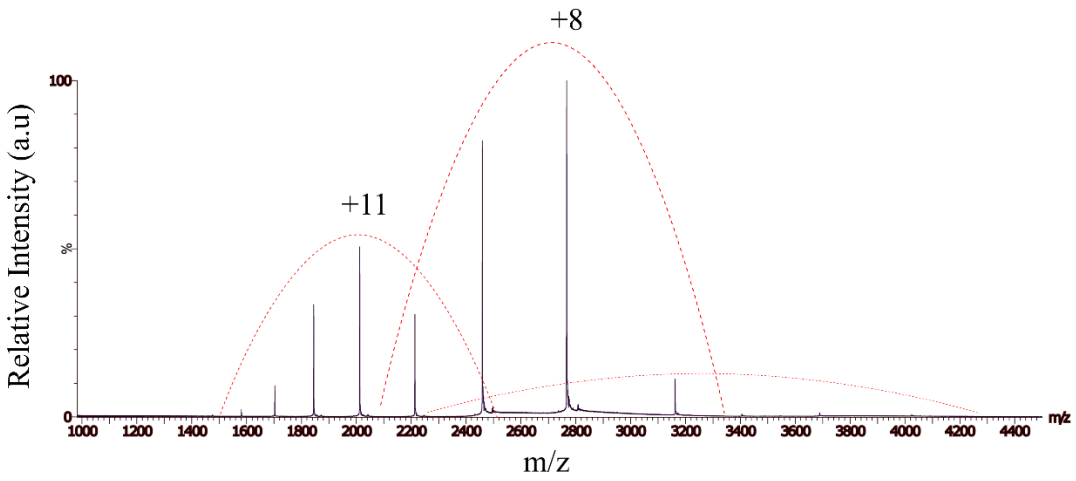


Figure 48. Native IMS-MS spectra of rhGH control. Multiple Gaussian CSDs suggest the presence of multiple conformations are indicated by the red dotted lines.

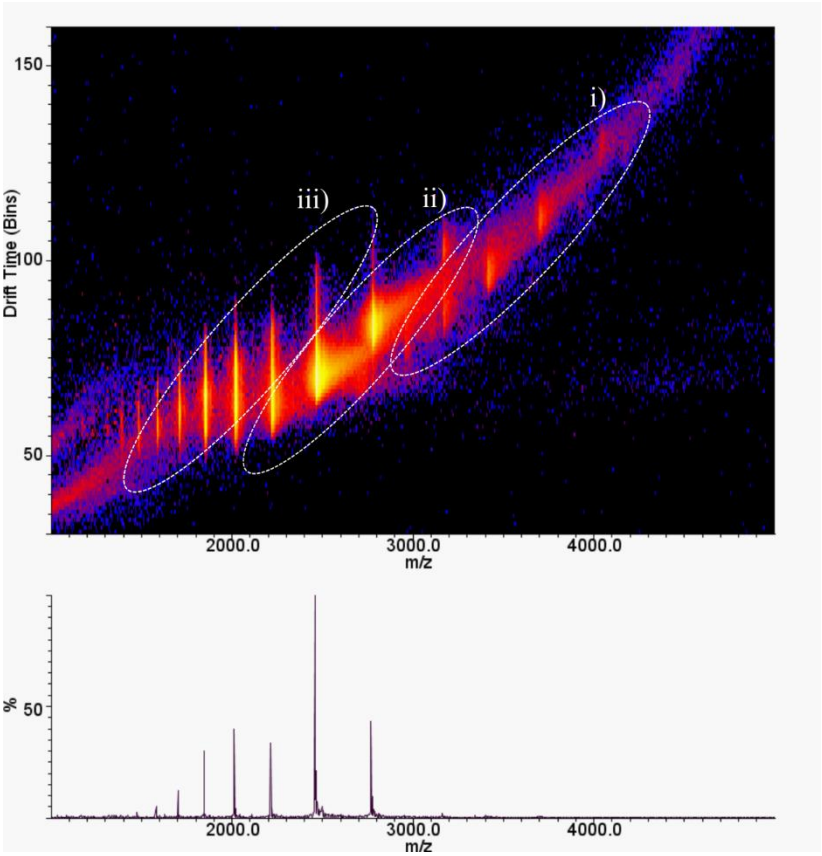


Figure 49. Mobilogram of rhGH control IMS-MS measurements. Drift time measurements reveal multiple conformations and are highlighted: i) dimer, ii) compact monomer, and iii) extended monomer species.

2.3.3.3 IMS-MS reproducibility

To assess the measurement variability of the optimised IMS-MS method, triplicate (within-day) repeat measurements of the rhGH control made across a series of three days (intra-day, IMP) were considered. For each given day a fresh aliquot of rhGH was removed, desalted and buffer exchanged into the optimised buffer.

Observed t_D measurements were assessed in terms of within-day repeatability and IMP, by considering the RSDs of average measurements. These values are expressed in Table 10 . Very high reproducibility was observed for both within-day repeatability (0.1-1.1% RSD) and IMP (0.0-1.8% RSD) measurements demonstrating both the IMS-MS measurements to be reproducible but also the conformations present in the rhGH sample to be consistent across the time span of the experiments and preparations.

Table 10. Measurement variability of IMS-MS measurements of rhGH control. Average observed t_D and RSDs for compact monomer, extended monomer and dimeric species are expressed for triplicate, n=3, intra-day (Rep) and inter-day (IMP) measurements.

			Rep		IMP	
	Charge state (z)	***m/z***	**t_D**	**RSD**	**t_D**	**RSD**
compact monomer	7	3160.7	14.0	0.3%	14.0	0.0%
	8	2765.6	11.2	0.2%	11.2	0.3%
	9	2458.3	9.4	0.2%	9.4	0.3%
	10	2212.5	8.6	0.4%	8.7	1.2%
	11	2011.3	8.5	0.3%	8.4	0.5%
extended monomer	10	2212.5	9.9	1.1%	10.2	1.8%
	11	2011.3	9.5	0.4%	9.7	0.3%
	12	1843.7	8.6	0.5%	8.6	0.0%
	13	1701.9	8.2	0.1%	8.2	0.2%
Dimer	12	3687.5	15.0	0.1%	15.0	0.3%
	13	3403.8	13.1	0.4%	13.1	0.1%
	14	3160.7	12.0	0.2%	11.9	0.4%

2.3.3.4 $^{TW}CCS_{N2}$ measurements

Conversion of t_D to $^{TW}CCS_{N2}$ measurements is common practice, facilitating the comparison of ions gas phase structures within theoretical structures or structures measured by XRC and also allowing for direct lab-to-lab comparison of IMS-MS measurements.

$^{TW}CCS_{N2}$ values for average IMP IMS-MS measurements described in Table 11 were calculated post-acquisition as described in 2.2.6.4. Taking an average of CCS across the charge states observed for a given species, $^{TW}CCS_{N2}$ measurements of rhGH are as follows ; rhGH compacted monomer (2519 ± 303 Å^2), rhGH extended monomer (3216 ± 271 Å^2) and rhGH dimer (3957± 179 Å^2).

Uncertainty was also calculated for each measurement including a contribution from the calibration line established. These uncertainty values appear relatively high for these types of measurements [206] and given the very high reproducibility of the extracted t_D measurements the main source of uncertainty of the TWIMS-MS measurements performed here is the contribution from the calibration.

For $^{TW}CCS_{N2}$ calculation, calibrants of known $^{TW}CCS_{N2}$ values are selected to bracket the molecular size of the compound of interest and measurements performed under the native conditions of the compound. The availability of such calibration values however is limited [52, 227–229]. A potential cause for the poor calibration fit could be the source of the calibrant $^{TW}CCS_{N2}$ values. The $^{TW}CCS_{N2}$ calibrant reference values used here, as reported by Bush et al. [52], where obtained using a modified earlier generation Synapt G1 TWIMS instrument of the, in which the TWIMS drift tube was replaced with a linear drift tube. Although the Synapt G2Si instrument used here is of similar instrument lineage, additional components such as the step-wave ion guide [30], mean the gas phase ions will experience non-identical voltages to those described by Bush et al. [52], which were acquired on an earlier generation instrument. This is corroborated by the fact that the $^{TW}CCS_{N2}$ measurements performed here for rhGH are consistently ~20% larger than previously published values which were performed using a Synapt G1 instrument [206]. This raises questions about the appropriate use of these calibration reference values for non-identical TWIMS instrumentation, which is beyond the scope of this thesis. In addition, recently, it has also been reported that space charge effects due to concentration deviations, can affect peak broadening and measured ATD peak tops, ultimately affecting the quality of CCS calibration in TWIMS [92, 230].

This is in accordance with the observation by Ferguson 2016 et al. [49], that CCS is a less sensitive metric for comparison of TWIMS measurements than t_D alone, similarly highlighting the need for better calibrants, or for HOS measurements to be performed relatively purely on t_D based metrics. Considering this, comparison of observed t_D was deemed more appropriate for HOS comparison across differing protein samples (rhGH and rhGH:Zn) in this work, aiding increased sensitivity towards smaller structural changes.

Table 11. $^{TW}CCS_{N2}$ values calculated for compact monomer, extended monomer and dimeric species of rhGH. The uncertainty of each measurement (U) is also given and expressed as a relative %.

			rhGH	
	Charge state	***m/z***	**$^{TW}CCS_{N2}$ (Å^2)**	**± U (rel%)**
compact monomer	7	3160.7	2170	343 (16%)
	8	2765.6	2321	267 (12%)
	9	2458.3	2481	162 (7%)
	10	2212.5	2682	221 (8%)
	11	2011.3	2940	239 (8%)
extended monomer	10	2212.5	2794	188 (7%)
	11	2011.3	3057	229 (8%)
	12	1843.7	3217	249 (8%)
	13	1701.9	3436	320 (9%)
Dimer	12	3687.5	3783.2	585 (16%)
	13	3403.8	3949.6	535 (14%)
	14	3160.7	4140.5	495 (12%)

2.3.3.5 IMS-MS measurements to identify HOS changes induced by zinc

Though structural changes were observed due to the presence of zinc using HDX-MS measurements, rhGH binding to zinc was only inferred. Comparison of spectra obtained from native IMS-MS measurements performed of rhGH control, rhGH:Zn 1:2 and 1:10 (Figure 50) confirm zinc binding to rhGH using the preparative conditions described.

Although a 1:1 binding stoichiometry (rhGH:Zn) was observed for the rhGH:Zn 1:2 sample, a significant portion of rhGH appears to be unbound (Figure 50 b ii). In contrast, both stoichiometric ratios of 1:1 and 1:2 are observed for the rhGH:Zn 1:10 sample, with a reduced presence of unbound rhGH, and the 1:1 species the most abundant, suggesting a greater proportion of bound rhGH-Zn. These findings are consistent with the reported Kd of ~ 1uM for the 1:1 rhGH:Zn complex and of positive co-operative binding of a secondary zinc molecule, of which no Kd is reported.

In order to rule out non-specific binding of Zn^{2+} to rhGH as a result of the ionisation process [231], a rhGH sample containing ten molar equivalents of Mg^{2+} (rhGH:Mg 1:10), an analogous small covalent bi-anion, was also analysed by native IMS-MS. Though some non-specific binding of Mg^{2+} was evident spectra revealed a greatly

reduced binding affinity of the Mg^{2+} compared to Zn^{2+} (Figure 50 b iii and iv) confirming the specific binding of zinc.

Comparison of t_D measurements showed no significant differences between rhGH and rhGH-Zn bound samples (both 1:2 and 1:10). This would suggest that under the conditions used here, there is no overall large differences in globular structure as a result of the binding of zinc and that the tertiary structure of the protein is largely consistent. Figure 51 illustrates the overlaid extracted mobilograms for each rhGH and rhGH-Zn charge state, these in turn are separated into relative conformer groups of dimer (+12 to +14), compact monomer (+7 to +10) and extended monomer (+11 to +13).

Considering the extracted mobilograms for each individual charge state, it is possible to observe multiple conformations for some charge states are evident and separated by IMS (Figure 51, D11+, M7+, M8+, M10+ and M11+). Comparison of the relative abundancies of these conformations reveal a predisposition towards the more folded monomeric conformations (Figure 51, M8+ and M10+ and M11+) and dimeric species (Figure 51, D11+, M7+) for Zn^{2+} containing samples relative to the rhGH control. This data suggests a shift towards more compact and dimeric species due to the presence of Zn^{2+}, which could arise due to a potential stabilisation in HOS by Zn^{2+}. This idea is supported by the correlation of the amount of Zn^{2+} present in the samples and the relative shift towards more compacted species, rhGH:Zn 1:10 > rhGH:Zn 1:2 > rhGH control, illustrated by the overlaid traces in Figure 51. Though no significant differences in t_D were observed across the samples this potential stabilisation in HOS was explored further using CIU-IMS-MS experiments.

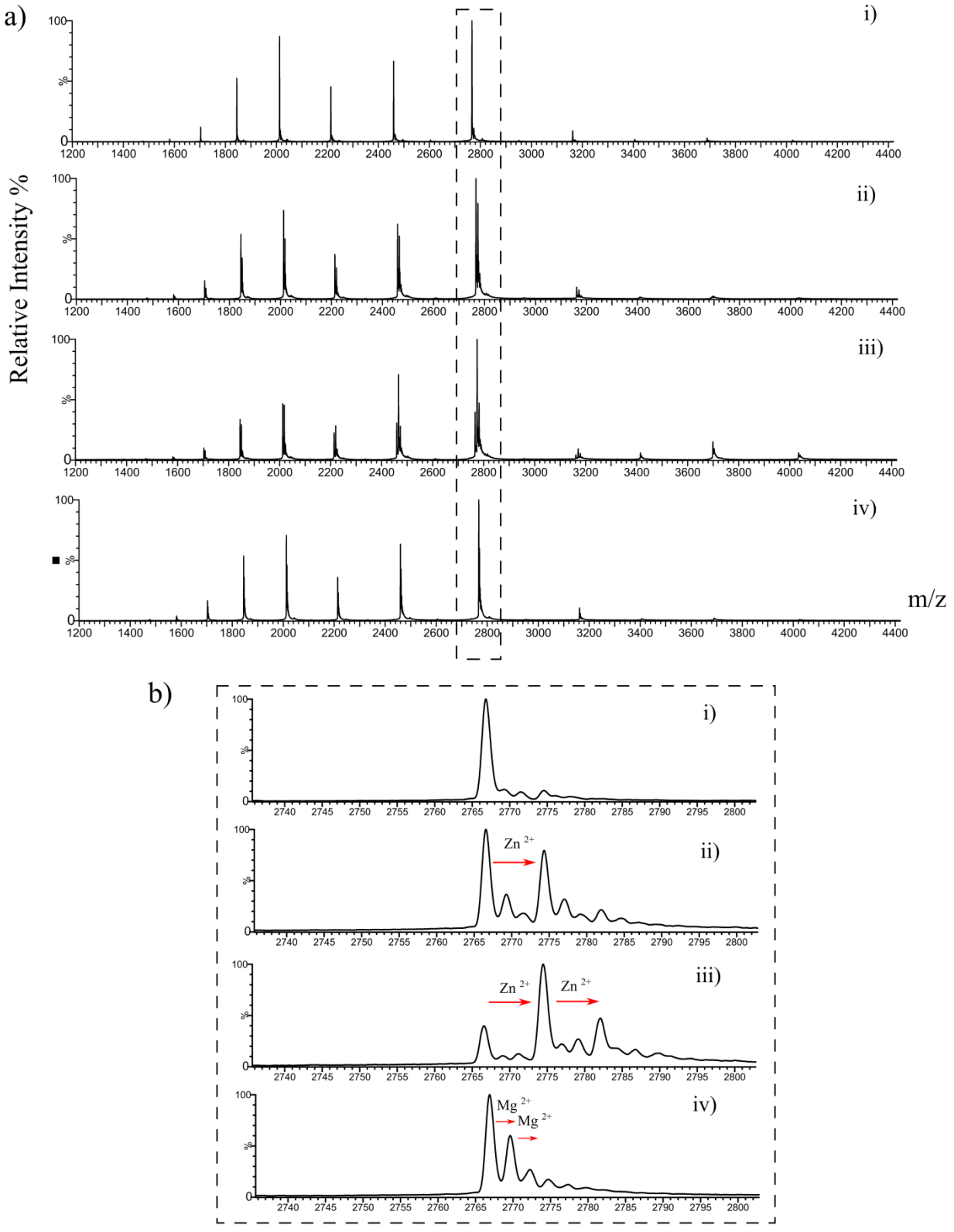


Figure 50. a) Native MS spectra of rhGH samples. i) rhGH, ii) rhGH:Zn 1:2, iii) rhGH:Zn 1:10, iv) rhGH Mg 1:10, b) zoom of +8 base peaks of rhGH and rhGH-bound species.

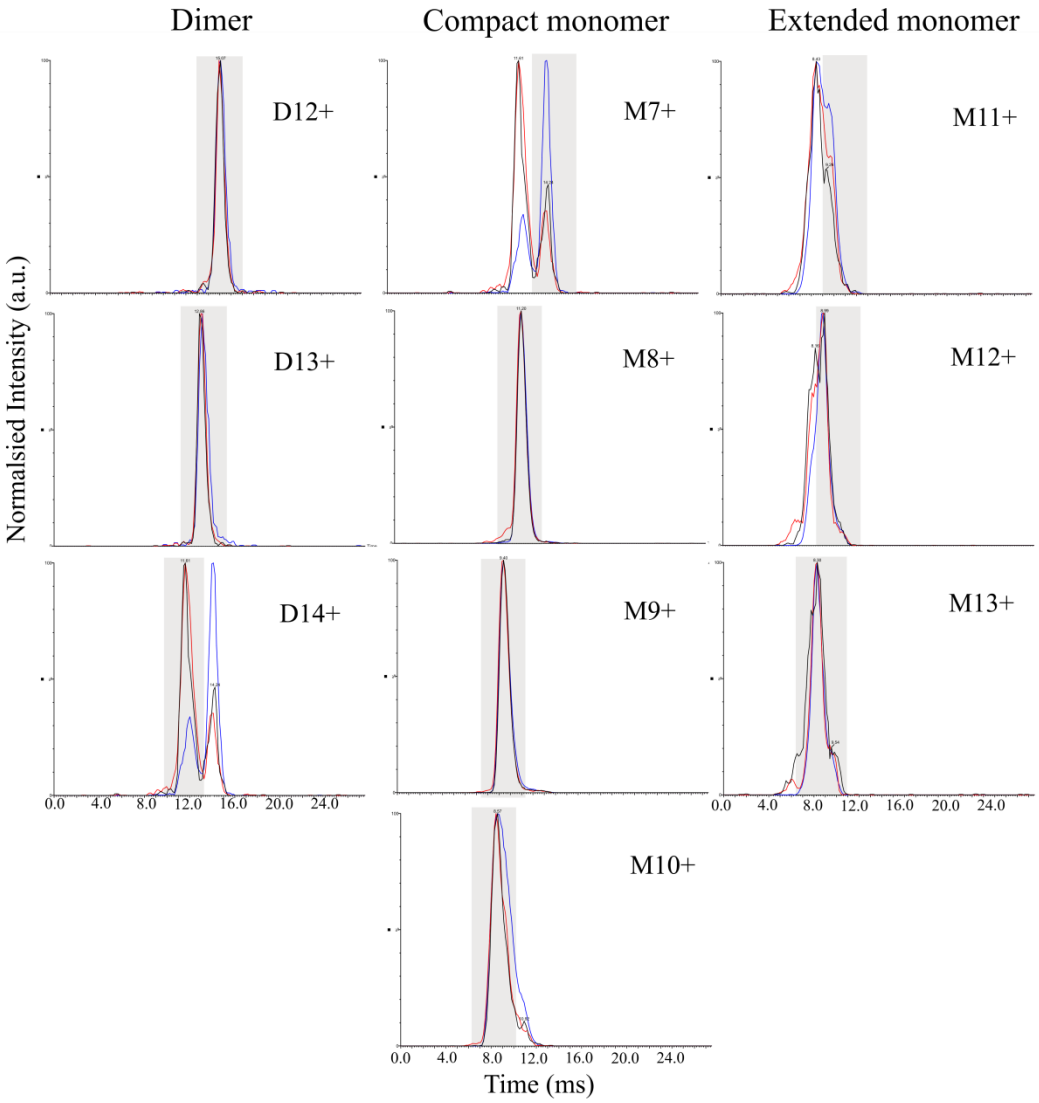


Figure 51. Overlay of extracted mobilograms associated with each species identified in IMS-MS measurements of rhGH control (blue), rhGH:Zn 1:2 (black) and rhGH 1:10 (red) samples. Each mobilogram is annotated as either dimer (D) or monomer (M) and the charge state indicated. Grey panels highlight the arrival time distribution of the species/charge state indicated.

2.3.3.6 CIU-IMS-MS comparisons

Previous studies [49, 95, 232] have illustrated that CIU-IMS-MS experiments are more sensitive than CCS measurements alone for comparing and differentiating between proteins with small differences in HOS. In light of this, CIU-IMS-MS experiments were perfromed on rhGH and rhGH:Zn samples. To avoid any bias associated with charge-state selection, the +8 charge state was selected for all samples (~ 2765.5 *m/z* for rhGH and 2773.6 *m/z* for the rhGH-zinc bound species for zinc

containing samples) and the trap voltage increased at 2V increments across the 6 – 38V range.

Similar unfolding pathways were observed across all three samples during the CIU experiment, in which at low collision voltages ~ 6 – 18 V, a single Gaussian ATD was observed, indicating a native folded state (T1‡, Figure 52 a-c) with a t_D ~ 11.3 ms. As voltages are increased a CIU event occurs, indicated by an additional species with an increased t_D, T2‡ ~ ~ 13.6 ms. The CIU event appears to proceed via a transitional voltage range where the folded and elongated conformer co-exist simultaneously. Though the unfolding profiles across the samples are similar, the appearance of the elongated ~ 13.6ms species (illustrated as T2‡ in Figure 52) occurs at a higher voltage for both rhGH-Zn bound species (22V) compared to the rhGH control (20 V) as represented in mobilograms Figure 52 d and e. Comparison of the relative proportion of T1‡/ T2‡ states at the transition point of 20 V also suggests a general trend of increased proportion of T2‡/ T1‡ observed across rhGH:Zn 1:10 > rhGH:Zn 1:2 > rhGH. These data indicate that under these conditions, not only does zinc-binding stabilise the globular structure of rhGH, representative of a difference in HOS between rhGH and rhGH-Zn, but that the amount of stabilisation is proportional to the amount of zinc in the sample. These data are in accordance with HDX-MS measurements which indicate higher structural changes at increased zinc concentrations. These data suggest that CIU IMS-MS experiments are also sensitive towards small structural changes, which is particularly attractive due to the high throughput nature of the experiments relative to HDX-MS measurements.

Differential HDX-MS analysis indicated two regions of increased HDX in rhGH:Zn 1:2 samples relative to rhGH. If these increased rates were solely due to the unfolding of helix A due to zinc binding, and hence decreased hydrogen bonding, one might expect a relative decrease in globular stability or CCS, yet, these CIU-IMS-MS data suggest the stabilisation of the structure due to zinc. This would therefore be in accordance not with the unfolding of helix A, but in fact the disruption of hydrogen bonding across either helix A and D or within the central cavity of the protein due to the presence of zinc. It is possible that the zinc ligand stabilises the structure overall through side chain interactions which is not reflected by the amide hydrogens monitored during the HDX-MS experiments described here.

Similarly for rhGH:Zn 1:10 differential measurements a mixture of both destabilising and stabilising changes in secondary structure were observed, though this time at a

much greater magnitude with six regions showing structural changes, illustrating the increased localised structural insights that can be achieved using HDX-MS relative to IMS-MS. The observed rhGH:Zn binding stoichometries of 1:1 and 1:2 revealed by native IMS-MS measurements of rhGH:Zn 1:2 and 1:10 samples respectively, support that ligand binding is indeed not saturated at the lower zinc concentration and the increased structural changes observed at higher zinc concentrations by both HDX-MS and IMS-MS. These data demonstrate the value of using multiple analytical techniques which span multiple HOS levels in order to maximise structural insights.

As aforementioned there is greater potential for use of IMS-MS measurements within a high throughput environment, and though lower resolution than HDX-MS measurements, these type of cross-platform validation suggest that this analytical technique is indeed applicable in a QC type environment such as those required using drug manufacture, in which HDX-MS may not be appropriate.

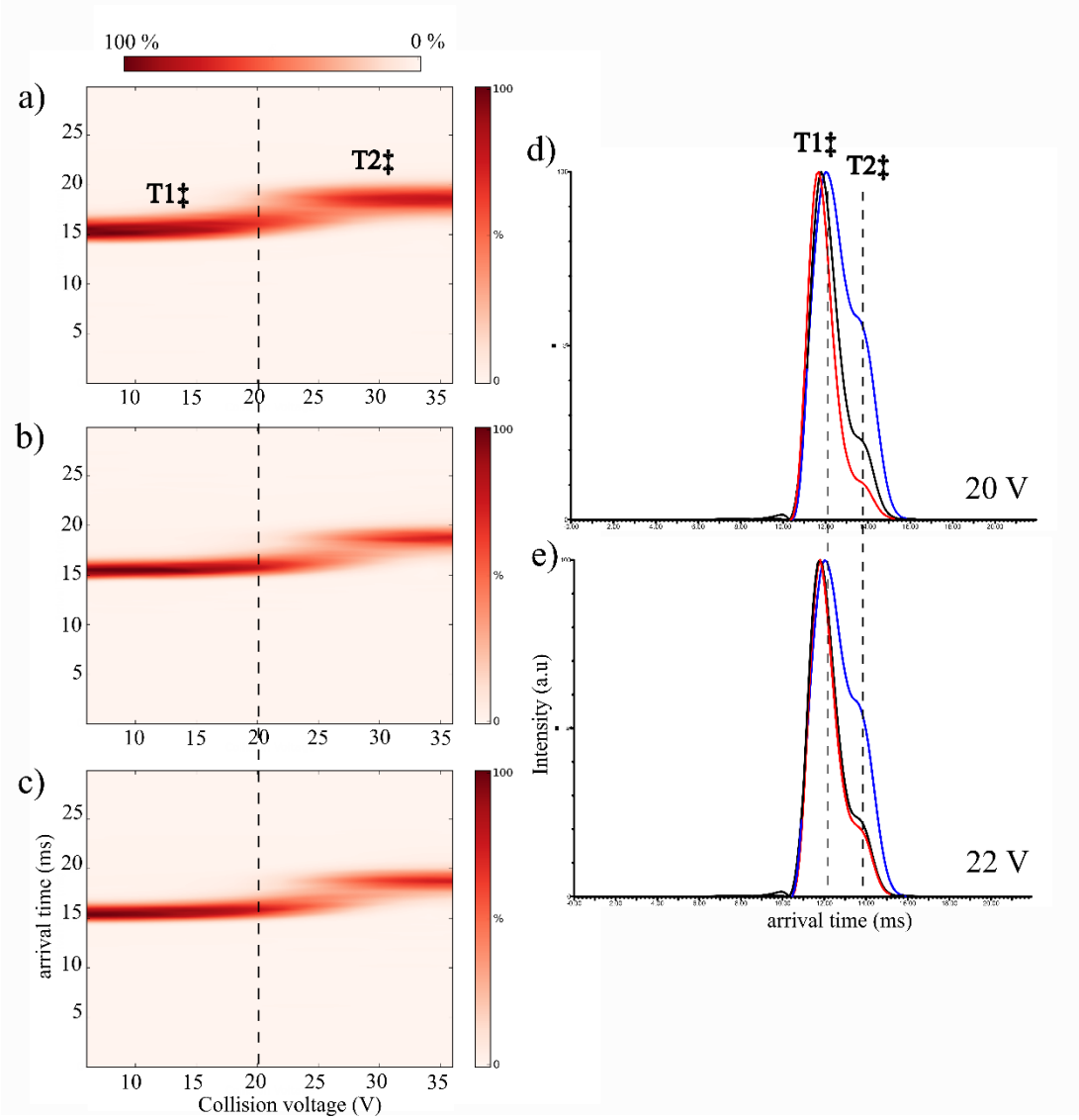


Figure 52. CIU-IMS-MS plots of a) rhGH control b) rhGH:Zn 1:2 and c) rhGH:Zn 1:10. Quadrupole selection of +8 charge state of free rhGH and rhGH-Zn bound states was performed and trap cell activation induced using a voltage range of 6 - 36 V. Data are displayed in three dimensions as arrival time (ms), trap cell collision voltage versus normalised intensity. Arrival time distributions (ATD) are indicated for two folded states compact (T1‡) and unfolded (T2‡) state. A decreased stability towards unfolding can be observed for rhGH (blue) relative to rhGH:Zn 1:2 (black) and rhGH:Zn 1:10 (red) when looking at the overlayed mobilograms for CIU-IMS-MS measurements at a) 20V and b) 22V, indicated by an increased relative amount of T2‡ state with an increased ATD. CIU-IMS-MS plots (a-c) were generated with use of PULSAR processing software [98].

2.4 Conclusions

With use of a commercially available RM (rhGH), a protocol to induce structural perturbations based on zinc ligands has here been developed. This RM has been used to systematically optimise a HDX-MS experimental methodology. A new statistical data analysis approach was developed to determine thresholds that enable discrimination between measurement variability and HOS changes and are specific for each peptide. The statistical approach proposed here is a generic methodology applicable to any protein of interest, that can be applied to improve sensitivity of measurements and increase confidence when defining structural changes detected by HDX-MS measurements. This approach was then applied to the rhGH and zinc model system, to evaluate sensitivity to structural changes of HDX-MS methods developed. The data here generated also provide detailed indications on what can be expected from an optimised HDX-MS system in term of back exchange, precision and accuracy and a comprehensive investigation of the major components affecting measurement uncertainty carried out.

Similarly the RM system was used to assess the applicability, reproducibility, robustness and sensitivity of IMS-MS measurements towards structural changes. Though structural changes were identified by IMS-MS measurements due to the presence of zinc, the sensitivity was lower than that of HDX-MS. HDX-MS measurements also provided much higher localised resolution towards structural changes. This model system is widely applicable to other analytical methodologies, which will allow the assessment of sensitivity and subsequent across platform comparisons. Outside the scope of this thesis, work is currently underway to validate this approach through the conduction of a cross-platform inter laboratory comparison based on defining structural differences of the rhGH model system [233].

Chapter 3

Application of higher order structure characterisation methodologies to a reference protocol for generation of NISTmAb Fc-glycan variants

Chapter 3
Application of higher order structure characterisation methodologies to a reference protocol for generation of NISTmAb Fc-glycan variants

3.1 Introduction

3.1.1 Monoclonal antibodies as biotherapeutics

The development of monoclonal antibodies (mAbs), was first introduced in 1975 by Köhler and Milstein [234], when they described using a clone generated from a single parent cell to produce a single antibody type, a feat that would later result in a share of the 1984 Nobel Prize for medicine and physiology.

Since the introduction of the first FDA approved mAb based drug, OKT3 (Muromonab-CD3), on the market in 1986, more than 60 therapeutic mAbs had been approved by 2017 with many more in the development pathway. Their therapeutic use supports treatment in areas such as oncology, inflammatory and autoimmune diseases, infectious diseases, and cardiovascular indications. The high specificity, low toxicity, long serum half-life and relative ease of manufacture of mAbs have since led them to become the major leading class of the therapeutic market [59] and such their characterisation is of great interest to industry and regulators.

3.1.2 Immunoglobulin (Ig) structure

Due to their prolificacy within the circulatory system, immunoglobulins (Igs) have become the dominant protein platform for the development of mAbs. These ~ 150 kDa glycoproteins facilitate an immune response by the recognition of and binding to a specific foreign antigen, resulting in the ultimate destruction of the antigen. Igs are "Y" shaped hetero-proteins made up of two identical units of a heavy chain (~ 50 kDa each) and two units of light chain (~ 25 kDa each) (Figure 53).

Across the Ig protein-class, slight variances in the heavy chain sequences exists, creating several different subclasses (Ig- A, D, E, G and M), these subclasses are diversified further still with two additional light chain subclasses (κ and λ) [235]. IgG type mAbs are by far the most abundant, accounting for some 75% of all circulatory Igs, and four further subclasses exist for IgG type mAbs, IgG1-4, in which subclass

varies based on the number of disulphide bonds present. The vast array of Ig based structures available are what facilitate the specificity of the antigen immune response, and of all these various isoforms, IgG1 are the most predominant species.

The structure of Igs consist of twelve repeating units, or domains, each made up of 7-9 anti-parallel β-sheet structures which are further stabilised by intra-chain disulphide bonds [236, 237]. Four and two of these domains link up to form both heavy and light chains respectively, and the entire structure is then unified by additional inter-chain disulphide bonds at the centre of the complex termed the hinge region. Figure 53 displays the conventional structure of an IgG based mAb.

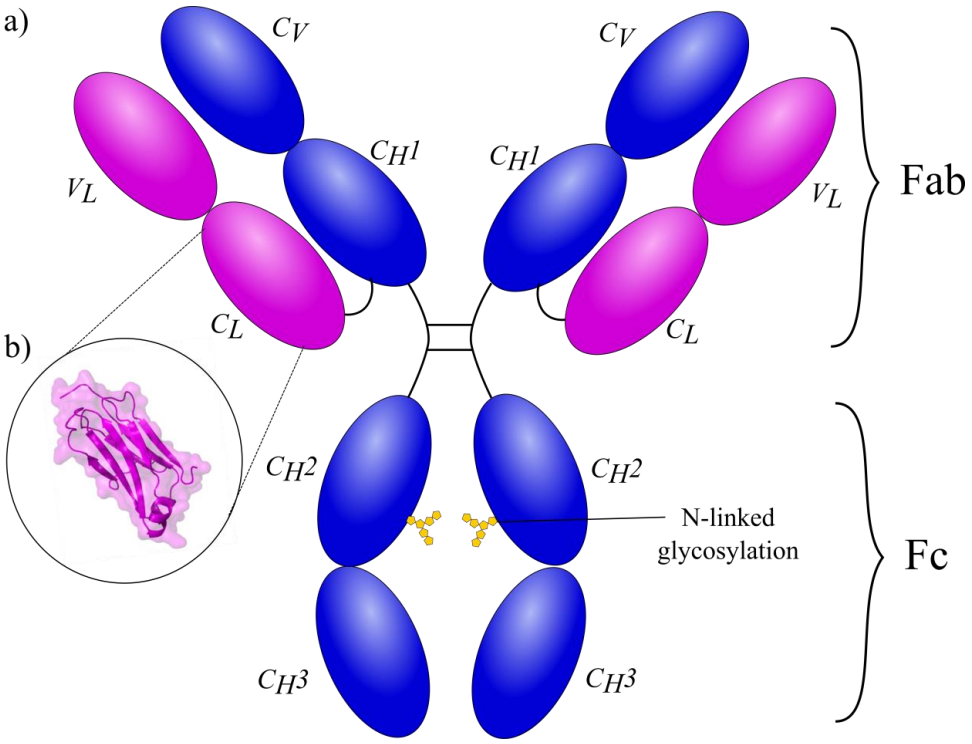


Figure 53 a) Structure of IgG1 mAb. Domains are represented in blue (heavy chain) and pink (light chain), black lines represent inter-chain disulphide bonds, b) crystal structure of constant light chain domain (adapted from PDB: 5k8a [238])

The top half of the molecule consists of two identical antigen binding fragments (Fab) each containing one variable (V_H/V_L) and one constant (C_H1/C_L) domain from both heavy and light chains. The bottom half of the molecule, the crystallisable fragment (Fc), contains two constant domains (C_H2 and C_H3) from each heavy chain present. Recognition of the antigen occurs via interaction with the complementary-determining regions (CDR), which are loop-like structures, found in the variable

domains of the Fab structure. Sequence and inherent structural variation of these CDRs are what yield the high specificity of mAbs.

3.1.3 Issues of heterogeneity

Further to sequence variation and subclasses mAb heterogeneity also arises in the form of PTMs and aggregation, introduced during the complex synthetic pathways used in their production. These PTMs can include glycosylation, oxidation, deamidation, isomerisation, N-terminal pryo-glutamine cyclization and incomplete disulphide bond formation, to mention but a few. Though their presence can be advantageous, by offering the opportunity to tune the properties of a candidate drug. when they arise in an uncontrolled manner during manufacturing processes, they can lead to differences in stability, efficacy and potency of a therapeutic mAb. For these reasons, there must be a tight control of product quality and batch-to-batch consistency, and these heterogeneities deemed CQAs which must be monitored for regulatory requirements [60–62].

Characterisation of this array of heterogeneities and understanding their impact represent a real analytical challenge for the biopharmaceutical industry, both during the high resolution characterisation performed in developmental stages and high throughput monitoring during QC. As aforementioned in the introduction (section 1.2.1.1), the availability of RM, standards and protocols to ensure the effectiveness of the analytical methods developed and help support regulatory documentation for a therapeutic drug are needed for mAbs.

To address this, the first mAb based reference material, RM 8761 (NISTmAb RM), was commercialised in 2016 by the National Institute for Standards and Technology (NIST) with the intended purpose of use for system suitability tests and establishing analytical method performance, variability, comparability and ultimately the qualification of new methodologies for determining physicochemical and biophysical attributes of mAbs. To support the use of this RM, reference analytical datasets have been published through an ACS book series as a collaborative effort across governmental, academic and bio-industrial settings [239–241].

3.1.4 NISTmAb reference material

The NISTmAb RM is a recombinant humanized IgG1κ, with the purpose of simulating a ‘drug-like’ mAb, which, as a part of RM validation, has been

characterised in detail by established analytical platforms [239–241]. These data are inclusive of the determination of NISTmAb primary structure [242], sequence variance analysis [243], structural elucidation of PTMs [244] and N-glycan structure [245], separation science analysis of CQAs [246], biophysical characterisation of HOS [247], a developability assessment [248], particle/aggregate analysis [249–251] plus host cell protein and process-related impurity analysis [252]. In addition, the NISTmAb RM, has been used as a case study to highlight the potential use of state-of-the-art and emerging technologies for mAb characterisation including assessment of multiple bioinformatics workflows [241, 253, 254]. HOS structural information datasets have been produced using technologies including NMR [255, 256], XRC [238], HDX-MS [256], oxidative footprinting [257], chemical XL [257], IMS-MS and high resolution Orbitrap mass spectrometry [258]. Though these provide reference datasets, the published manuscripts do not offer methodologies for assessing the sensitivity of these technologies for use in comparability studies of mAbs, such as detecting HOS changes occurring as result of storage induced degradation and changes in manufacturing processes.

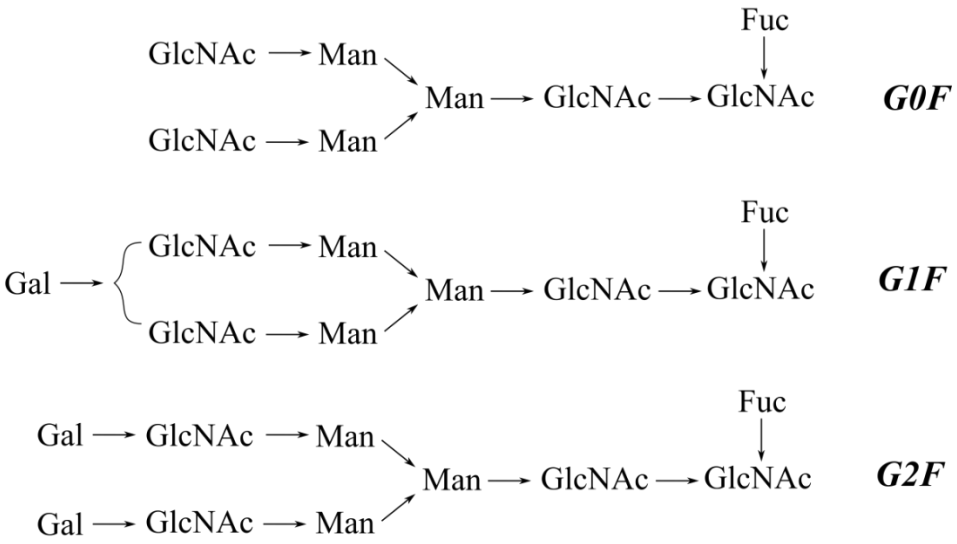


Figure 54. Schematic representation of the most abundant N-glycan structures found in the NISTmAb RM. Nomenclature corresponds to the following oligiosaccharide residues Gal, galactose; GlcNAc, N-acetylglucosamine, Man, mannose; Fuc, fucose.

3.1.5 Glycosylation of IgG1

N-linked glycosylation is present in many native and recombinant mAb biotherapeutics. The complex biosynthetic pathways involved mean producing a single glycoform in vivo is almost impossible, hence any glycosylation present is almost always heterogeneous, further adding to the mAbs structural heterogeneity. Variation in N-glycosylation is significant as it has been demonstrated to influence both antibody-dependent cellular cytotoxicity and complement-dependent cytotoxicity (CDC) [259], meaning this type of heterogeneity can ultimately affect therapeutic dosage of mAbs in the clinic. Variance in mAb glycosylation has also been demonstrated to have significant impact on a wide range of critical attributes including product stability [260], immunogenicity [261, 262], serum clearance [263], pharmacokinetics [264] and potential anti-inflammatory response [265, 266] of biotherapeutics. For these reasons, glycosylation is defined as a CQA of mAb based biotherapeutics by regulators. Furthermore understanding the impacts and mechanisms of glycan variation on the therapeutic use of mAbs is of great interest to the biopharmaceutical industry, for exploiting potency and hence minimising manufacturing costs of a drug.

For IgG type mAbs, glycosylation is often present as N-linked glycosylation at the Asn300 residue of the Fc region. Typically this is in the form of a biantennary complex type oligosaccharide, with a core fucose (Fuc), bisecting N-acetylglucosamine (GlcNAc) structure and terminal β-1,4-galactose (Gal) residues. The batch-to-batch heterogeneity of these terminal β-1,4-galactose linkages is considered to be one of the most common glycan modifications in humanized IgGs, with variation the number of terminal (0, 1 or 2) Gal residues, termed G0F, G1F and G2F glycan isoforms (Figure 54), present on both of the two heavy chains. Slight variations in pathology, culturing and species of the chosen cell expression lines can generate changes in the glycan structure such as the number of terminal Gal residues. These types of variations have been shown to have implications of a biotherapeutics immunogenicity, effector function modulation [267] and CDC [259]. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity [268, 269]. Similarly complete absence of glycan structure alone is known to affect immunogenicity, thought to be as a result of enhancement and stabilisation of the C_H2 domain necessary for FcγR binding [269].

3.1.6 Alternative proteases for on-line digestion of mAbs

The resistance to proteolysis often observed for mAbs [270], particularly by the light chain [271], is known to make obtaining HDX-MS peptide maps of sufficient resolution to give high resolution HOS insights, challenging. For this reason it was anticipated that the use of alternative acidic proteases to pepsin may need to be considered when optimising HDX-MS peptide mapping of NISTmAb RM to circumvent the known specificities of pepsin [142, 145, 272, 273] or to maximise the potential spatial resolution achieved when using a single protease. Examples of such alternative proteases used in HDX-MS workflows include: aspergillopepsin (protease type XIII, EC 3.4.23.18) [141, 273, 274], rhizopuspepsin (protease type XVIII, EC 3.4.23.6) [141, 272, 273], and plasmepsin (EC 3.4.23.39) [275]. Contrasting data from studies that compare the performance of different proteases [141, 142, 272, 276], tend to suggests a lack of a one-shoe-fits-all approach for the selection of an ideal protease to digest a protein of interest. Yet several studies have suggested the attractiveness of the short peptides produced by aspergillopepsin for yielding higher resolution of HDX-MS measurements[141, 276].

Nirudodhi et al. [277] demonstrated the effectiveness of using a commercially available co-immobilized enzymatic column, containing both pepsin and aspergillopepsin, for the digestion of IgG2s, compared to pepsin immobilised alone. It was thought that the close proximity of the proteases and complementary cleavage preferentiality of the enzymes facilitates more superior proteolysis, with high digestion coverage and redundancy (≥95% and ≥ 6.5) also observed for IgG1 based mAb using the dual protease column. Similarly the benefits of multiplexing proteases for increased resolution of HDX-MS measurements have been demonstrated in other studies [272, 273].

3.1.7 Model system: NISTmAb RM and Fc-glycan variants

The N-linked glycan structure of the NIST mAb RM has been characterised extensively [245] and has been indicated as having an ensemble population of G0F, G1F, G2F isoforms, as illustrated in Figure 54. Recently a study by Arbogast et al. [255], demonstrated, using the application of multivariate analysis of 2D NMR data, that enzymatically induced changes in the glycosylation pattern of the NIST mAb RM led to changes in HOS.

The work described the use of simple enzymatic bio-chemistry to generate Fc-glycan variants of NISTmAb RM, in the form of deglycosylation (complete glycan removal) and degalactosylation (removal of terminal Gal residues), simulating such heterogeneities that may exist as a result of changes in manufacturing processes or cell culture lines. These model systems were then used to validate a statistical approach for screening 2D NMR fingerprinting data for significant changes in HOS, a potential methodology for industrial comparability studies.

Although structural changes were identified, the high spectral complexity and absence of a 15N labelled protein resulted in unresolved amino acidic chemical shifts and hence the data gave no resolution as to where the HOS changes occur. Using the enzymatic methodology described, a model system consisting of the 3 Fc-glycan variants of NISTmAb, (control, deglycosylated and degalactosylated) has here been described and used to test, validate, and build on the analytical methodologies (HDX-MS and IMS-MS) developed in Chapter 2, using a more complex and potentially a more biologically relevant molecule to the biopharmaceutical industry as a whole.

Additionally these data are used to provide localised and globular resolution to changes in HOS as a result of Fc-glycan variation previously reported by NMR [255]. These data will ultimately provide a reference data set to support use of NISTmAb RM model system for use in the validation and assessment of analytical methodologies ability to detect changes in HOS of biotherapeutic mAbs [224].

3.2 Methods

3.2.1 Materials

All chemicals and standards were purchased as stated in section 2 unless otherwise stated. The NISTmAb RM 8761, humanized IgG1κ mAb, was purchased from NIST [Gaithersburg, Maryland, USA] and ammonium hexafluorophosphate (AHFP) was purchased from ACROS Organics [Fisher scientific, Loughborough, UK].

3.2.2 Protein sample preparation

3.2.2.1 Generation of Fc-glycan variants

A stock solution of 67 µM NISTmAb RM 8761 (10 mg/mL) IgG1κ mAb in 12.5 mmol/L L-histidine, 12.5 mmol/L L-histidine HCl (pH 6.0), was received and split into 50 µL aliquots upon arrival before storage at -80°C for single use.

3.2.2.2 Generation of Fc-glycan variants

Endo-deglycosylated (DeGly) NISTmAb was prepared by dilution of 50 µl of NISTmAb RM stock (67 µM) in 50 µL 100 mM potassium phosphate buffer saline (KPBS, pH 7.4) and incubated with 4 µL peptide:N-glycosidase (PNGase F) [10 ku/mL, Promega, Southampton, UK] at 37 °C for 24 hours. Exo-degalactosylated (ExoGal) NISTmAb was prepared by diluting 50 µl of NISTmAb stock (67 µM) in 50 µL 1X GlycoBuffer (50 mM sodium acetate, 5 mM $CaCl_2$ pH 5.5) and incubated with 4 µL β 1-4 galactosidase [8 ku/mL, both New England Biolabs, Ipswich, Massachusetts, USA] a ratio of 1mg:0.07 KU protein:enzyme at 37 °C for 72 hours [255].

PNGase F (~36 kDa) and β 1-4 galactosidase (~ 94 kDa) were removed from their respective samples via ultrafiltration using Amicon Ultracel-100 regenerated cellulose membranes with 100 kDa molecular weight cut-off [Sigma-Aldrich, Dorset, UK] and a 5810R centrifuge [Eppendorf, Stevenage, UK]. The membrane filter was preconditioned using 400 µL 10 mM KPBS (pH 6.2) and centrifuged at 4 °C at 13,000 g for 8 min, as were all ultrafiltration steps. Following filter conditioning, 200 µL sample was loaded followed by two sequential washing steps consisting of 200 µL 10 mM KPBS (pH 6.2). Post centrifugation and sample concentration, samples were diluted to 13.2 µM sample concentration using 10 mM KPBS (pH 6.2).

For both enzyme treatment procedures, a control glycosylated (NISTmAb RM) sample was treated in parallel, where in place of the relevant enzyme 4 µL of equivalent incubation buffer was added. All prior sample handling, heated incubation and ultrafiltration steps between the enzyme containing and control samples were otherwise identical. These samples will be referred to as “DeGly NISTmAb control” and “ExoGal NISTmAb control” herein.

3.2.2.3 Sample preparation for IMS-MS measurements

Prior to native IMS-MS analysis, NISTmAb RM, DeGly NISTmAb and ExoGal NISTmAb samples were desalted and buffer exchanged twice into 50 mM ammonium acetate as described in section 2.2.1.2, using Bio-spin micro P6 pre-packed size exclusion columns. Following desalting samples were diluted further to 5 µM in 50 mM ammonium acetate. Samples were left to equilibrate at room temperature for 1h before analysis. For inter-day measurements each sample was freshly buffer exchanged and analysed on the same day.

3.2.3. Intact analysis of Fc-glycan variants

Glycosylation, deglycosylation and degalactosylation was verified at the intact protein level by LC-MS analysis as described below.

3.2.3.1 Chromatography

Chromatography was performed using a Vanquish UHPLC system [Thermo Fischer Scientific, Bremen, Germany]. 1.5 pmol of sample was injected onto an Aeris Widebore XB-C18 column [2.1 x 100 mm, 3.6 µm, Phenomenex, Torrance, California, USA] and chromatographic separation achieved using a 350 µL/min flow rate, 50 °C column temperature and a linear gradient starting from 95% A/ 5% B to 2% A/ 98% B over 6 minutes. Mobile phases consisted of aqueous, 0.5% formic acid (A) and 0.5% formic, CH_3CN (B).

3.2.3.2 MS conditions

Following chromatographic separation intact mass analysis was performed a Q-Exactive Plus Orbitrap instrument, using a HESI-II probe source [both Thermo Fischer Scientific, Bremen, Germany] in positive ESI mode. Source conditions included a 320 °C source temperature, 3 kV capillary voltage, 90 S-lens RF, 25 a.u. sheath gas, 12 a.u. aux gas, 80 V in-source CID. MS experiments were performed using the following parameters; 200 ms maximum IT, $3E^6$ AGC target value, 10 microscans, high mass resolution (HMR) mode. Data were acquired in a resolution mode of 17.5 K over a range of 1700 – 4500 *m/z*. The instrument was calibrated in the high mass range of 1000 – 5000 *m/z* using a 1 mg/mL solution of AHFP.

3.2.3.3 Deconvolution of intact protein measurements

Protein deconvolution was performed using BioPharma finder v 2.0 software [Thermo Fischer Scientific, Massachusetts, USA] using the ReSpect™ algorithm for isotopically unresolved spectra. Deconvolution was performed in the 140-150 kDa range, considering 10-100 charge states, a range of 6-10 minimum charge states and a target mass of 148 kDa.

3.2.4. HDX-MS experiments

3.2.4.1 Instrumentation and experimental settings

Sample handling and mixing steps were performed as described in section 2.2.3 unless otherwise stated. 13.2 µM protein sample solutions were diluted 10-fold in 10 mM

KPBS (pH 6.2). For exchange experiments, deuterium labelling was performed by diluting 15 µL of protein 10-fold in 10 mM KPBS (pH 5.8, uncorrected for isotope effect) prepared in D_2O. Exchange times were 5 and 60 min with the exchange occurring at room temperature (21 ± 2 °C) and each incubation period run in triplicate. Following the appropriate incubation period, exchange was quenched by a 2-fold dilution with 50 µL of 1 M TCEP, 8 M Urea in 100 mM KPBS (pH 2.3) at 4°C. Following a 5 min quench delay, 95 µL of quenched sample was injected onto a refrigerated nanoACQUITY UPLC System with HDX technology for on-line pepsin digestion and chromatographic separation, as described previously.

Method variation in digestion conditions from section 2.2.3.2 include a digestion temperature of 15 °C, a flow rate of 70 µL/min (mobile phase 0.5% v/v formic acid) and a digestion pressure of ~ 7000 psi applied using a PEEK restrictor placed after the trap column. Otherwise chromatographic separation, MS and data analysis workflows were followed as described previously in sections 2.2.3.2 and 2.2.4.

3.2.4.2 Alternative protease evaluation

Using the final optimised peptide mapping conditions described above an additional commercially available on-line proteolysis column, a 2.1 x 30 mm protease column containing a 1:1 mixture of immobilised pepsin and aspergillopepsin (Type XIII protease) [NovaBioAssays, USA] was also evaluated. All sample handling, quenching and repeats were as previously described. Lower pressures of ~ 800 psi were applied for the NovaBioAssay column due to manufacturers recommendations.

3.2.5. Native IMS-MS experiments

3.2.5.1 Instrumentation and experimental settings

nESI-IMS-MS experiments were performed as described in section 2.2.6.1 unless otherwise stated. Experimental settings were optimised as follows: source temperature 30°C, sampling cone 75 V, source offset 50 V, capillary voltage 1.4 kV, trap CE 3 V, trap gas flow 3 mL/min, helium cell has flow 180 mL/min, IMS gas flow 60 mL/min, trap DC entrance 3 V, trap DC bias 45 V, trap DC -2 V, trap DC exit 0 V, IMS DC entrance 10 V, helium Cell DC 25 V, helium exit -5 V, IMS bias 3 V, IMS DC exit 0 V, transfer DC Entrance 4 V, transfer DC exit 5 V, IMS wave velocity 800 m/s, IMS wave height 40 V, transfer wave velocity 47 m/s, transfer wave height 4 V, stepwave1 RF offset 300, stepwave2 RF wave offset 350, 200 µsec mobility trapping release

time; 20 V trap height; 0 V extract height. Data were acquired in resolution mode in the 2000-8000 *m/z* range, using a 2 sec scan time for 2 minutes. No cone or purge gas were used. Parameters were systematically optimised by infusion NISTmAb RM to ensure effective transition of gas ions whilst avoiding unfolding of the gas ions within the instrument.

3.2.5.2 CIU-IMS-MS experiments

CIU-IMS-MS experiments were conducted using IMS-MS experimental conditions stated above but with the quadrupole selection of the + 24 charge state. The ions 6180 *m/z*, 6050 *m/z* and 6180 *m/z* were quadrupole selected during CIU-IMS-MS experiments for NISTmAb RM, DeGly NISTmAb and ExoGal NISTmAb samples respectively. The trap collision voltage was increased in 1 V increments between the range of 30 – 55 V rate and 5 V increments between 55-100 V and the data acquired for each voltage. Data processing was carried out as described previously in section 2.2.6.2.

3.2.6. NISTmAb structural models

Due to the current unsolved XRC status of the intact NISTmAb RM, two solved XRC IgG1 fragment structures of both the Fab (PDB:5K8A [238]) fragment and a fucosylated-Fc (PDB:3AVE, [278]) IgG fragment are herein used to represent the structure of NISTmAb RM and Fc-glycan variants. These structures were solved to 1.99 Å and 2 Å for Fab and Fc respectively and the PyMOL fragments combined to illustrate a complete structure using Inkscape 0.92.3 (Copyright (C) 1989, 1991 Free Software Foundation, Inc.)

3.3 Results & Discussion

3.3.1. Intact protein characterisation of Fc-glycan variants

Intact LC-MS of the NISTmAb RM and Fc-glycan variants was performed to confirm the presence of the expected glycoforms for NISTmAb RM and the successful enzymatic generation of deglycosylated and degalactosylated forms. Subsequent spectra were then deconvoluted to experimental intact masses (Figure 55), which were compared with theoretical and previously reported masses [242] as stated in Table 12 and Table 13.

The presence of all five major expected glycoforms (Figure 54) was observed upon deconvolution of the NISTmAb RM sample, with the most abundant being G0F/G1F type. Complete deglycosylation due to PNGase F treatment was confirmed for the DeGly NISTmAb sample.

For the ExoGal NISTmAb sample, treatment with β 1-4, galactosidase saw a shift from the predominant species of G0F/G1F type towards predominantly G0F/G0F type, indicating removal of both terminal galactose residues at both C_H2 glycoyslation sites present in NISTmAb RM. Complete absence of all glycosylated species containing both terminal galactose residues (G2F type, Figure 54) was observed, and though there was evidence of the presence of single terminal glycan species (G1F type), this G1F type was evident with reduced abundancy to that of the NISTmAb RM control, this is similar as previously reported [255]. The degalactosylation demonstrated here simulates the type of micro heterogeneity that can arise as a result of differing manufacturing practices (such as changed cell lines) which could result in subtle but real changes in HOS of mAbs [268, 279].

Comparison of the deconvoluted masses with theoretical masses[242], average Δ mass tolerances were found to be within $\Delta \sim 5.8\text{-}22.8$ ppm across all the glycoforms and $\Delta \sim 9.1\text{-}22.3$ ppm for the deglycosylated sample.

Table 12. Average deconvoluted (n=3 measurements) intact masses of glycoforms observed in LC-MS analysis of the a) NISTmAb RM control and ExoGal samples. Measurement error (standard deviation) from theoretical values are stated in ppm. * Theoretical masses are taken from Formolo et al. [242].

		Deconvoluted Mass (Da)		Error (ppm)	
Isoforms	**Theoretical ***	**NISTmAb RM control**	**ExoGal**	**control**	**ExoGal**
G0F/G0F	**148036.4**	148038.4	148038.7	13.5	15.3
G0F/G1F	**148198.6**	148200.7	148200.9	13.9	15.2
G1F/G1F	**148360.7**	148361.8	148361.4	7.3	4.5
G1F/G2F	**148522.8**	148521.9		-5.8	
G2F/G2F	**148685.0**	148681.8		-21.8	

Table 13. Average deconvoluted (n=3 measurements) intact masses of main isoforms observed in LC-MS analysis of the DeGly NISTmAb samples. Measurement error (standard deviation) from theoretical values are stated in ppm. * Theoretical masses are taken from Formolo et al [242].

Isoforms	**Theoretical***	**Deconvoluted Mass (Da)**	**Error (ppm)**
DeGly	**145148.5**	145150.1	11.1
+ Hex	**145310.6**	145311.9	9.1
+ 2Lys	**145404.0**	145407.2	22.3

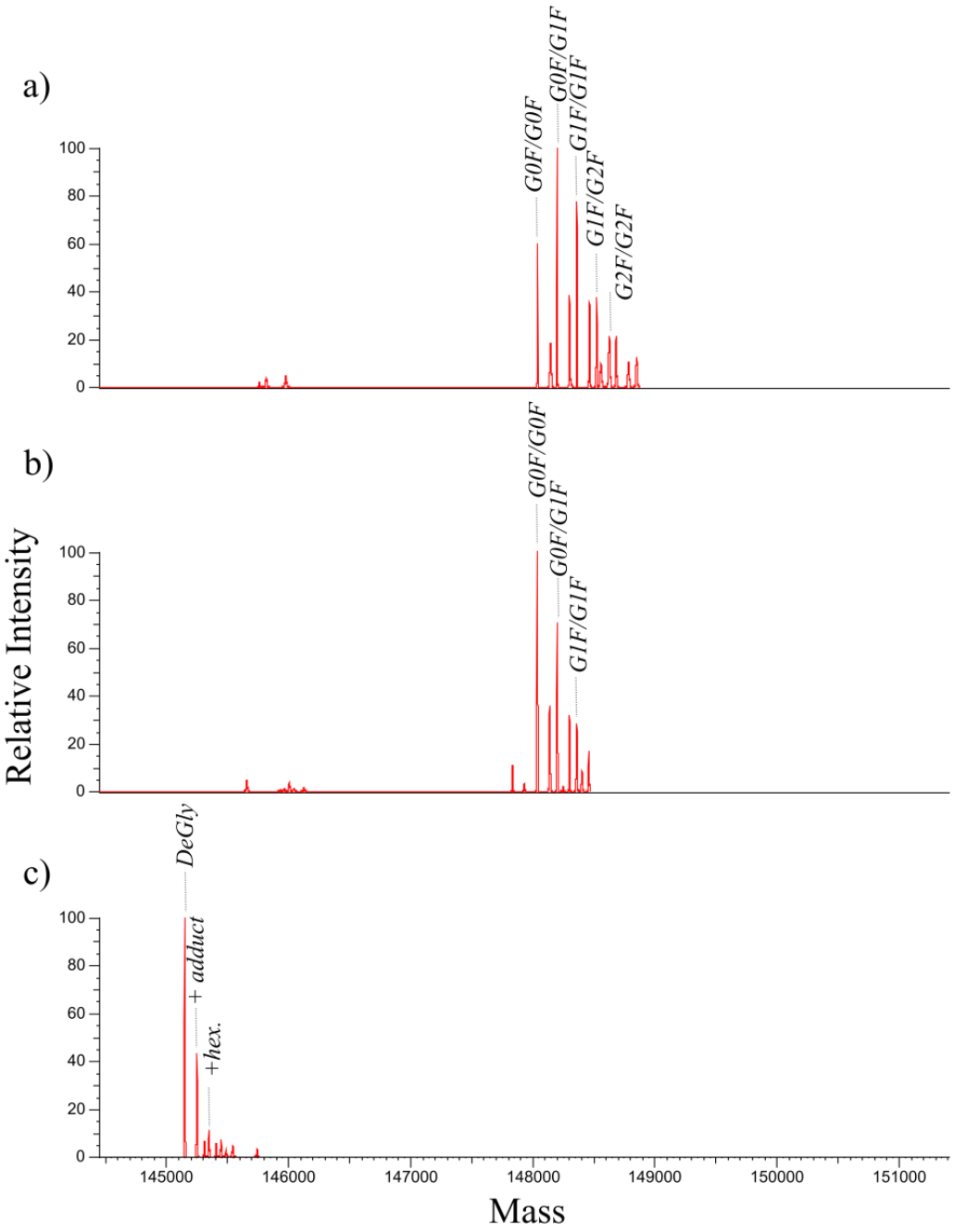


Figure 55. Deconvoluted intact mass spectrum of a) native glycosylated NISTmAb RM, b) β-1-4,-galactosidase treated (ExoGal) NISTmAb and c) PNGase F treated (DeGly) NISTmAb. The abundant glycoforms are indicated which are of biantennary complex type oligosaccharides (GXF), that varying the number of terminal β-1,4-galactose residues as indicated by numbers (X = 0,1 and 2), a hexose (hex.) containing impurity is also indicated.

3.3.2. HDX-MS method optimisation

As described for the rhGH protein in Chapter 2 the peptide mapping stage was optimised for NISTmAb. Systematic initial screen of a range of digestion conditions was made for the on-line pepsin digestion of NISTmAb RM, including variation of chemical denaturing agent concentration, digestion pressure, temperature and flow rate, assessing the resulting peptide maps using the digestion factors previously described section 2.2.4.3.

3.3.2.1 Optimisation of peptide mapping of NIST mAb RM

The peptide maps produced for both heavy chain (HC) and light chains (LC) under each of the various digestion conditions screened are summarised in Table 14 .

Increased molarity of chemical denaturants, Gnd.HCl (4M and 6M) and TCEP (200mM, 400mM and 600 mM), Table 14, conditions i-iv, facilitated higher sequence coverage and redundancy for both HC and LC of the protein. However, even under these relatively strong quenching conditions there was still incomplete digestion of both HC and LC (~ 85% and ~ 93% sequence coverage, Table 14 , condition iii).

Additionally, there was evidence of protein carry-over (data not shown) during the organic washing stage of the chromatographic separation, indicating not only incomplete digestion but the presence of carry-over can have significant impact on the magnitude and reproducibility of measured deuteration rates due to contributions towards back exchange and ultimately isotopic distributions [280].

An increase in HC redundancy was observed when increasing the digestion temperature from 15 to 20 °C, (Table 14, conditions iv and v), though little gain in sequence coverage was observed. Yet, at 20 °C, there was an apparent slight decrease in number of LC peptides produced, though poor reproducibility and low digestion efficiency observed under both these conditions at both temperatures. In light of this 15 °C was deemed a more appropriate digestion temperature to use.

Increasing pepsin flow rate across 70, 90 and 115 µL/min (Table 14, conditions vi-viii) led to a decrease in percentage sequence coverage of the HC, due to reduced effective digestion caused by the increased flow rate, though LC sequence coverage was less significantly affected by the change in flow rate. Digestion indicators for the LC peptide map were still poor and combined with the presence of intact LC in the

chromatograms produced as described previously, suggested incomplete LC digestion.

Increased sequence coverage and redundancy were observed when the digestion pressure was increased from 4000 to 7000 psi (Table 14, conditions v and ix) with use of a PEEK restrictor after the trap column.

Comparison of initial condition i with resultant condition ix shows a significant improvement in peptide map coverage through simple screening of conditions. However peptide map coverage was still ~ 90% for LC and HC and intact undigested protein was evident upon appraisal of the resulting chromatograms. This was considered to be most likely due to incomplete reduction and denaturation of the protein, therefore optimisation of the quench hold time [281], that is the time protein is held in the quench solution, was explored as a potential way to facilitate better digestion.

Table 14. Summary of the resultant peptide maps produced from the on-line digestion of NISTmAb using a range of initial screening conditions. Only those peptides observed in n ≥ 4/5 injections were retained for peptide map.

Condition	Quenching composition		On-line digestion conditions			Heavy chain			Light chain		
	Gnd.HCl (M)	TCEP (mM)	Temp. (°C)	Flow rate (µL/min)	Pressure (psi)	# Peptides	Sequence Coverage (%)	Redundancy	# Peptides	Sequence Coverage (%)	Redundancy
i	4	200	15	70	4000	106	74.0	4.15	11	40.4	1.65
ii	4	400	15	70	4000	97	73.1	4.05	19	57.7	1.74
iii	4	600	15	70	4000	129	85.1	4.82	29	93.4	2.09
iv	6	400	15	70	4000	108	82.7	4.04	33	85.4	2.59
v	6	400	20	70	4000	116	83.8	4.44	31	77.5	2.42
vi	4	600	20	70	4000	117	92.7	3.85	30	73.2	2.41
vii	4	600	20	95	4000	123	82.2	4.7	31	73.7	2.47
viii	4	600	20	115	4000	85	73.1	3.62	24	62.0	2.11
ix	4	600	20	70	7000	137	86.7	4.81	43	90.1	3.03

3.3.2.2 Effect of quench holding times

Peptide maps were produced using 0.5 and 5 min quench hold times and further variation of reducing and denaturing reagents and these are summarised in Figure 56.

Much improved digestion was observed for both HC and LC across all digestion factors, when quench hold time was extended from a default 0.5 min to 5 min (Figure 56 conditions i and ii), suggesting incomplete denaturation and reduction of the 16 disulphide bonds present when using the shorter quenching hold times.

Increasing the molarity of the TCEP reducing agent from 600 mM to 1 M (Figure 56, conditions iii and iv) gave further gains in all digestion indicators, corroborating that at the lower molarity, there was indeed still incomplete reduction of disulphide bonds present within the NISTmAb structure. Almost complete sequence coverage, 99% and 98% HC and LC respectively, was maintained when the denaturing agent was switched from 4 M Gnd.HCl to 8 M urea, yet a 1.4 fold increase in redundancy and 1.2 fold increase in the number of peptides was observed with use of urea.

In hope of reducing any further unreduced disulphides and/or preventing any reformation of reduced disulphides during the pepsin digestion stage, a low molarity of TCEP (8 mM, Figure 56 condition v) was added to the on-line pepsin mobile phase. However, this resulted in a detrimental effect on all digestion indicators, likely due to increased ion suppression of the MS signal and hence a lower number of peptide identifications made from the resultant spectrum.

Overall increasing quench hold time, reducing agent molarity and the use of an alternate chaotropic agent resulted in an increase of sequence coverage, redundancy and number of peptides for the NISTmAb HC by 20%, 50% and 62% respectively. More significant gains were observed for the LC with values of 40%, 165% and 245% respectively, demonstrating a far superior pepsin digestion, peptide map coverage and ultimately significantly increased resolution of HDX-MS measurements using these optimised conditions.

3.3.2.3 Effect of back exchange rates

To check for negative effects on back exchange due to increased quench hold times, a sample of NISTmAb was incubated in deuterium for 5 minutes before being subsequently quenched using both 0.5 and 5 min quench hold times, and analysed using bottom-up HDX-MS workflow.

Considering all peptides observed in both data sets produced across both quench hold times, increased relative uptake measurements were in the range of 0.04 – 5.33% for NISTmAb analysed using the 0.5 minute quench hold time, with a mean increased value of 0.7% across all peptides.

The large gains observed in peptide map quality due to increased quench holding were therefore deemed to outweigh a small increase in back exchange. Unless otherwise stated, for all proceeding HDX-MS on-line digestion was performed using a quench containing 8 M Urea, 1M TCEP, a quench hold time of 5 min, 15°C and a flow rate of 70 µL.

Figure 57 and Figure 48 illustrate the resultant peptides maps produced for NISTmAb using these optimised conditions.

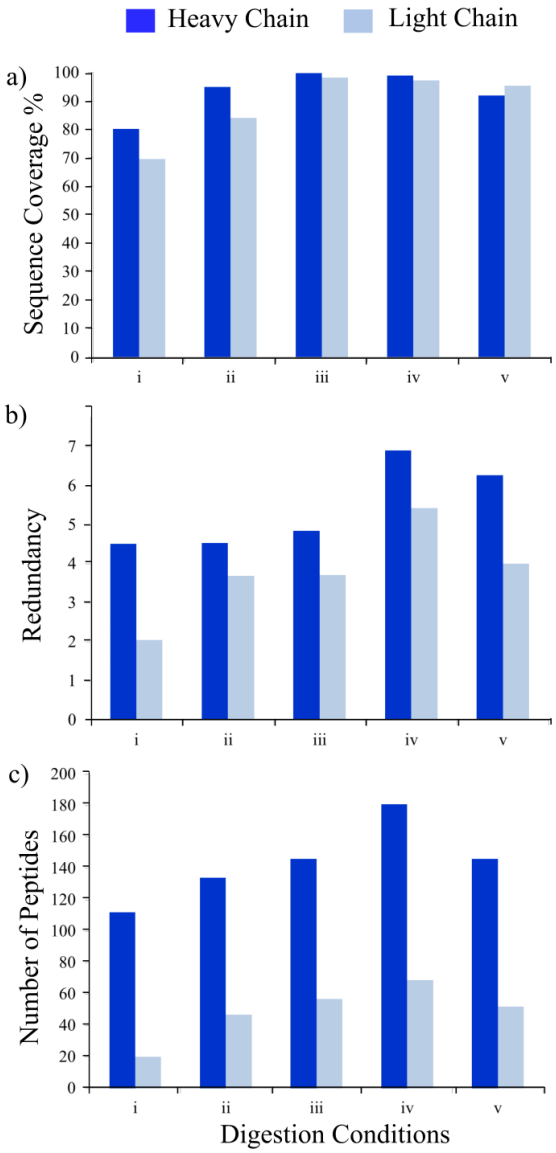


Figure 56. Representation of NISTmAb RM peptide maps from on-line pepsin digestion under a range of digestion conditions (i-v). Peptide maps are evaluated in terms of a) sequence coverage percentage b) redundancy c) number of peptides identified. Quenching composition, quench hold time and pepsin mobile phase were varied to optimise pepsin digestion. Digestion conditions were i) quench - 4 M Gnd.HCl, 600 mM TCEP, 0.5 min quench hold time ii) quench - 4 M Gnd.HCl, 600 mM TCEP, 5min quench hold time iii) quench - 4 M Gnd.HCl, 1M TCEP, 5min quench hold time iv) quench - 8 M Urea, 1M TCEP, 5min quench hold time v) quench - 8 M Urea, 1M TCEP, 5min quench hold time, 8 mM TCEP pepsin mobile phase.

Figure 57. Peptide map of NISTmAb RM HC, using optimised conditions. These conditions include; quenching solution containing 8 M Urea, 1M TCEP, a 5min quench holding time, 15 °C, 70 µL/min. Only peptide observed in n=4/5 repeat injections are displayed. In total some 182 peptides are displayed, giving 98% sequence coverage and redundancy score of 7.36.

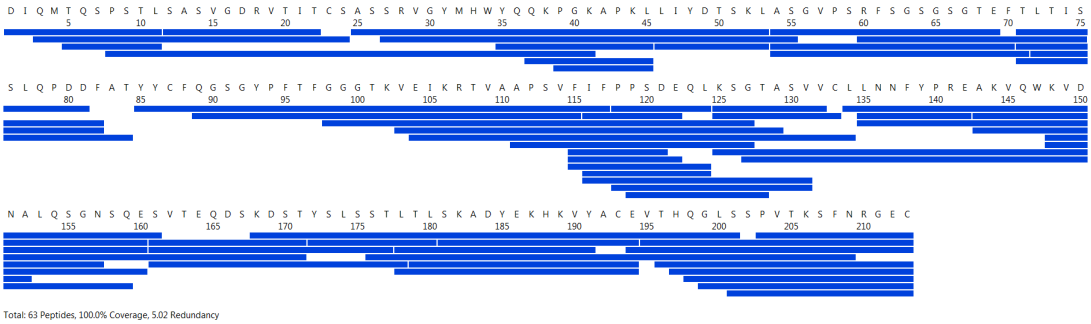


Figure 58. Peptide map of NISTmAb RM LC, using optimised conditions These conditions include; quenching solution containing 8 M Urea, 1M TCEP, a 5min quench holding time, 15 °C, 70 µL/min, Only peptide observed in n=4/5 repeat injections are displayed. In total some 63 peptides are displayed, giving 100% sequence coverage and redundancy score of 5.02.

3.3.2.4 Alternative protease evaluation

Two commercially available on-line proteolysis columns, one containing only immobilised pepsin and the other containing a 1:1 mixture of immobilised pepsin and Type XIII protease (from *Aspergillus saitoi*), referred to as Enzymate™ and NovaBioassay respectively herein, were also evaluated for the digestion of NISTmAb RM.

Table 15 summarises the resultant peptide maps produced from repeat injections (n=5) of 10 µM NISTmAb RM solution using both columns. Both columns demonstrated high sequence coverage with HC and LC coverage in excess of 90 %, consistent with application of the proteases to on-line mAb digestion mAbs in previous studies [277]. However, the NovaBioassay dual protease column yielded a higher number of peptides and higher redundancy rates, 6.37 and 5.56 for HC and LC, respectively, compared to 5.57 and 3.93 redundancy scored for the pepsin column alone. These increased proteolysis and redundancy rates reported are consistent with previous studies [277].

In light of the similar sequence coverage and high redundancy achieved using both columns and ultimately due to lab practicalities and costs, the Enzymate™ column was selected for all further experiments described in this Chapter.

Table 15. Summary of the resultant peptide maps produced of NISTmAb RM using two different commercially available on-line protease columns. Peptide maps are summarised by the number of peptides identified, the percentage sequence coverage and redundancy score.

	Heavy Chain			**Light Chain**		
Pepsin column	**# Peptides**	**Sequence coverage (%)**	**Redundancy**	**# Peptides**	**Sequence coverage (%)**	**Redundancy**
NovaBioassay*	176	93.1	6.37	82	94.4	5.56
Enzymate™	136	91.1	5.57	48	96.2	3.93

3.3.3 HDX-MS measurements and NISTmAb RM structure prediction

Using the optimised peptide mapping method described, HDX-MS measurements were made for the NISTmAb RM control and the uptake measurements were interpreted and compared against known structures for the protein.

Following manual curation of the exchange data, sequence coverage was 89.6% and 91.1% for HC and LC, respectively. Relative HDX-MS peptide rates were mapped onto the solved XRC structures of the NISTmAb Fab (PDB:5K8A [238]) and an Fc-fragment (PDB:3AVE, [278]] solved for an alternative IgG1, as illustrated in Figure 59. Overall regions of predominantly lower relative exchange (cooler colours) was observed, suggesting a relative high amount of structure, consistent with the high levels of secondary structure in the form of the β-sheets known to exist in each domain. However there are some regions of higher relative exchange (warmer colours) tending to populate the bottom of the C_H3 domain and the lower region of the C_H1 domains.

Very low exchange rates (dark blue) were measured for peptides spanning HC 95-99 and HC 221-230 suggesting high structure and/or low solvent accessibility (indicated in Figure 59). Considering this in the context of the XRC structure confirms indeed these peptides are buried within the Fab domain of the crystal structure (Figure 59). Other regions of low exchange kinetics include HC 84-94, HC 111-117, HC 150-159, HC 315-322, HC 394-408, LC 72-74 and LC 173-201; this is in accordance with low kinetics observed during a previous HDX-MS study of NISTmAb RM [256].

Higher kinetics were measured for the following regions: HC 190-220, HC 280-303, HC 352-359, HC 385-394, LC 5-11, LC 142-160, LC 201-214, with these regions consisting of a mixture of loop regions plus β-secondary structures within the constant region of the Fab. Albeit, differences observed in exchange conditions, back exchange rates and peptides make direct and absolute comparisons of relative HDX-MS measurements between different laboratories difficult, though general trends of kinetics can be observed across laboratories.

Considering the 3D heat map, overall higher structural rigidity was observed for the Fab region compared to the Fc region of the NISTmAb RM, in particular the variable domains appeared to have more structural rigidity compared to the constant domains in the Fab region. This rigidity could reflect the need of residual structure for specificity of binding at these variable regions which house the CDRs, the associated

antigen-binding regions. Additionally lower relative exchange rates were measured for those residues within the central cavity of the Fc-region compared to those residues located on the outside and C-terminal end of the domain, reflecting the lower solvent accessibility experienced by residues within the cavity. These data are in good accordance with predicted and published data for the structure of NISTmAb RM validating the effectiveness of the HDX-MS method optimised [256].

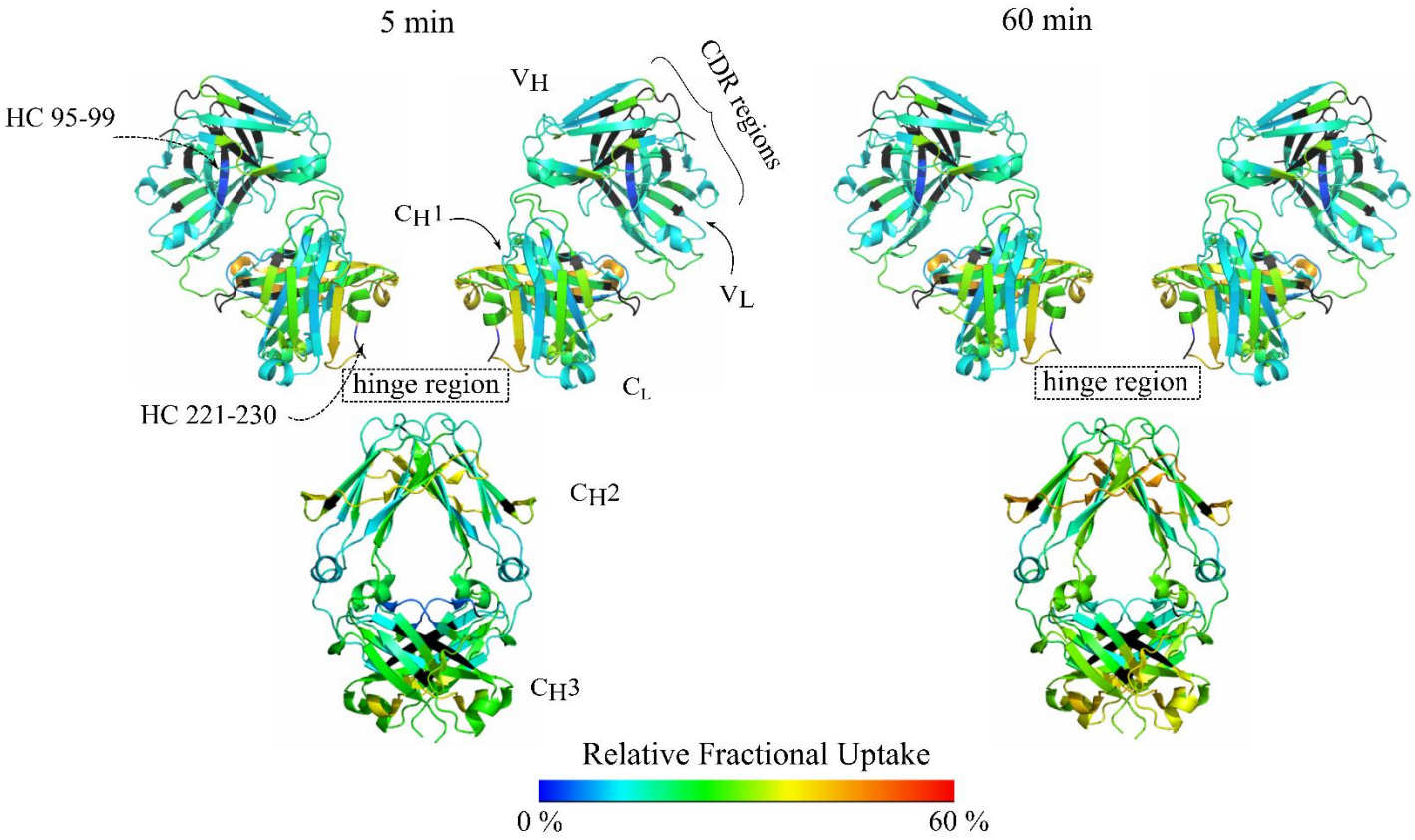


Figure 59. Relative deuterium uptake mapped on to crystal structure of combined Fab (PDB:5K8A[238]) and Fc (PDB:3AVE [278]) regions of IgG1 structure and for two deuterium incubation periods of 5 and 60 minutes. Relative uptake is represent by the colour scale ranging from low uptake (blue) to high uptake (red). Black regions indicate regions of sequence not identified in HDX-MS experiments.

3.3.4 HDX-MS measurements to identify changes in HOS of Fc-glycan variants

3.3.4.1 Definition of significance thresholds for NISTmAb RM

Changes in HOS structure as a result of deglycosylation and degalactosylation of NISTmAb have been reported previously as measured by multivariate analysis of 2D NMR [255]. In this study, however, the specific resonance shifts were not resolved at an amino acid level and localised characterisation of these changes is therefore incomplete. It was hoped that relative HDX-MS measurements of this model system could be able to provide localised insights into where these structural changes occur, validating the applicability of the model system.

Prior to this comparison, as demonstrated by the rhGH model system described in section 2.3.1.5, confidence and sensitivity towards small structural changes in HOS are increased by definition of measurement thresholds using vial-to-vial comparability studies. Prior to the measurement of Fc-glycan variants, first a vial-to-vial HDX-MS comparability was performed using a series of NISTmAb RM control samples.

To account for any structural changes that could occur during the digestion and centrifugal clean-up of the NISTmAb Fc-glycan variants, a sample of the NISTmAb RM control sample was first subjected to heated incubation (minus enzymatic treatment), centrifugal clean-up and dilution, before splitting into two equal vials and HDX-MS measurements from both vials performed concurrently. As for rhGH vial-to-vial, each measurement was performed in triplicate on the same day and a relative vial-to-vial comparison was repeated three times on three different days (n=9 in total measurements).

So as to put the data into context with the final Fc-glycan variant HDX-MS comparison, only those peptides, i, identified in all datasets including a final NISTmAb RM vs DeGly NISTmAb differential study (n = 4 datasets, 10 samples) were considered. Measurement thresholds were then generated from the vial-to-vial comparability measurements using the statistical approach previously described in section 2.3.1.5. Following manual curation of HDX-MS measurements across all datasets, measurement thresholds, represented by the dotted lines in Figure 60 and Figure 61, were established for 87 (76.2% sequence coverage, 200.9% redundancy) and 34 (84.5% sequence coverage, 116.4% redundancy) peptides observed for NIST

mAb samples for HC and LC respectively. All raw differential measurements are detailed in Table A-26 located in the appendices.

3.3.4.2 Comparability of NISTmAb RM and DeGly NISTmAb HDX-MS measurements

Comparison of differential HDX-MS measurements between the NISTmAb RM control and DeGly NISTmAb with thresholds established from the vial-to-vial comparability study (Figure 60 a, Figure 61 a and Table 16), revealed seventeen HC peptides and a single LC peptide exhibiting statistically significant structural changes following deglycosylation. All other regions appeared unaffected and the large majority of the HOS structure was unchanged.

Figure 61 a highlights these regions identified as exhibiting significant structural changes on the XRC structures of NISTmAb RM Fab and Fc regions. Two distinct regions within the C_H2 domain of the Fc region (Figure 60a regions III and IV) revealed increased structural flexibility as a result of deglycosylation, corresponding to 11 peptides spanning residues HC 238-254 and HC 265-280. Relative changes in HDX-MS rates for these specific peptidic sequences, as a result of deglycosylation, are illustrated in Figure 62 b. Located within the hinge and C_H2 domains, all six peptides monitored within the HC 238-254 region exhibit increased exchange kinetics (S ~ between 0.79 – 1.011 Da) following deglycosylation. Resolution afforded by the overlapping peptides, lack of exchangeable amides in proline residues and loss of deuteration of the two N-terminal residues during analysis [125], suggest these changes in kinetics are most significant for residues Phe-246, Lys-249, and residues 251-255 (Figure 62 b).

Increased HDX-MS rates upon deglycosylation have been attributed previously to the loss of an interaction between the phenyl ring present in Phe-246 and the GlcNAc moiety present within the glycan structure [282]. Removal of this interaction indeed seems to lead to reduced stabilisation of these residues which in turn leads to greater flexibility across the HC 238-254 region. Increased flexibility of the specific sequence FLFPPKPKDTLM (residues 244-255, Fc-region) is widely reported in the literature, in the HDX-MS analysis of other IgGs, as a result of undergoing a variety of other perturbations including distal Fc Met-oxidation [129, 283] and chemical destabilisations [284]. These data are concurrent with the identification of this Fc

region as a potential aggregation spot [285] and reported increased aggregation rates as a result of mAb deglycosylation of mAbs [286].

Further still, increased structural flexibility was measured for residues HC 265-280, situated between the C_H2 and C_H3 domains of the Fc region, revealing that the loss of this stabilising carbohydrate interaction appears to also have long-range effects (Figure 62 b and c). Structural changes, in the form of increased flexibility of the C_H2 domain of IgG1-based proteins upon deglycosylation appears to be a repeated observation [130, 279, 287, 288] demonstrating not only the homology of structure across the IgG1 class but also the reproducibility and robustness of HDX-MS measurements for protein conformation dynamics.

Three regions of the Fab portion of the NISTmAb RM, HC 5-20, HC 159-177 and LC 122-131 also show increased structural flexibility as a result of deglycosylation (Figure 60 a, regions I and II and Figure 61 a, region I). Of particular interest is the increase in structural flexibility in the region of HC 159-177, situated in the C_H1 domain (Figure 62 a). This region has previously been indicated as being involved in the interaction between an IgG1 in its native glycosylated form and an FCγRIIIa type receptor by both HDX-MS [258,] and FPOP-MS studies[279]. If this region is indeed part of the epitope of this interaction, then loss of structure in this area is consistent with the well documented loss of activity observed between an IgGs and FCγRIIIa type receptors upon deglycosylation [282].

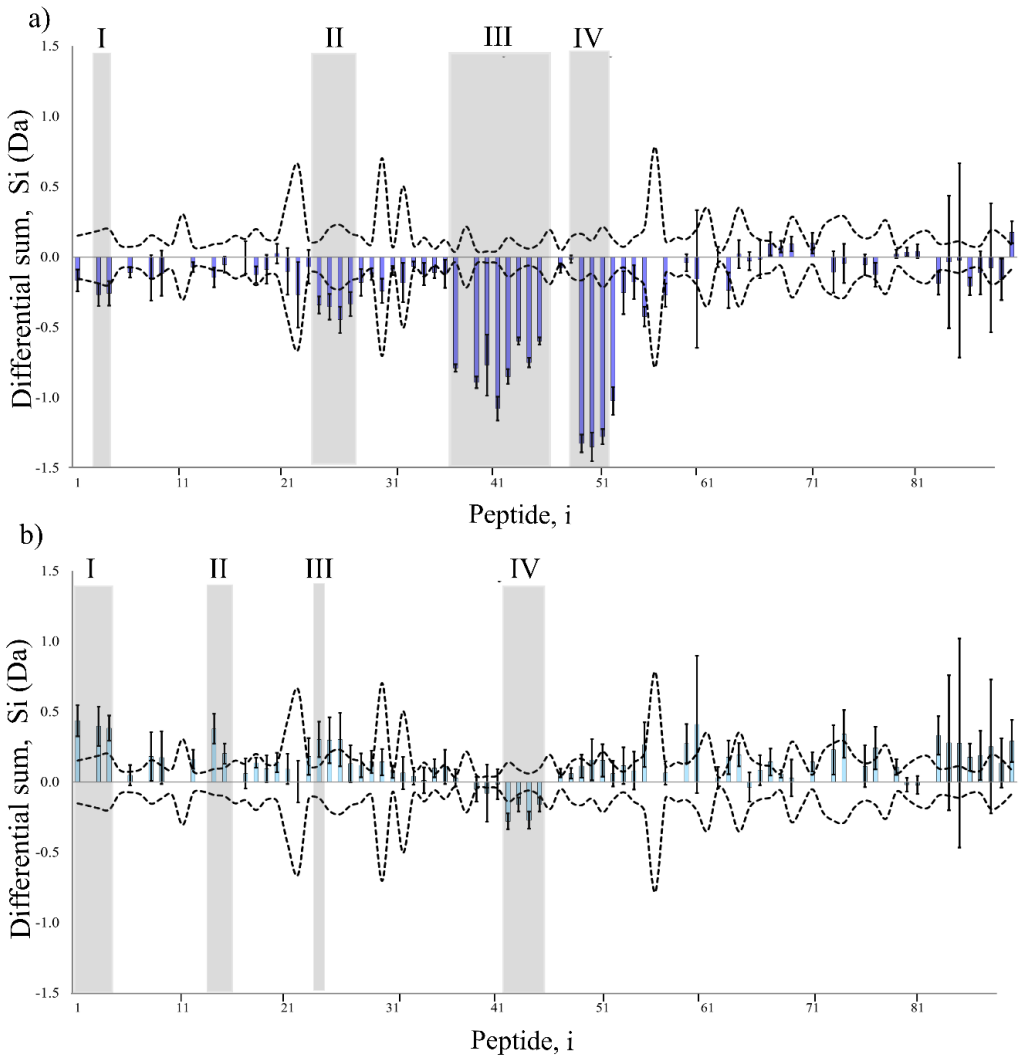


Figure 60. Differential HDX-MS measurements across the heavy chain sequence of NISTmAb RM between the native NISTmAb RM and a) Degly NISTmAb and b) ExoGal NISTmAb. Dotted lines illustrate the measurement thresholds for each peptide, i, calculated from a vial-to-vial comparability study using the NISTmAb RM. Those differential measurements which are in excess of these thresholds demonstrate robust structural changes and are highlighted in grey and correspond to the following residues a) I) HC 5-20, II) HC 159-177, III) HC 238-254 and V) HC 265-280 and b) I) HC 5-20, II) HC 66-82, III) HC 159-177 and V) HC 248-255. Error bars represent standard deviations of triplicate measurements.

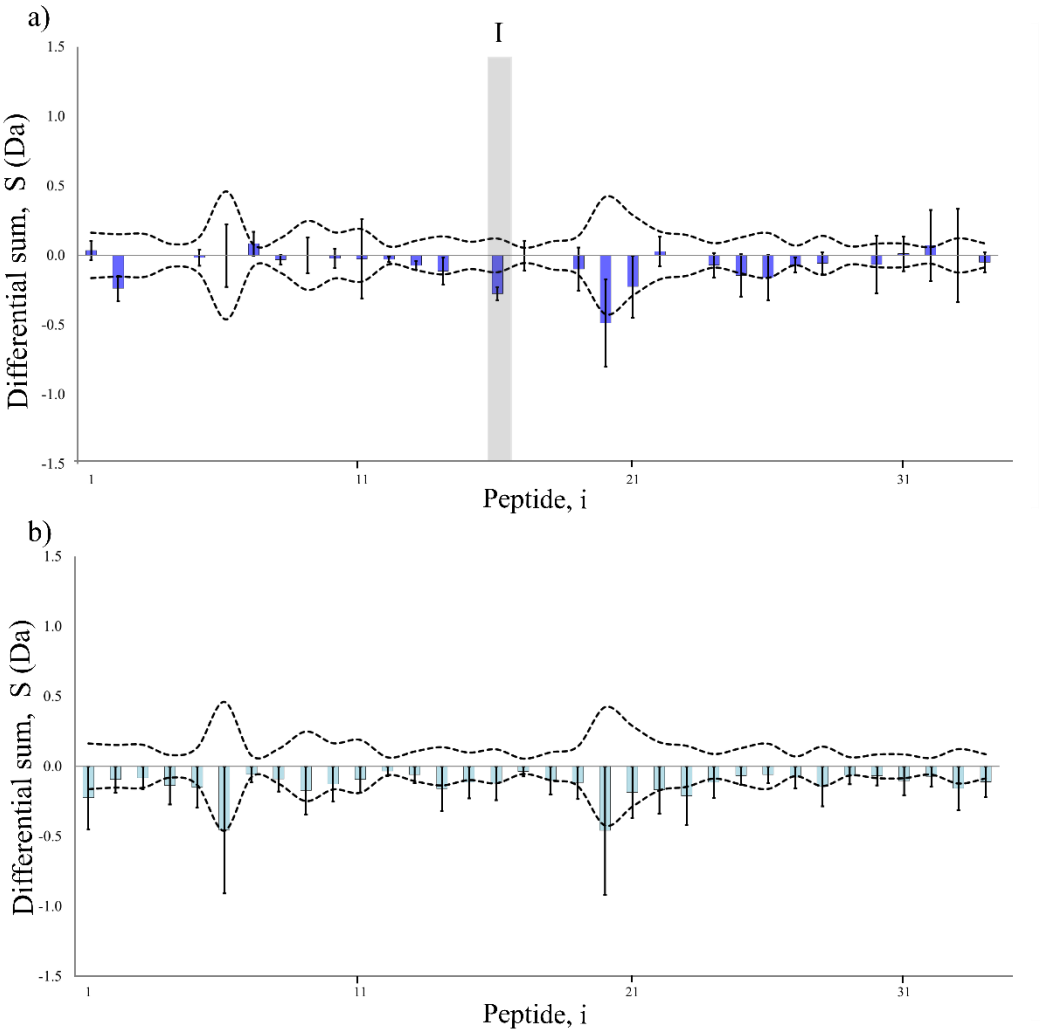


Figure 61. Differential HDX-MS measurements across the light chain sequence of NISTmAb RM between the native NISTmAb glycosylated control and a) Degly NISTmAb and b) ExoGal NISTmAb. Dotted lines illustrate the measurement thresholds for each peptide, i, calculated from a vial-to-vial comparability study using NISTmAb glycosylated control. Those differential measurements which are in excess of these thresholds demonstrate robust structural changes a) region I corresponding to residues LC 122-131. Error bars represent standard deviations of triplicate measurements.

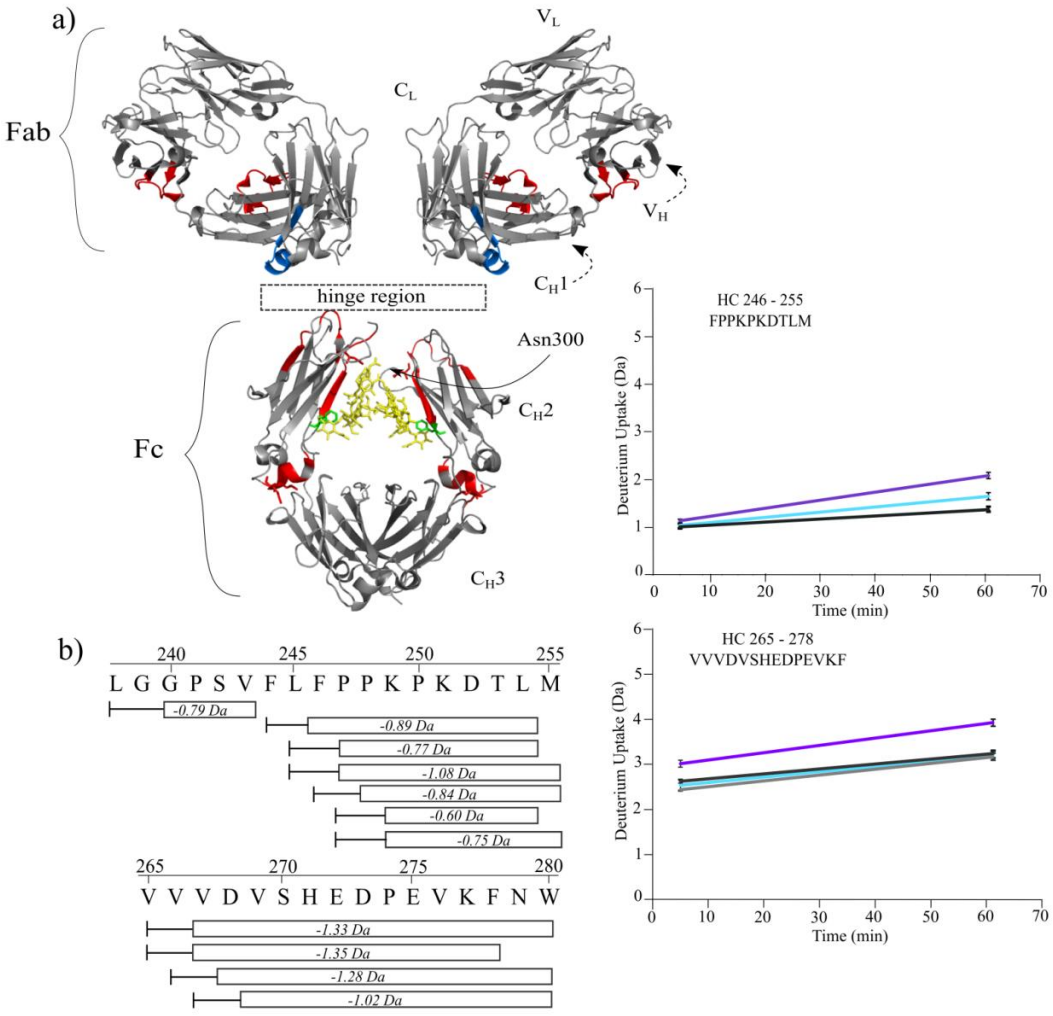


Figure 62 a) Resolved areas of structural change upon endo-deglycosylation of NISTmAb mapped onto Fab and Fc crystal structures (PDB:5K8A [238] and 3AVE [278] respectively). Regions identified as showing reduced (blue) and increased (red) HDX rates are highlighted; residue Phe-246 is highlighted in green sticks, indicated to stabilise the C_H2 domain via interaction with GlcNAc residues. b) Overlapping fragments of interest monitored by HDX-MS. Differential measurements between NISTmAb RM and DeGly NISTmAb are indicated in Da. c) Relative uptake plots of two peptides in which structural changes were identified. Data presented show NISTmAb RM vial 1 (grey), NISTmAb RM vial 2 (black), DeGly NISTmAb (purple) ExoGal NISTmAb (cyan), standard deviations of triplicate measurements are represented by error bars.

3.3.4.3 Comparability of NISTmAb RM and ExoGal NISTmAb HDX-MS measurements

Comparison of differential HDX-MS measurements between the NISTmAb RM control and ExoGal NISTmAb with thresholds (Figure 60 b, Figure 61 b and Table 16), revealed eight HC peptides exhibiting statistically significant structural changes following β-1,4-degalactosylation. None of the LC and all other HC regions appeared unaffected, demonstrating the large majority of the HOS structure was unchanged. These regions of structural change are indicated in Figure 63.

Increased structural flexibility for the region HC 248-255 (Figure 60 b, region IV) located in the C_H2/C_H3 domain interface was also observed for the NISTmAb RM upon degalactosylation, though with HDX rates increased between 0.44-0.57 Da less magnitude relative to Degly NISTmAb. No structural changes were observed for HC 238-245 and HC 265-280 (Figure 63), in contrast to the DeGly NISTmAb, suggesting that the conservation of the interaction between the phenyl ring of Phe-246 and the GlcNAc moiety results in better overall stabilisation of the C_H2 domain. In spite of this conserved interaction, the data suggest a small destabilisation of structure for the HC 248-255 region consistent with the loss of an interaction between the terminal galactose moiety and Lys-249 residue [268]. In contrast to the effect of total glycan removal for the Degly NISTmAb study, dynamics in the Fab region, HC 5-20, HC 66-82 and HC 159-177 (Figure 60 b regions I, II and III), appear to be decreased due to galactose removal. Figure 61 b illustrates that NISTmAb RM degalactosylation has no significant change on LC HOS structure.

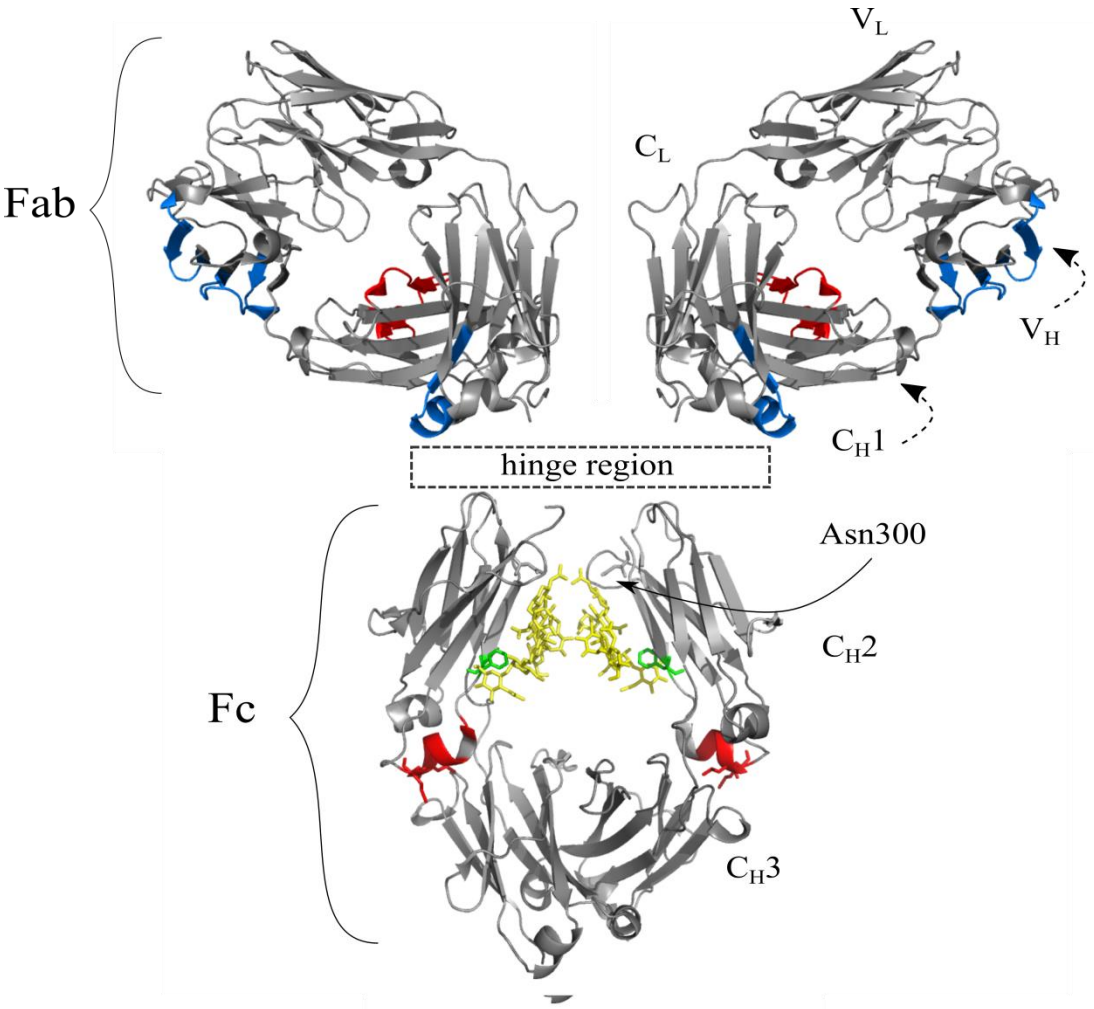


Figure 63. Resolved areas of structural change upon β 1,4-degalactosylation of NISTmAb RM mapped onto Fab and Fc crystal structures (PDB:5K8A [238] and 3AVE [278] respectively). Regions identified as showing reduced (blue) and increased (red) HDX rates are highlighted; residue Phe-246 is highlighted in green sticks, and glycan structure in yellow.

3.3.4.4 Effect of glycosylation on intrinsic HDX-MS rates

The acetoamido groups and N-linking asparagine of the N-glycosylation site will undergo deuteration during HDX [279, 290], and in light of the absence of both of these in the DeGly NISTmAb peptides that contain the glycosylation N300 were therefore not considered in the overall differential HDX-MS analysis.

However comparing the uptake of the peptide HC 281-303 (Figure 64), which spans the glycosylation site Asn300, we observe, as expected, increased uptake for the NISTmAb RM and ExoGal NISTmAb samples compared to that of the DeGly NISTmAb.

It is worth noting deuterium incorporation was equivalent across the two different glycosylation states, this is consistent with the experimental and theoretical observations that the major glycoforms present across the NISTmAb RM and ExoGal NISTmAb samples both share 5 acetoamido groups known to incorporate deuterium (Figure 54) [290]. These data would tend to suggest also that structural changes within the C_H2 domain, associated with removal of galactose residues are conserved to the HC 248-255 region and do not affect the glycosylation site directly.

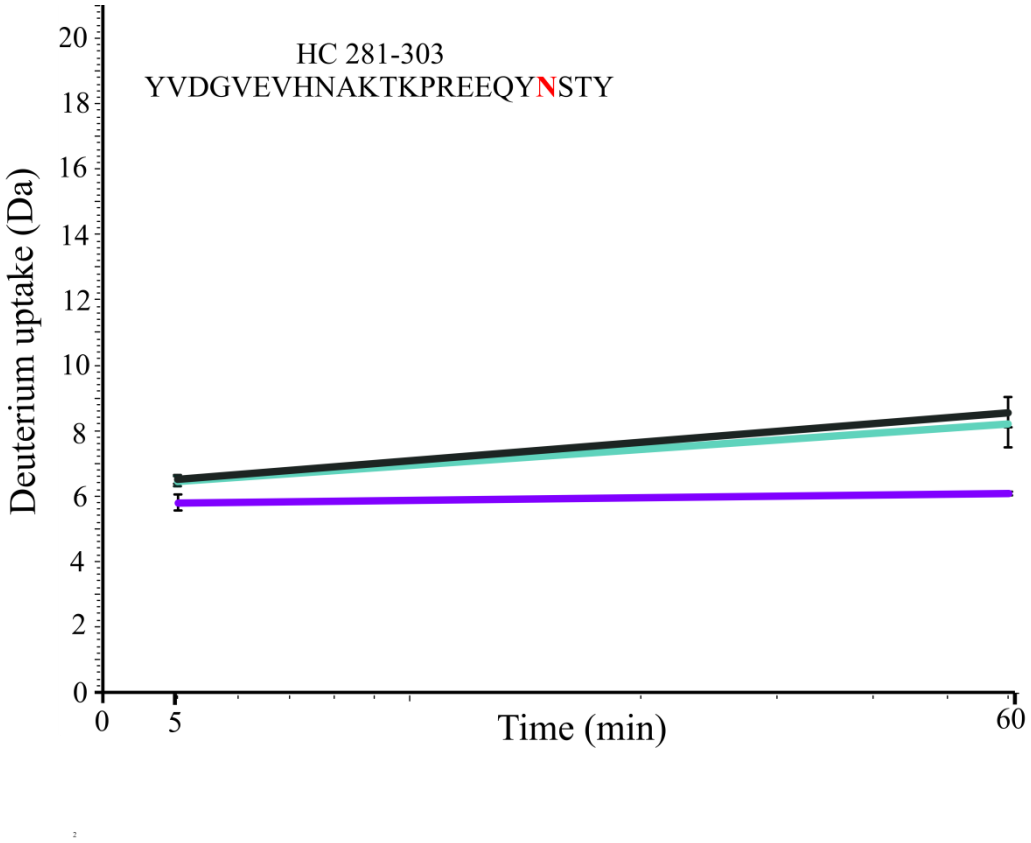


Figure 64. Deuterium uptake plot for the Fc-glycan site of NISTmAb RM. Data presents the uptake rates for the HC peptide spanning residues 281– 303 for the NISTmab RM (black), ExoGal NISTmAb (blue) and Degly NISTmAb (purple) following incubation in deuterium for 5 and 60 min, error bars represent standard deviations (n=3).

3.3.4.5 Comparability of HDX-MS measurements of Fc-glycan variants

Comparison of the localised relative structural changes induced by variation in Fc-glycan structure of the NISTmAb RM as identified in both NISTmAb DeGly and ExoGal NISTmAb samples (Figure 62 a and Figure 63, Table 16), demonstrates a larger HOS change as a result of complete removal of the glycan structure.

Structural changes were identified in both Fc and Fab portions of the molecule for both Fc-glycan variants, indicating potential stability and efficacy implications for the two portions respectively. Such as in the Fab region, where contrasting structural changes, of increased and decreased flexibility for regions HC 5-20 and HC 159-177 were observed for Degly and ExoGal samples respectively, both of which could influence the interaction with FCγRIIIa type receptors [268, 289]. However, increased flexibility of the region situated at the C_H2/C_H3 domain interface was observed in both samples, demonstrating the significance of the interaction between the terminal galactose moiety and Lys-249 residue for the HOS of IgG1s. In the degalcosylated form, the HOS remains unchanged for HC 238-255, a region associated with IgG1 structure degradation, implicating potential different stabilities across the two variants.

Overall the variation in number and magnitude of structural changes observed across the variants, i.e. larger HOS changes upon deglycosylation compared to degalactosylation, promotes the use and suitability of this model system for establishing the sensitivity of analytical technologies towards structural changes.

Table 16. NISTmAb RM peptides identified as showing significant changes due to variation in Fc-glycan structure by HDX-MS differential analysis. Changes are expressed as increases (+) and decreases (-) of HDX relative to NISTmAb RM, with no significant change is indicated by blank input.

Identifier			Relative HDX change	
Peptide, i	**Peptide ID**	**Sequence**	**DeGly**	**ExoGal**
1	HC 5-17-	RESGPALVKPTQT		-
3	HC 5-19-	RESGPALVKPTQTLT	+	-
4	HC 5-20-	RESGPALVKPTQTLTL	+	-
14	HC 66-82-	KDRLTISKDTSKNQVVL		-
15	HC 70-82-	TISKDTSKNQVVL		-
24	HC 159-166-	VSWNSGAL	+	
25	HC 159-170-	VSWNSGALTSGV	+	
26	HC 159-177-	VSWNSGALTSGVHTFPAVL	+	-
27	HC 162-177-	NSGALTSGVHTFPAVL	+	
37	HC 238-244-	LGGPSVF	+	
39	HC 244-254-	FLFPPKPKDTL	+	
40	HC 245-254-	LFPPKPKDTL	+	
41	HC 245-255-	FLFPPKPKDTLM	+	
42	HC 246-255-	FPPKPKDTLM	+	+
43	HC 247-254-	PPKPKDTL	+	+
44	HC 247-255-	PPKPKDTLM	+	+
48	HC 265-278-	VVVDVSHEDPEVKF	+	
49	HC 265-280-	VVVDVSHEDPEVKFNW	+	
50	HC 266-280-	VVDVSHEDPEVKFNW	+	
51	HC 267-280-	VDVSHEDPEVKFNW	+	
16	LC 122-131-	QLKSGTASV	+	

3.3.5 Native IMS-MS method optimisation

Gas-phase unfolding and increased peak resolution as a result of increased sample and trap cone voltages was reported in previous IMS-MS studies of NISTmAb RM using a previous generation instrument (Synapt G2) [291]. To establish source conditions suitable for retaining a native conformation whilst achieving maximum resolution, the

NISTmAb RM control was infused under a range of sample cone voltages between 50 - 150 V (Figure 65). An increase in peak resolution is observed with sample cone voltage, culminating in the appearance of the individual Fc-glycan states at the highest sample cone voltage of 150 V (Figure 65 b), consistent with the previous study [291]. A single Gaussian CSD (+ 20 – 25) was evident across the whole range of sample cone and trap voltages, suggesting no large scale global unfolding. However there does appear to be a slight shift from a centroid of + 23 value to + 24 at sample cone voltages in excess of 150 V suggesting a slight change in globular folding due to increased energy consistent with the charge state residue model [76]. Extraction of the IMS-MS t_D for the + 23 base peak, across the voltage range (Figure 66), reveals the appearance of a secondary ATD peak, of increased t_D (~ 19 ms) for the 150 V cone voltage confirming gas-phase unfolding of the protein at this voltage.

Improved resolution was also observed when increasing the instrument trap voltage (Figure 65 b), with the appearance of glycoforms evident for trap voltages in excess of 55V. Similarly, gas-phase unfolding was also apparent with the a appearance of a secondary longer ATD at trap voltages in excess of 50 V (Figure 66) when considering the + 24 charge state. For cone voltages $\leq$ 100 V and trap voltages $\leq$ 40 V, it is possible to observe a single ATD peak. Therefore to retain the native conformation of NISTmAb RM moving forward sample cone voltage was maintained at 75 V and trap voltage at 3 V, unless states otherwise.

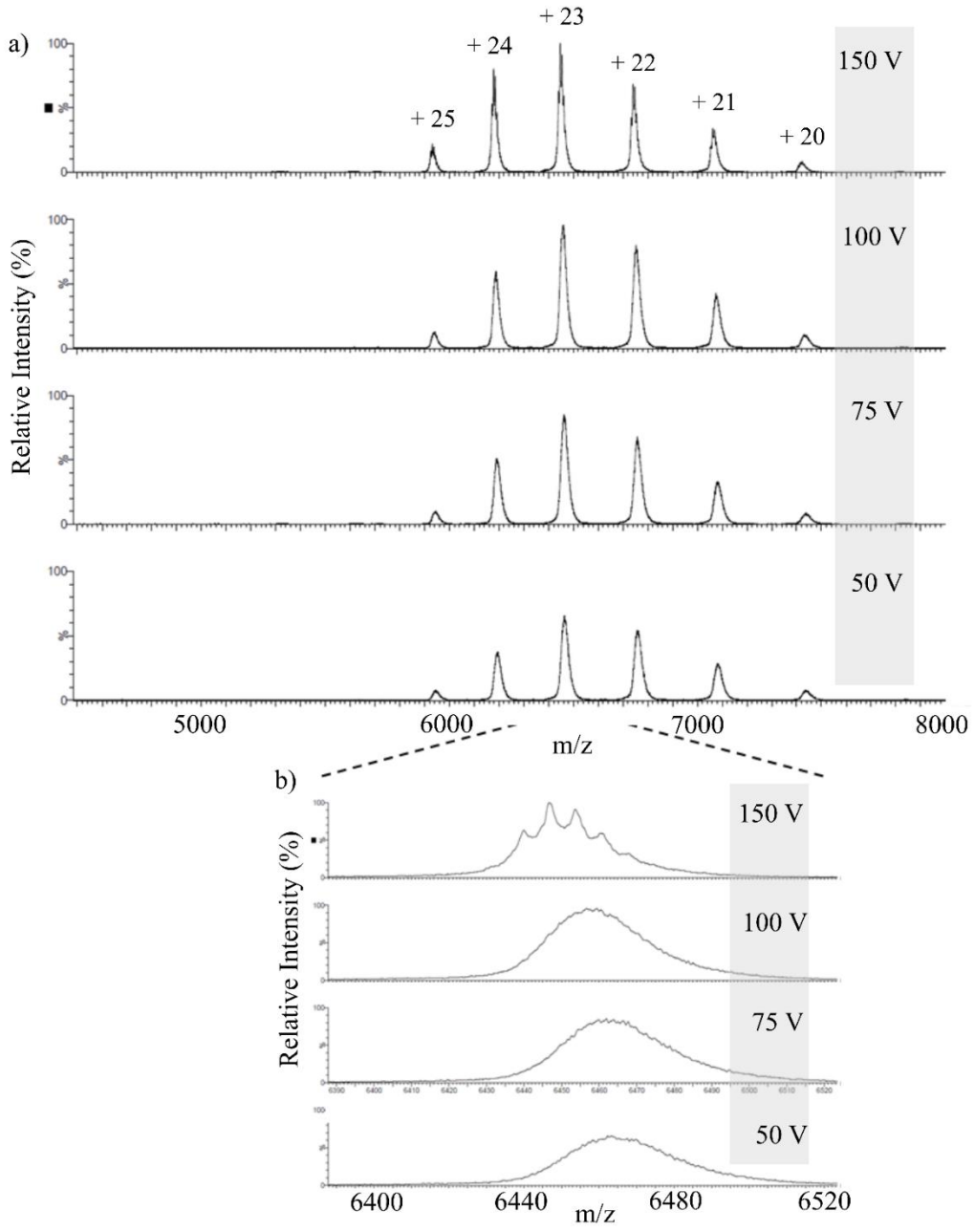


Figure 65. Comparison of the native MS spectra of NISTmAb RM acquired with increasing sampling cone voltages across the 50 – 150 V range. Spectra were acquired under native buffer conditions a) the full charge observed CSD observed b) MS-spectra of the + 23 charge state. Increased resolution of the major glycoforms is observed with increased sampling cone voltage.

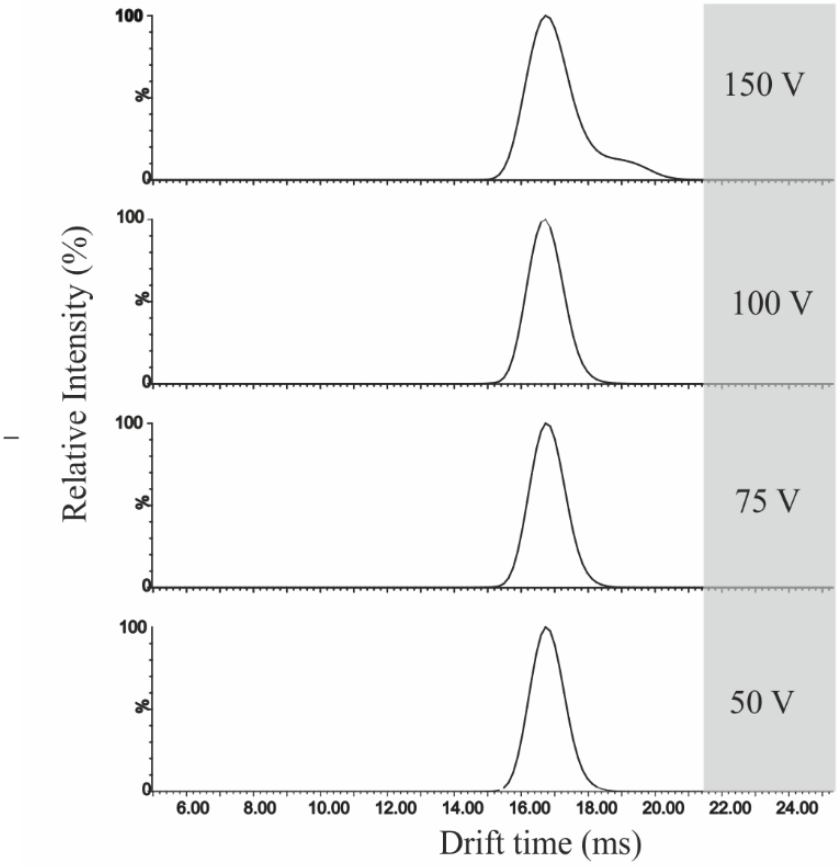


Figure 66. Comparison of the extracted ATD of the + 23 charge state of NISTmAb RM, acquired with increasing sampling cone voltages across the 50 – 150 V range, under native buffer conditions. At sampling voltages > 100 V a second ATD with increased t_D ~ 19ms is observed, indicative of gas-phase unfolding.

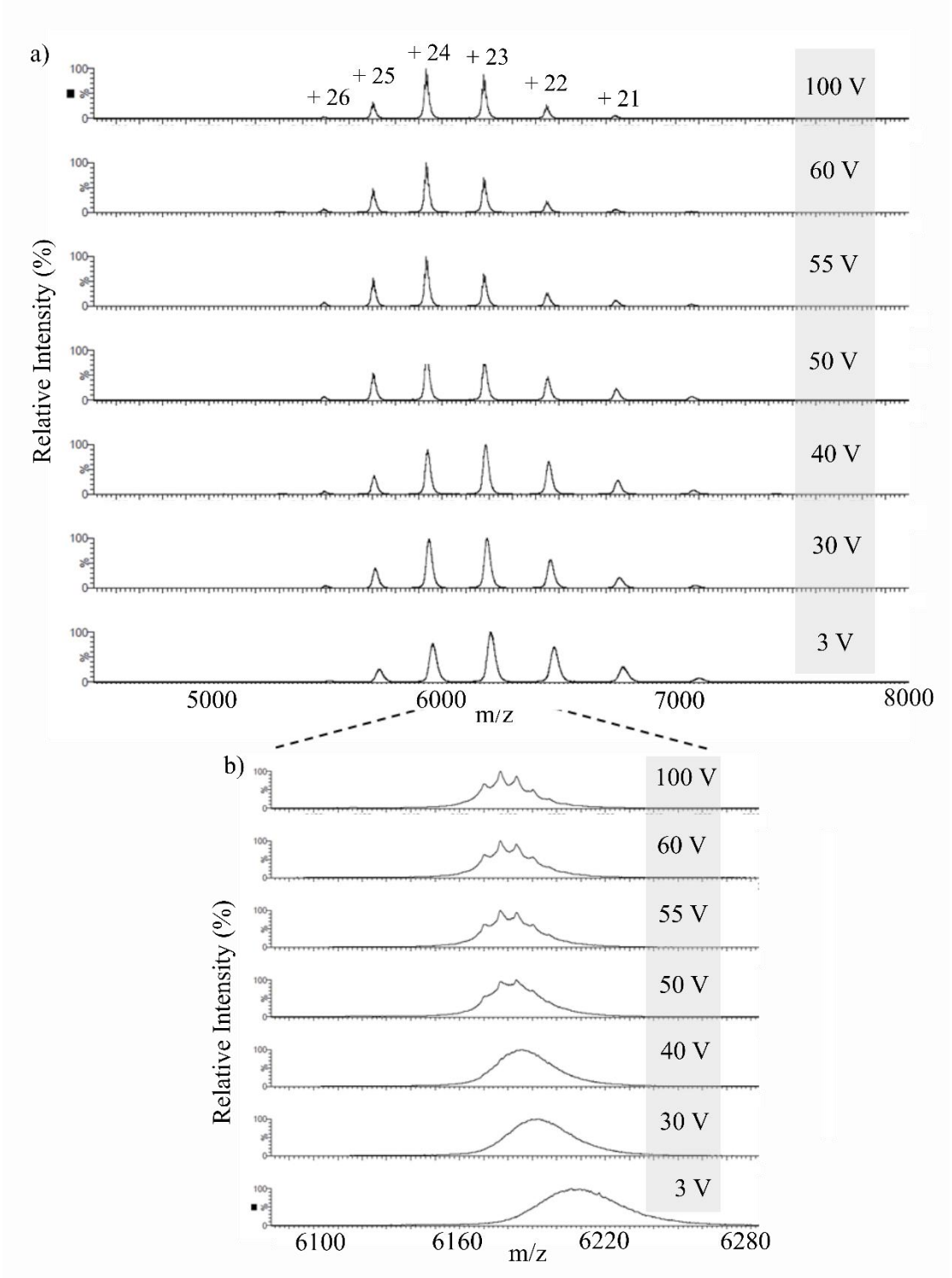


Figure 67. Comparison of the native MS spectra of NISTmAb RM control acquired with increasing trap voltages across the 3 – 100 V range. Spectra was acquired under native buffer conditions a) the full charge observed CSD observed b) MS-spectra of the +24 base peak. Increased resolution of the major glycoforms is observed with increased sampling cone voltage.

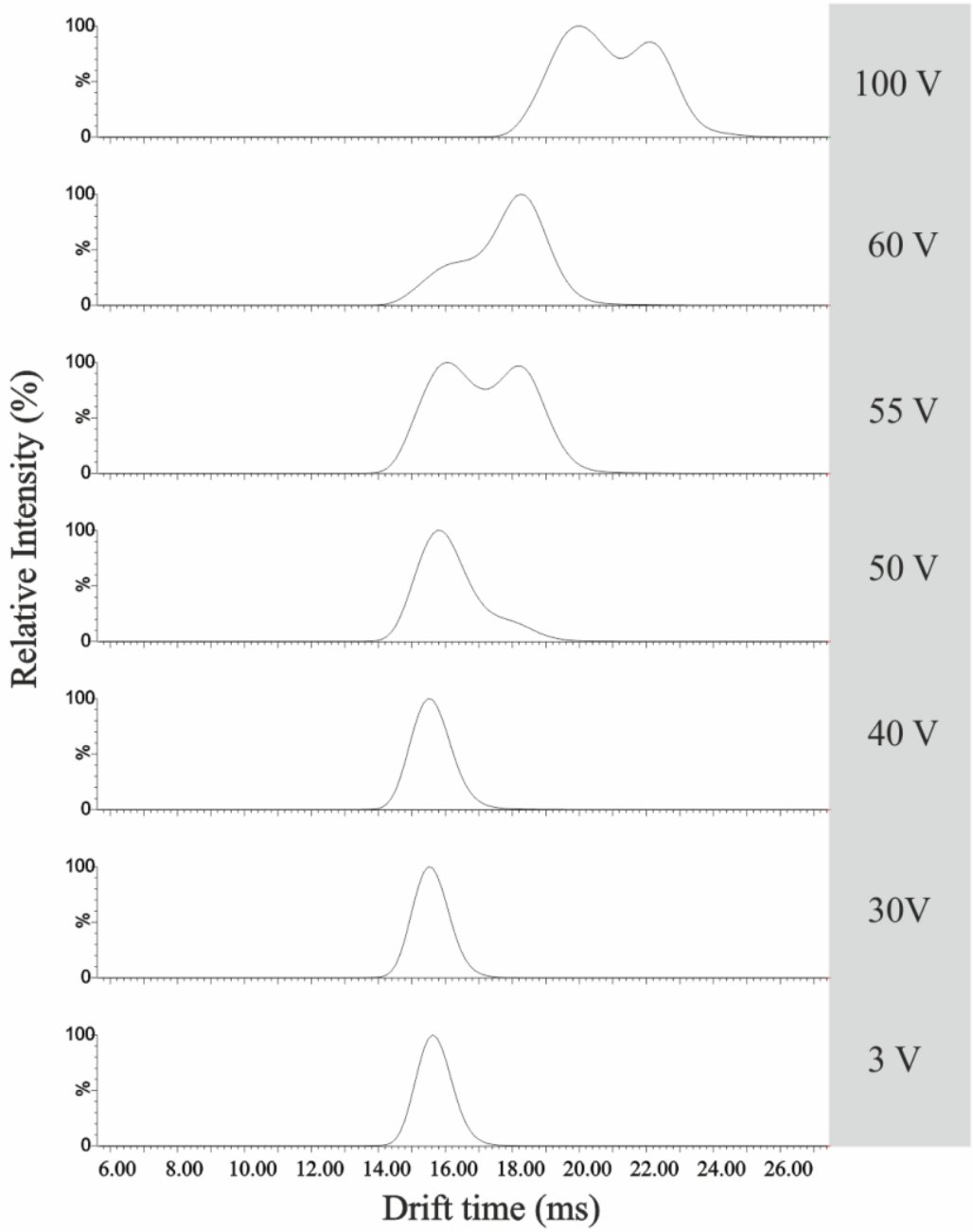


Figure 68. Comparison of the extracted ATD of the + 23 charge state of NISTmAb RM, acquired with increasing trap voltages across the 3 – 100 V range, under native buffer conditions. At trap voltages > 40 V a second ATD with at an increased t_D ~ 19ms is observed, indicative of gas-phase unfolding A second unfolded state, with further increase ATD ~ 22 ms is observed at 100 V.

3.3.6 Native IMS-MS measurements of NISTmAb RM and structure prediction

The native IMS-MS spectra of NISTmAb RM (Figure 69 b) reveal a single narrow Gaussian CSD with charge states spanning from + 20 - 27, indicative of a single solution conformation, and a globular-like folded protein. This is corroborated by the single ATD observed for each charge state (Figure 69 a) .

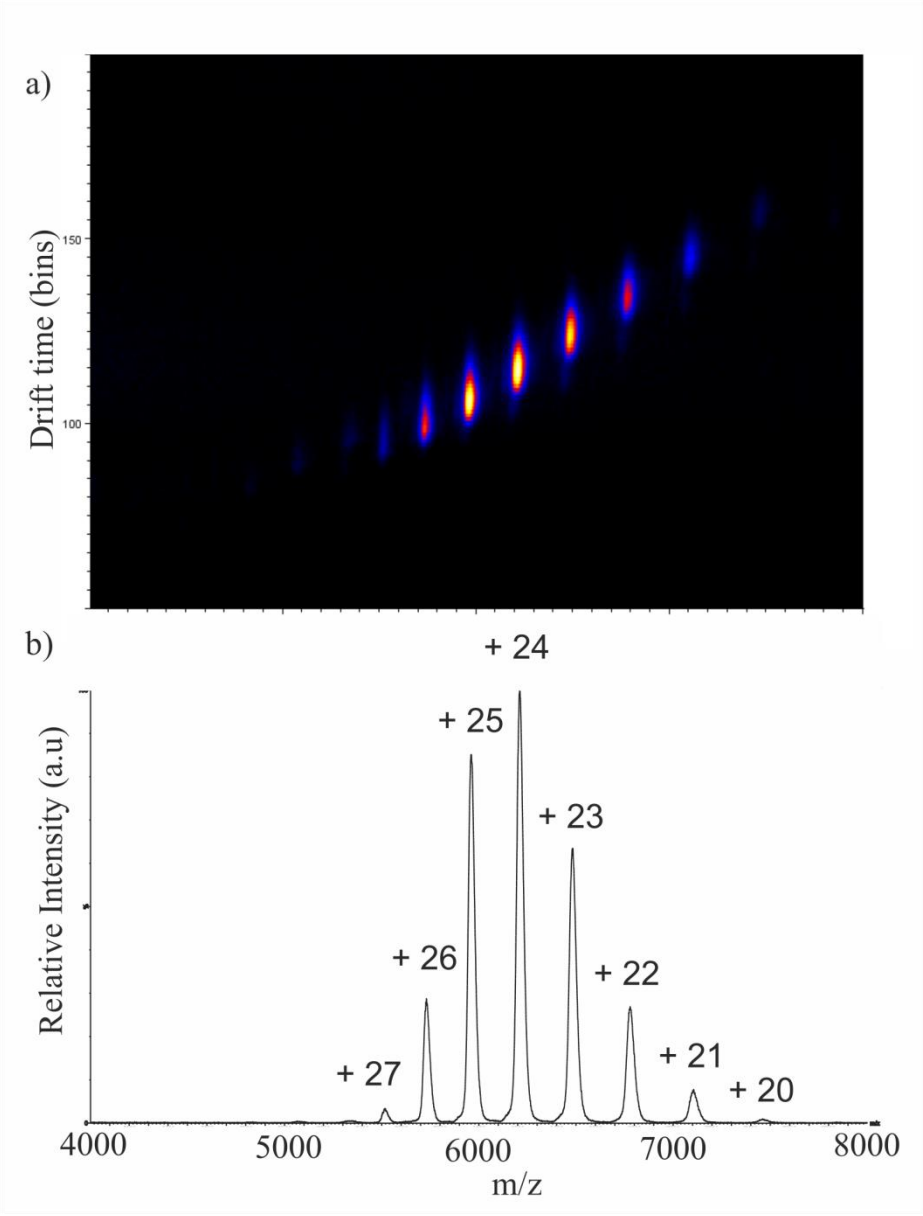


Figure 69. Native IMS-MS analysis of NISTmAb RM. a) Drift time mobiliogram of NISTmAb RM IMS-MS measurements reveal a single conformational species for each charge state b) a single Gaussian CSD is observed for the corresponding mass spectra.

3.3.7 Native IMS-MS to identify changes in HOS induced by Fc-glycan variation

IMS-MS t_D measurements of the NISTmAb RM and Fc-glycan variants were compared for differences in HOS and globular stability Table 17. All measurements are expressed as an average of three inter-day measurements and the standard deviation given.

The same single narrow Gaussian CSD (+20 - 27) was observed for all three samples, indicating no large scale unfolding due to change in Fc-glycan structure and the existence of a single conformation under these native-like conditions. No significant differences in t_Ds were observed across NISTmAb RM and ExoGal NISTmAb measurements, showing equivalent global structures between the samples.

However under these native-like conditions, an average 3.6% decrease in ATD (across all charge states) was resolved for the DeGly NISTmAb relative to the NISTmAb RM, demonstrating a decrease in $^{TW}CCS_{N2}$ and a more compacted globular structure for the deglycosylated form.

It is worth noting that though mobility is proportional to the reduced mass of the ion, μ_{red} , the 2.1% mass difference between glycosylation states equates to a 0.0002% difference in μ_{red} which has a negligible effect on t_D for these large mass proteins. This indicates that this decrease in ATD is indeed as a result of a change in HOS [53]. Measured ATDs were equivalent for NISTmAb RM and ExoGal NISTmAb under these conditions, revealing similar Ω values.

The presence of a glycan structure within the cavity between the two C_H2 domains may increase the distance between the two domains as a result of steric hindrance, removal of these glycans could result in the gas-phase collapse of this cavity and the globular Ω compaction observed for deglycosylated samples. Similarly Upon et al. [292] observed a ~ 3.5% decrease in ATD upon the deglycosylation of the therapeutic Herceptin®, suggesting this phenomenon is consistent across other IgG1s. These data are consistent with observations from HDX-MS analysis showing significant increase in flexibility within the interface of C_H2/C_H3 domains for the deglycosylated form, which may result in the gas-phase collapse described. Further still this would support the idea that the presence of glycosylation enhances and stabilise, the open conformation of C_H2 domain necessary for FcγR binding [269].

Table 17. Comparison of IMS-MS measurements of NISTmAb RM control, DeGly NISTmAb and ExoGal NISTmAB. Average observed drift times (t_D) and SDs expressed for triplicate, n=3, interday measurements. **m/z* values are expressed considering the mass of the most abundant glycoforms of NISTmAb RM control, G0F/G1F glycosylation state.

Charge state (z)	*m/z**	NISTmAb RM		DeGly NISTmAb		ExoGal NISTmAb	
		t_D (ms)	SD (ms)	t_D (ms)	SD (ms)	t_D (ms)	SD (ms)
20	7410.0	21.7	0.3	21.1	0.3	21.8	0.1
21	7057.1	20.0	0.2	19.4	0.2	20.1	0.1
22	6736.3	18.5	0.2	17.8	0.1	18.5	0.0
23	6443.4	17.0	0.1	16.4	0.1	17.0	0.1
24	6175.0	15.7	0.1	15.1	0.1	15.7	0.2
25	5928.0	14.5	0.1	14.0	0.1	14.5	0.2
26	5700.0	13.6	0.1	13.1	0.0	13.6	0.2
27	5488.9	13.0	0.2	12.5	0.3	13.1	0.1

3.3.7.1 CIU-IMS-MS comparisons

Quadrupole selection was performed of the corresponding +24 base peak under low voltage native conditions, with selection made based on the criteria of ion intensity, compact structure [95], and CIU-IMS-MS experiments performed by ramping trap cell collision voltages from 3 – 100 V. At the lower collision voltage of 3 V, a single Gaussian ATD was observed indicating a native folded state (T1‡) (Figure 70) present in all samples. Similarly as stated previously the ATD of DeGly NISTmAb showed a ~ 3.6% compaction at native conditions for a trap voltage of 3 V, measured ATDs were equivalent for NISTmAb RM and ExoGal NISTmAb under these conditions, again revealing similar Ω values, under these conditions.

For measurements of all samples performed at voltages 3 –40 V, only the T1‡ state was observed (Figure 71). Increasing voltages to ~ 44 V, reveals the presence of a tailing peak with increased ATD, indicating the presence of a less folded species (T2‡) at this higher voltage, though the T1‡ state remains the predominant species. Figure 70, displays the overlayed mobilograms measured at collision voltages of 44 - 55V

for all samples. Comparison of the relative proportion of T1‡/ T2‡ states between these voltages, reveals a higher proportion of unfolded T2‡ state for the DeGly NISTmAb compared to the NISTmAb RM. A general trend of increased proportion of T2‡/ T1‡ is observed across DeGly > ExoGal > NISTmAb RM, presenting a correlation between the amount of collisional induced activation required for unfolding the native conformation and increased glycan structure, similar to studies of other IgG1s [95, 292]. This decreased stability is apparent across the whole 44 – 55V range, corroborating the destabilisation of the C_H2/C_H3 domains revealed by HDX-MS measurements caused by deglycosylation and degalactosylation. Higher similarity ATD profiles across the CIU-IMS-MS experiment between the NISTmAb RM and ExoGal NISTmAb samples, indicates more comparabile HOS structures. Finally an additional more unfolded species (T3‡) was evident at higher collision energies ~ > 75 V for all samples (Figure 71) consistent with previous IMS-MS studies of NISTmAb RM [291].

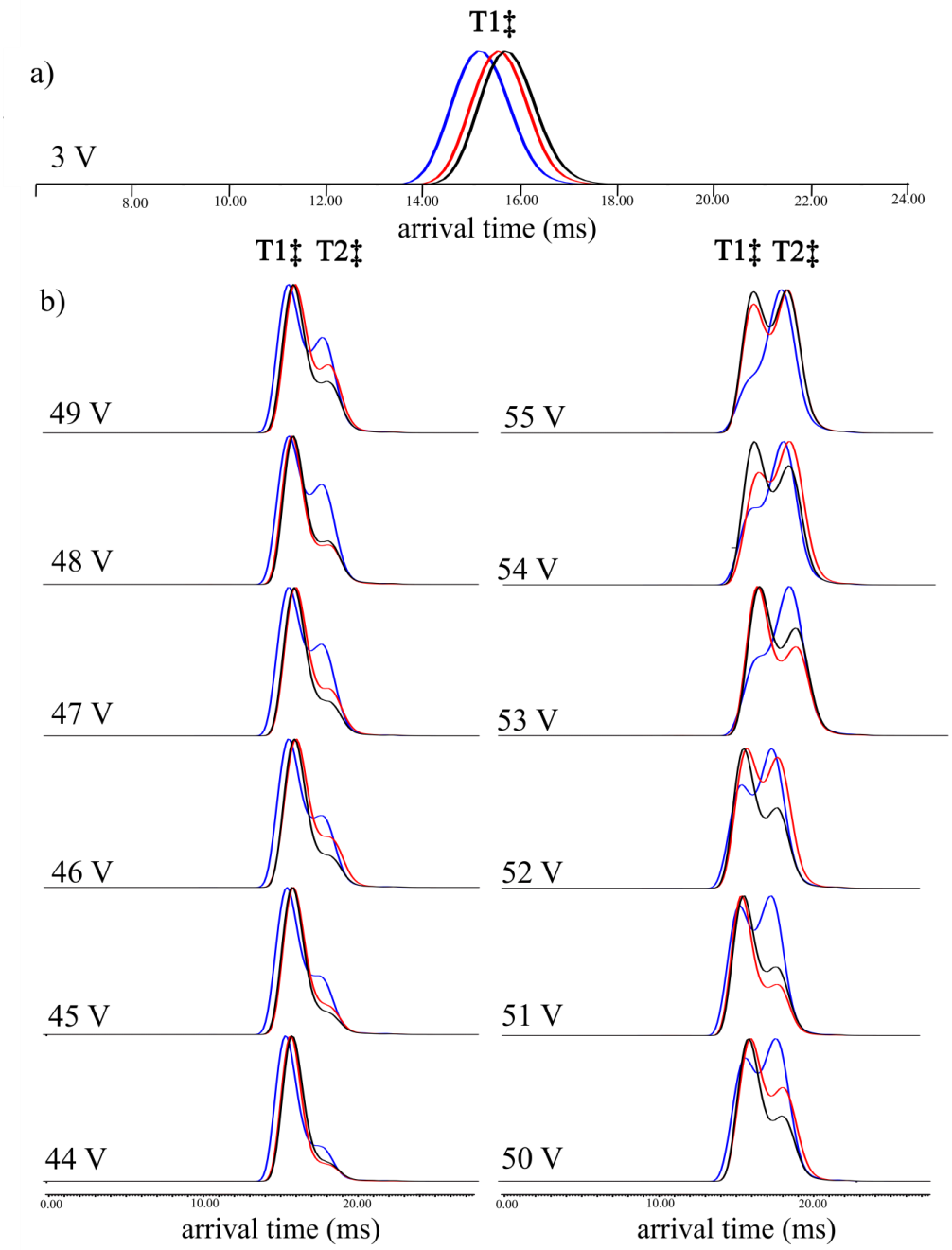


Figure 70. Collision-induced unfolding (CIU-IMS-MS) plots of the + 24 charge state of NISTmAb RM (black), DeGly NISTmAb (blue) and ExoGal NISTmAb (red). a) A single folded gas-phase conformation (T1‡) is observed at low collision voltage 3 V b) increased presence of a unfolded gas-phase conformation (T2‡) with increased arrival time is observed as collision voltages are ramped from 44 – 55 V. Samples were infused using equivalent experimental conditions and under native buffer conditions (50 mM ammonium acetate). Data is displayed as arrival time (ms) versus normalised intensity.

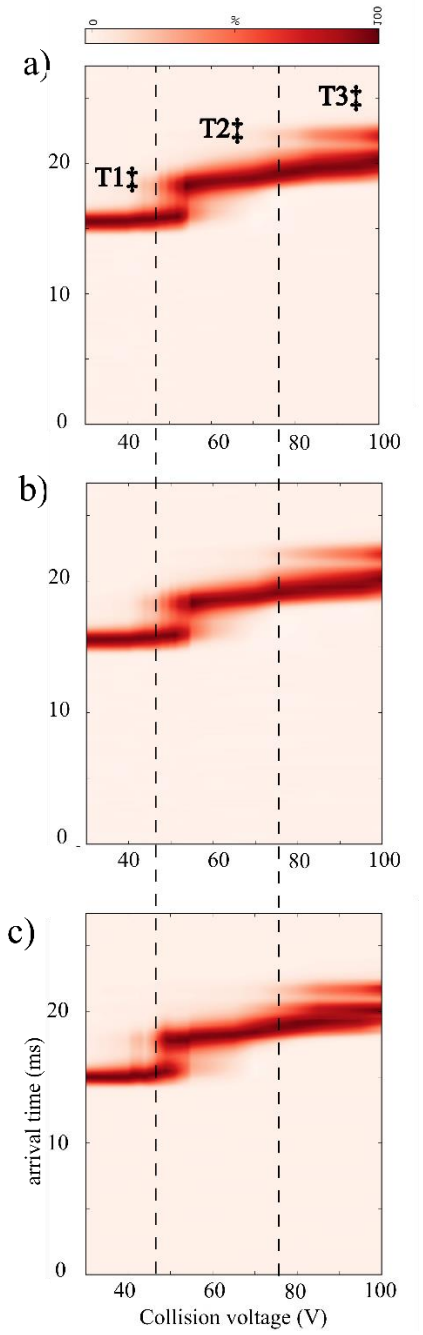


Figure 71. Collision-induced unfolding (CIU-IMS-MS) plots of a) NISTmAb RM b) ExoGal NISTmAb and c) DeGlyl NISTmAb. Quadrupole selection of +24 charge state was performed and trap cell activation induced using a voltage range of 30 - 100 V. Data are displayed in three dimensions as arrival time (ms), trap cell collision voltage versus normalised intensity. Arrival time distributions are indicated for three folded states compact (T1‡), unfolded intermediate (T2‡) and unfolded (T3‡). Plots were generated with use of PULSAR processing software [98].

3.4 Conclusions

mAbs have revolutionised modern-day therapeutics, with their complex and tuneable HOS structures opening the gateway for specific targeted therapies for the treatment of a vast array of diseases. Current regulation supports the findings that changes in CQAs, such as glycosylation, of therapeutic mAbs can result in changes in HOS and effect safety, efficacy and potency of the therapeutic. Due to their complex nature, monitoring for the HOS changes of mAbs can be analytically challenging and the technologies employed must be both validated and demonstrated to provide adequate measurements. There is a need for reference datasets to support the use of the NISTmAb RM in method validation, QC monitoring and comparability studies of mAb-based biotherapeutics. Indeed the NISTmAb RM and the protocol described here to generate Fc-glycan variants, has facilitated the validation of the statistical approach developed, in Chapter 2 using rhGH, to establish HDX-MS measurement thresholds, on biomolecules of increased sized and complexity, relative to rhGH RM.

The simple enzymatic preparation of two Fc-glycan variants from the NISTmAb RM described here offers a candidate system for assessing the suitability of more novel analytical methodologies, such as HDX-MS and IMS-MS, towards the characterisation of small structural changes. Whilst considering the methodological measurement variability, localised, mainly destabilising, structural changes have here been characterised by HDX-MS for both Fc-glycan variants. Differences in global structure and decreased stability between the NISTmAb RM and fully deglycosylated form were also corroborated by CIU-IMS-MS experiments, with more facile CIU following deglycosylation validating the destabilisation of NISTmAb RM structure as indicated by HDX-MS analysis. Both analytical measurements showed complete deglycosylation to have a larger effect on HOS relative to degalactosylation. These data promote the use of these Fc-glycan variants for validating analytical methods for the detection of HOS measurements of mAbs and also the suitability of HDX-MS and IMS-MS as valuable analytical methods within the biopharma toolbox.

Chapter 4

Assessment of fast photochemical oxidation of proteins (FPOP) mass spectrometry for structural analysis of biotherapeutics

Chapter 4
Assessment of fast photochemical oxidation of proteins (FPOP) mass spectrometry for structural analysis of biotherapeutics

4.1 Introduction

MS protein footprinting methodologies have been shown to be effective tools for providing high resolution and highly sensitive protein HOS measurements, making them strong candidates for use in the analysis of protein biotherapeutics. Contrary to IMS-MS measurements which provide global information, both HDX-MS and FPOP-MS provide localised structural insights into secondary and tertiary structures.

FPOP-MS also offers its own unique advantages over alternative footprinting technologies, such as HDX-MS. For instance, the irreversible nature of the labelling, allows for the storage of labelled samples, the opportunity to use more traditional proteomic workflows that offer higher sensitivity, flexibility (including proteases, chromatography and fragmentation techniques) compared to more restricted HDX-MS workflows. In addition samples can be reanalysed and even be subjected to sample clean up strategies such as solid phase extraction. These elements in particular open the potential of the technique towards more complex and physiologically relevant matrices. Indeed, FPOP-MS, in recent years, has been used to study protein-protein interactions *in-cell* [189, 293] and within multi-organ systems, *in vivo* [294].

This flexibility of analysis also makes quantitative downstream workflows, such as the use of isotopically labelled internal standards, more applicable, opening the door to quantitative structural measurements. If this was achieved, this in turn could help bridge the traceability chain between the active folded form of a protein and the primary sequence of a protein. A feat that would be particularly advantageous in diagnostic areas such as neurodegeneration or amyloidosis, where intermediates of protein folding states, folded and unfolded state can be clinical biomarkers. Isolation and quantification of specific folding state can represent an important step in understanding the implications of this. So far no methods can quantify the structure of a protein and it is possible that FPOP-MS could bridge this gap. However, the technique is very much in its infancy, and though it shows potential, requires more

validation and standardisation of workflows in order to move from academia into an industrial setting to support the bio-therapeutic area.

A “traditional” capillary flow FPOP labelling set up, involves the delivery of a sample through a silica capillary via a syringe pump and irradiation occurs through a transparent window on the capillary [182, 183, 186, 188] (Figure 72 a). This workflow is non-automated and at times a laborious manual task. It was hypothesised that in order to simplify the FPOP labelling experimental workflow, the irradiation of the sample could occur in a vessel, such as an Eppendorf, with the irradiation occurring by in a single shot of irradiation to the sample. This workflow would bypass the need to flow the sample via a capillary. Quenching would then occur via pipette and the mixture vortexed. It was also hypothesised that due to the “static” nature of the irradiation, the labelling would be more reproducible and easier to optimise. Figure 72 b gives a schematic of the proposed novel labelling procedure. If effective it was thought that this “single-shot Eppendorf” FPOP labelling approach would be higher throughput and potentially more reproducible due to the simplified sample handling. To current knowledge, no instances of this type of FPOP labelling workflow appear to exist in the literature.

In this Chapter, the applicability of FPOP-MS measurements to define protein structural changes was assessed using the rhGH model system developed in Chapter 2. Two FPOP labelling strategies, capillary flow and a novel single-shot Eppendorf methodology, first illustrated here, were optimised and implemented for the analysis of the rhGH model system at both the intact and peptidic level. Two differing MS instrumental workflows were utilised for peptide level analysis and a semi quantitative approach applied. The results of the two labelling strategies were compared and discussed in order to validate the single shot Eppendorf approach. Data acquired here were also compared with HDX-MS and IMS-MS measurements, and the benefits and drawbacks of the measurements discussed. Finally, this data set provides further characterisation of the structural changes incurred by rhGH in the presence of zinc, further promoting the use of the RM and protocol for protein HOS measurements.

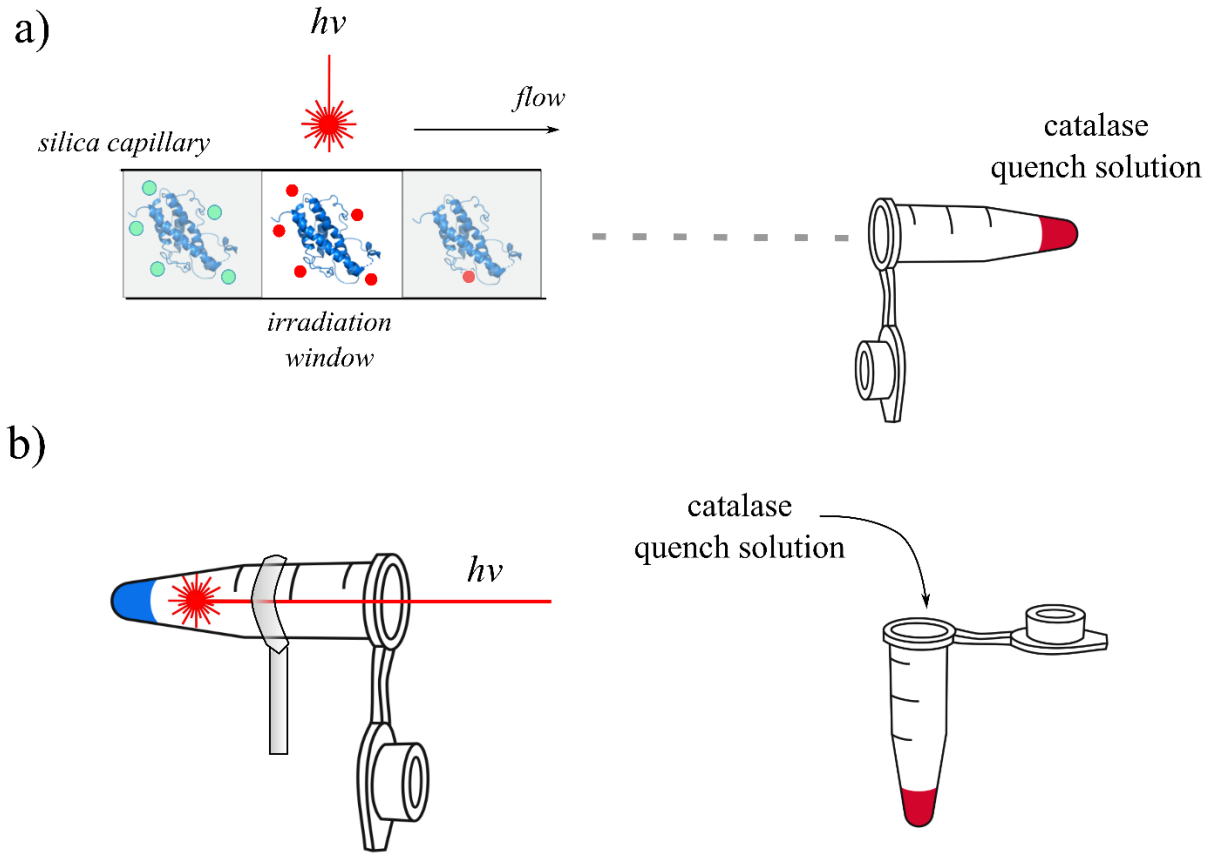


Figure 72. Schematic of two FPOP irradiation workflows. The standard a) capillary flow set up, introduces the sample (protein, scavenger and H_2O_2), into a laser beam path via a capillary flow. Exposure to the laser source occurs through an irradiation window, and the irradiated sample is subsequently quenched as it flows into the catalase containing quenching solution. Alternatively a novel labelling workflow, termed the b) single-shot Eppendorf set up, is proposed, in which irradiation occurs as a single shot directly into the sample-containing Eppendorf and the catalase quenching solution is pipetted into the sample immediately post-irradiation, by-passing the need for any capillary or flow system.

4.2. Materials and Methods

4.2.1. Materials

All chemicals and standards were purchased as described in section 2, unless otherwise stated. MS grade trypsin was purchased from Promega [Madison, USA].

4.2.2. Protein sample preparation

Prior to FPOP labelling, 7.2 µM rhGH and rhGH:Zn samples were prepared as described in section 2.2.1.1 to obtain rhGH control and rhGH:Zn 1:2, 1:10 and 1:20 samples in 10 mM KPBS (pH 7.4). 28 µM protein samples were prepared using the appropriate scale up. Protein solutions were left to equilibrate for 1 hour on the bench prior to FPOP labelling.

4.2.3. FPOP irradiation source set up

The FPOP laser set-up used [106, 186], consists of a Complex 50 Pro KrF excimer laser operating at a 248 nm wavelength (Coherent Inc, Ely, UK) and laser beam width of < 3 mm at point of irradiation. Two different experimental workflows for FPOP labelling were utilised, as described in Figure 72, one using a traditional capillary flow FPOP system in which the sample (protein + H_20_2 + scavenger) is delivered into the irradiation path via a syringe pump and a silica capillary and the other, a single-shot Eppendorf workflow, in which a single shot of irradiation is delivered directly onto a sample held in an Eppendorf and positioned directly into the laser beam.

4.2.4. FPOP labelling procedures

4.2.4.1 Capillary flow FPOP labelling procedure

Immediately prior to irradiation, 0.93 µL 5.5% H_20_2 (v/v) was added to 50 µL of protein sample (7.2 µM rhGH) (final concentration H_20_2 of 0.1% v/v) containing 350 µM L-histidine in 10 mM KPBS (pH 7.4). The sample was then infused through a fused silica capillary (100 µm, i.d), at a flow rate of 20 µL/min. Hydroxyl radicals were generated by exposure of the sample (through an etched window) to irradiation from the laser operating at a 15 Hz frequency. Flow rate and laser frequency were set so as to only irradiate the protein sample once. Following laser irradiation, the capillary outflow was collected in an 1.5 mL Eppendorf tube containing 20 µL quenching solution (100 mM L-methionine, 1 µM catalase in 10 mM KPBS, pH 7.4), vortexed and immediately placed on ice.

4.2.4.2 Single-shot Eppendorf FPOP labelling procedure

Immediately prior to irradiation, 0.5 µL 0.5% H_20_2 (v/v) was added to 10 µL of protein (28 µM rhGH) (final concentration H_20_2 of 0.025% v/v) containing 175 µM L-histidine in 10 mM KPBS (pH 7.4), in an Eppendorf. Hydroxyl radicals were generated by exposure of the sample to a single shot of irradiation from the laser

operating at a 1 Hz frequency. Immediately (~ within 20s) following laser irradiation, the sample was combined with 5 µL quenching solution (100 mM L-methionine, 1 µM catalase in 10 mM KPBS, pH 7.4), vortexed and immediately placed on ice.

Post labelling samples (both FPOP strategies) were stored at -20 °C, awaiting intact LC-MS analysis or proteolysis followed by LC-MS/MS analysis.

To account for any background oxidation that may occur due to low levels of hydroxyl radicals present in the H_20_2 solution, control samples were handled in a similar fashion, in parallel to each irradiated sample, yet without exposure to laser irradiation.

4.2.5 Tryptic Digestion

4.2.5.1 Capillary flow FPOP labelled samples

40 µL of each FPOP protein sample (5.76 µM rhGH) was added to 20 µL of digestion buffer (1 M urea, in 500 µM ammonium bicarbonate, pH 8, 0.6% SDS). Disulphide bond reduction was performed by addition of 2 µL of 50 mM TCEP (60 °C for 1 hour), followed by alkylation using an addition of 1 µL of 200 mM methyl methanethiosulfonate (MMTS) (RT for 10 min), before trypsin digestion (37 °C overnight) using a 1:50 (w/w) enzyme:protein ratio. The digested samples were frozen and stored at -20 °C until analysis.

4.2.5.2 Single-shot Eppendorf FPOP labelled samples

5 µL of each FPOP protein sample (11 µM rhGH) was added to 20 µL of digestion buffer (0.1 M urea, in 50 mM ammonium bicarbonate, pH 8, 0.6% SDS). Disulphide bond reduction was performed by addition of 1 µL of 10 mM DTT (60 °C for 1 hour, followed by alkylation by addition of 1 µL of 200 mM idoacetamide (IAM) (RT for 30 min in the dark), before trypsin digestion (37 °C overnight) using a 1:50 (w/w) enzyme:protein ratio. The digested samples were frozen and stored at -20 °C until analysis.

4.2.6 FPOP-MS analysis

4.2.6.1 Intact LC-MS analysis

The total extent of oxidation was measured at the intact protein level by LC-MS analysis. Chromatography was performed using an Acquity H-Class UHPLC system [Waters, Milford, USA]. 1.5 pmol of rhGH samples was injected onto an ACE C4 column [2.1 x 150 mm, 3 µm, Hichrom Reading, U.K]. Chromatographic separation

was achieved using a 300 µL/min flow rate, 50 °C column temperature and a linear gradient starting from 80% A/ 20% B to 35% A/ 65% B over 6 minutes. Mobile phases consisted of aqueous, 0.5% formic acid (A) and 0.5% formic, CH_3CN (B).

Following chromatographic separation intact mass analysis was performed using a Xevo G2-XS Q-ToF instrument [Waters, Milford, USA] in positive ESI mode. Experimental settings were as follows: capillary voltage 3 kV, source temperature 100 °C, sampling cone 100 V, source offset 100 V, desolvation temperature 500 °C. Data were acquired in MS only mode, using sensitivity mode and in the 300-3000 *m/z* range with a 0.2 sec scan time.

4.2.6.2 Bottom up LC-MS/MS analysis

For the localisation and quantification of oxidation modification sites, two differing LC-MS/MS instrumental workflows were utilised for peptide level analysis, one for each labelling workflows as described.

4.2.6.2.1 Capillary flow FPOP labelled samples

Chromatography was performed using an Acquity H-Class UHPLC system. 12 pmol of digested rhGH samples was injected onto an ACE Excel Super C18, (2 µm, 2.1 x 100 mm i.d, both Hichrom Reading, U.K). Chromatographic separation was achieved using a 400 µL/min flow rate, 40 °C column temperature and a linear gradient starting from 97% A/ 3% B to 50% A/ 50% B over 19 minutes. Mobile phases consisted of aqueous, 0.1% formic acid (A) and 0.1% formic, CH_3CN (B).

Following chromatographic separation, MS/MS was performed using a Xevo G2-XS Q-ToF instrument in positive ESI mode. Experimental settings were as follows: capillary voltage 3 kV, source temperature 100 °C, sampling cone 90 V, source offset 80 V, desolvation temperature 300 °C. For peptide identification and oxidation localisation, MS/MS experiments were performed in both DDA and MS^E mode. DDA experiments using a Top 5 CID approach a 0.2 sec scan time, a threshold of 5000, a 4 s dynamic exclusion window in the 50-2000 *m/z* range. MS^E experiments were performed as described in section 2.2.3.2. To increase sensitivity and reproducibility, quantification measurement data were acquired in MS only mode, using sensitivity mode and in the 100-2000 *m/z* range with a 0.5 sec scan time.

4.2.6.2.2 Single-shot Eppendorf FPOP labelled samples

Chromatography was performed using a Vanquish UHPLC system. 3.5 pmol of digested rhGH sample was injected onto an Acquity C18 BEH column [1.0 x 100 mm, 1.8 µm, Waters, USA]. Chromatographic separation was achieved using a 30 µL/min flow rate, 50 °C column temperature and a linear gradient starting from 97% A/ 3% B to 60% A/ 40% B over 36 minutes. Mobile phases consisted of aqueous, 0.1% formic acid (A) and 0.1% formic, CH_3CN (B).

Following chromatographic separation intact mass analysis was performed on a Q-Exactive Plus Orbitrap instrument, using a HESI-II probe source in positive mode. Source conditions included a 320 °C source temperature, 3 kV capillary voltage, 90 S-lens RF, 25 a.u. sheath gas, 12 a.u. aux gas, 80 V in-source CID. Data were acquired in MS/MS mode, using a Top 5 CID approach. Full scan MS experiments: 120 ms maximum IT, $1E^6$ AGC target value, 140 K resolution over a range of 350 – 2000 *m/z*. MS/MS data were acquired in dd-MS^2 mode, 80 ms maximum IT, $2E^5$ AGC target value, 3s dynamic exclusion window, 17.5 K resolution and over a range of 100 – 1500 *m/z*.

4.2.7 Data analysis

4.2.7.1 Oxidation quantification

Deconvoluted mass distributions were extracted from intact protein LC-MS spectra, using MassLynx software v 4.1 and the MaxEnt 1 algorithm in the 18 – 25 kDa range, using a resolution of 0.1 Da/channel.

Quantification of FPOP modifications was achieved by integration of modified and unmodified peaks generated from XIC using the following equation:

Equation 35

$$\%modified = \frac{\sum modified_{Peak}}{\sum unmodified + \sum modified_{All\ Peaks}}$$

Each FPOP sample, i, was normalised against any low levels of oxidation present in the irradiated control samples or introduced during sample handling using the following equation:

Equation 36

$$\%modified_{i_{Norm}} = (\%modified_{FPOP} - \%modified_{Control})$$

Samples were labelled in triplicate, and final %modified measurements expressed as a mean value for these normalised replicate measurements;

Equation 37

$$\%modified_{Norm} = \frac{1}{n}\sum_{i}^{n=3} \%modified_{i_{Norm}}$$

The uncertainty of these measurements was expressed as a combined uncertainty using the following;

Equation 38

$$U = \sqrt{\frac{{SD_{FPOP}}^2 + {SD_{Control}}^2}{2}}$$

To determine statistically significant changes two tailed t-tests were applied, p values < 0.05 were determined to be statistically significant.

4.2.7.2 Peptide oxidation identification

Peptides were identified from Q-ToF DDA and Orbitrap dd-MS^2 experiments using PEAKS 8.5 software [Bioinformatics Solutions Inc., Waterloo, ON, Canada], searching against a database containing rhGH, catalase and trypsin sequences including a reverse database, a precursor mass error of ± 10 ppm and a fragment error of 0.1 Da. Similarly peptides were identified from Q-ToF MS^E experiments using PLGS v 3.02, using the processing parameters listed in Table A-24.

Mass additions of + 15.99, + 31.98 and + 13.98 Da were searched to identify modified peptides [191]. Quantification of FPOP modifications was achieved by integration of modified and unmodified peaks generated from XICs using Equation 35 and Equation 36. Modified peaks with RTs indistinguishable from that of the unmodified peaks were excluded from quantification due to the inability to distinguish their presence as

a result of in-source oxidation of the unmodified peptide and not as a direct result of FPOP modifications.

XICs were generated by extracting the *m/z* of the base peak of each peptide isotope distribution, using Xcalibur software [v4.0.27.19; ThermoFisher, Bremen, Germany] and QuanLynx software [MassLynx v4.1; Waters, Milford, USA] for Orbitrap and Q-ToF data respectively.

4.3 Results and discussion

Both the capillary flow and single-shot Eppendorf labelling strategies and MS workflows have been individually optimised using rhGH and utilised for the analysis of rhGH model system. The repeatability and sensitivity of each strategy were then critically assessed and results compared. FPOP-MS measurements were then used to provide structural insights into changes in protein HOS of the rhGH model system and compared with insights gained from HDX-MS and IMS-MS based measurements from this thesis.

4.3.1 Capillary flow FPOP labelling

Oxidative conditions using the capillary flow set up were first optimised, as was a MS-based workflow, using a high resolution Q-ToF instrument, for the peptide level analysis of FPOP labelled samples. This workflow was then applied to identify and characterise peptide oxidation sites, before finally being applied to identify structural differences and provide structural insights into the rhGH model system.

4.3.1.1 Optimisation of oxidative conditions

Using the capillary flow FPOP set-up, samples of rhGH were labelled in the presence of two different scavenging agents, 17.5 mM glutamine and 350 µM histamine, at three different H_20_2 % concentrations; 0.02, 0.05 and 0.1%. Intact LC-MS analysis was used to confirm the presence of any oxidative species. Figure 73 b-d and Figure 74 b-d, display the rhGH base peak spectra of irradiated rhGH under increasing concentrations of H_20_2, peaks are labelled according to the number of oxygen molecules incorporated (~ +15.99 Da). At both scavenger conditions, increased oxidation was observed with increased levels of H_20_2. Glutamine and histidine scavenger solutions, in the presence of 0.1% H_20_2, resulted in maximum oxygen incorporation of 2 and 5 oxygen molecules respectively. The increased oxidation rates

observed for samples incubated with the histidine scavenger, is consistent with the higher reaction rate constant reported for hydroxyl radicals in the presence of glutamine compared to the histidine (~ 5.6 times calculated using reaction rates constants as measured by ~ Xu and Chance[191]), facilitating lower oxidation to the protein sample.

A Poisson distribution was observed for oxidation in all irradiated samples, indicating the absence of unwanted secondary oxidation reaction, implying oxidation levels observed should be indicative of the fully folded protein state of rhGH [196]. This reiterates that hydroxyl lifetime in the presence of histidine and glutamine scavengers is approximately ~ 1μs and faster than that of protein unfolding. It is beneficial to maximise the amount of oxidative labelling that occurs during FPOP labelling, whilst avoiding these secondary oxidation events, so as to increase the sensitivity of LC-MS/MS analysis and maximise the number of amino acid residues that can act as potential reporter molecules towards structural changes. In light of this, the optimised FPOP labelling conditions were selected as 350 μM histidine, 0.1% H_20_2 and utilised for all further samples labelled using capillary flow FPOP experiments .

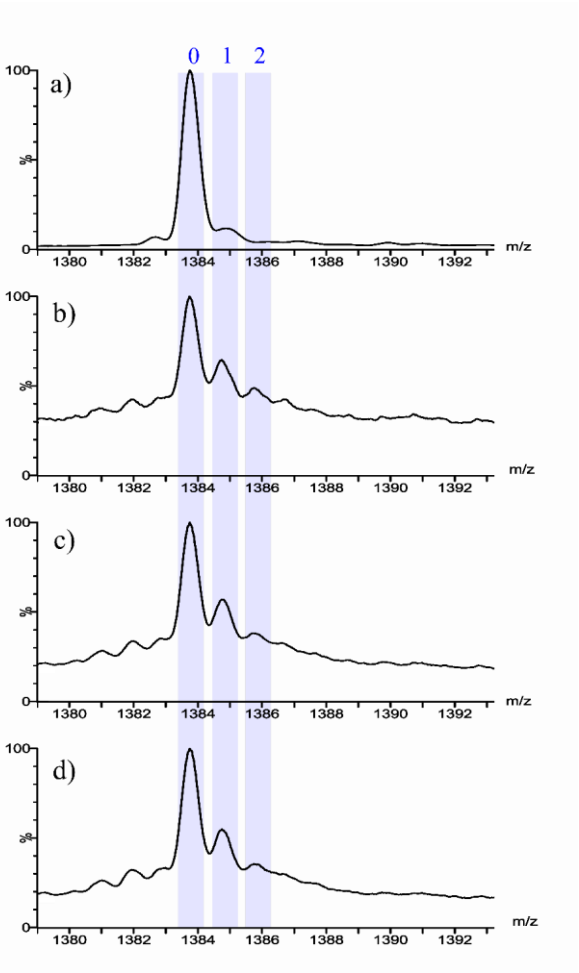


Figure 73. Intact protein LC-MS spectra of the rhGH +13 base peak for FPOP samples in presence of 17.5 mM glutamine scavenger. Spectra represent an a) irradiated control sample and irradiated rhGH in varying concentrations of H_20_2, b) 0.02% c) 0.05% and d) 0.1%. Peaks corresponding to unlabelled protein sample are labelled 0 and oxidated peaks are labelled according to the number of oxygen molecules incorporated (1,2).

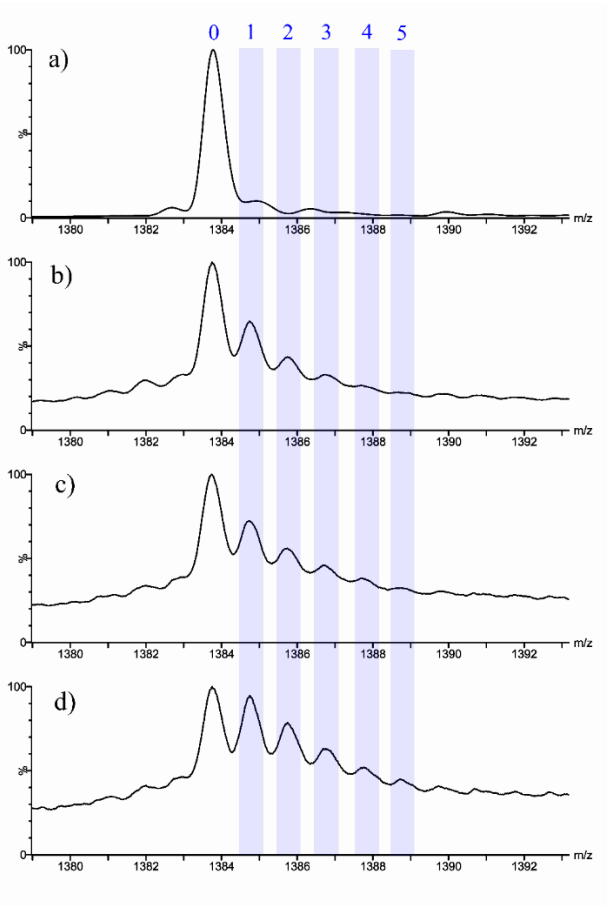


Figure 74. Intact protein LC-MS spectra of the rhGH +13 base peak from FPOP samples in presence of 350 µM histidine scavenger. Spectra represent an a) irradiated control sample and irradiated rhGH in varying concentrations of H_20_2, b) 0.02% c) 0.05% and d) 0.1%. Peaks corresponding to unlabelled protein sample are labelled 0 and oxidated peaks are labelled according to the number of oxygen molecules incorporated (1- 5).

4.3.1.2 Optimisation of peptide oxidation analysis

In order to obtain localised insights to FPOP-MS measurements, analysis of the labelled samples must occur at the peptide level, via a proteolysis step. Before application to the rhGH model system, oxidised rhGH samples were subjected to tryptic digestion and LC-MS/MS analysis, to identify and localise oxidation sites.

4.3.1.2.1 LC-MS/MS method development [Xevo Q-ToF]

In order to optimise the LC-MS/MS method for the bottom-up FPOP quantification of the rhGH and rhGH:Zn samples, a range of LC-MS/MS conditions was screened (Table 18) in which flow rate, desolvation temperature, elution gradient and scan speed were varied. Triplicate injections were made of tryptic digests of a capillary

FPOP labelled rhGH samples. The repeatability was assessed in terms of the RSDs of the extracted oxidative peak area (Figure 75) and RT (Figure 76) for four oxidated tryptic peptides, T2*, T4*, T11-1* and T11-2* (Table 19).

So as to best cover the analytical space across the range of peptides in terms of repeatability, conditions ii, Table 18, were selected for further bottom up analysis of capillary flow FPOP labelled samples.

Table 18. Range of LC-MS conditions used to screen for optimised method for bottom up analysis of FPOP samples

Condition	**LC flow rate (µL/min)**	**Gradient**	**Desolvation Temp (°C)**	**Scan speed**
i	300	3 - 50%	250	0.6 sec
ii	400	3 - 50%	250	0.7 sec
iii	400	3 - 45%	250	0.7 sec
iv	400	3 - 50%	300	0.7 sec

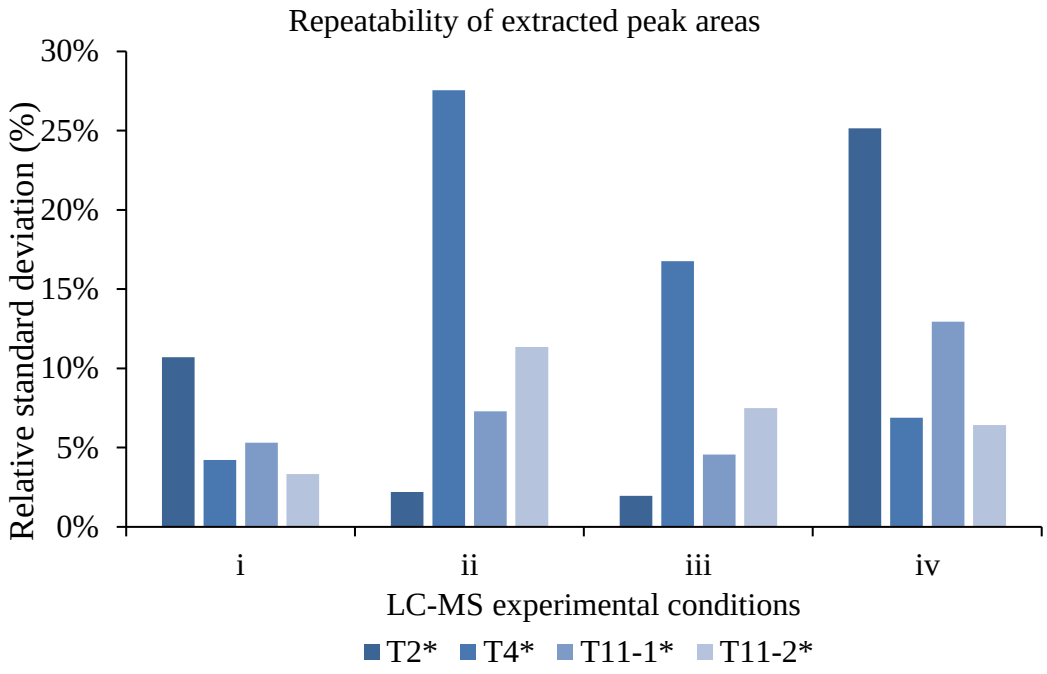


Figure 75. XIC repeatability of four oxidised rhGH tryptic peptides using a range of experimental LC-MS conditions (i-iv, Table 18).

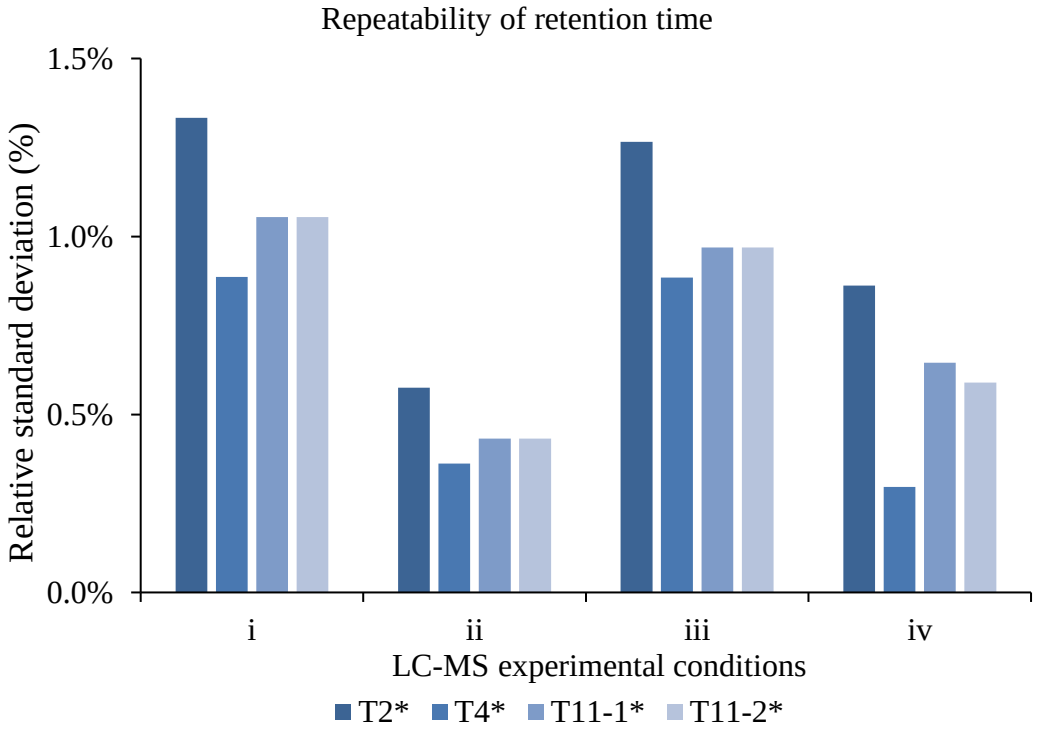


Figure 76. RT repeatability of four oxidised rhGH tryptic peptides using a range of experimental LC-MS conditions (i-iv, Table 18).

4.3.1.2.2 Identification and residue localisation of peptide oxidation

To identify and localise the oxidation sites, tryptic digests of triplicate FPOP-labelled rhGH samples were analysed using the optimised LC-MS/MS Xevo Q-ToF method. A combined approach of MS^E and DDA analysis was used for peptide identification of the tryptic digests and sequence coverage rates were in the range of ~ 95.3% and ~ 83% for the two MS/MS methodologies, respectively.

In total, ten oxidatively modified tryptic peptides were identified reproducibly across all the samples, as summarised in Table 19. A list of unmodified peptides produced from an *in-silico* digest of rhGH is found in the appendices, Table A-27. MS/MS spectra were manually examined to localise oxidation to amino acid resolution, and in total five different amino acid residues were oxidised, correlating with a mixture of oxidatively labelled populations in the range of 1-5 oxidation sites revealed by the global analysis of FPOP labelled rhGH (Figure 74 d). These five oxidated residues account for ~ 18% of the side chains within the sequence that are the most reactive to hydroxyl radical species (Cys, Trp, Tyr, Met, Phe and His) [191].

Two different oxidative species corresponding to the formation of a carbonyl moiety (~ +13.98 Da) and addition of an oxygen atom (~ +15.99 Da) were observed for the T1* peptide. Analysis of MS/MS spectra resolved these modifications to the first two N- terminal amino acid residues, Phe-1 and Pro-2 to give the oxygen and carbonyl products respectively. The coexistence of multiple forms of oxidative species for the T1 peptide is testament to the presence of multiple oxidisable residues in the peptide sequence (such as Phe and Pro) giving rise to singly modified peptide species which have been modified at different residues.

There is evidence of several isobaric peaks corresponding to the XIC spectra of +15.99 Da modifications of several peptides, T1, T4 and T11 across the rhGH sequence. Several modified peaks exist for the T1 peptide, as illustrated in Figure 77, by the XIC for the 473.77 *m/z* species, correlating with a +15.99 Da modification of the T1 peptide. Closer inspection of the mass spectra underneath the chromatographic peaks T*-1, T*-2, T*-3 (Figure 77 a) revealed three isobaric peaks (RT 9.37, 9.74, 10.06) that fulfil the correct isotope pattern and charge state for a z =2 ion, yet result in different RTs. MS/MS spectra revealed identical fragment spectra, consistent with +15.99 Da modification of the Phe-1 residue. These data suggest the presence of structural isomers of the oxidated T1*(+15.99) species, occurring due to hydroxyl

radical oxidation at the meta-, ortho- and para-positions of the benzene ring within the Phe-residue (Scheme 1) [110, 191, 202, 295]. It has been documented that for RP-UPLC, the elution order of oxidised phenylalanine, both as free amino acids and as residues within a peptide chain, occurs in the order of para < meta < ortho, therefore these assignments have been made based on elution time (Table 19) [102, 295].

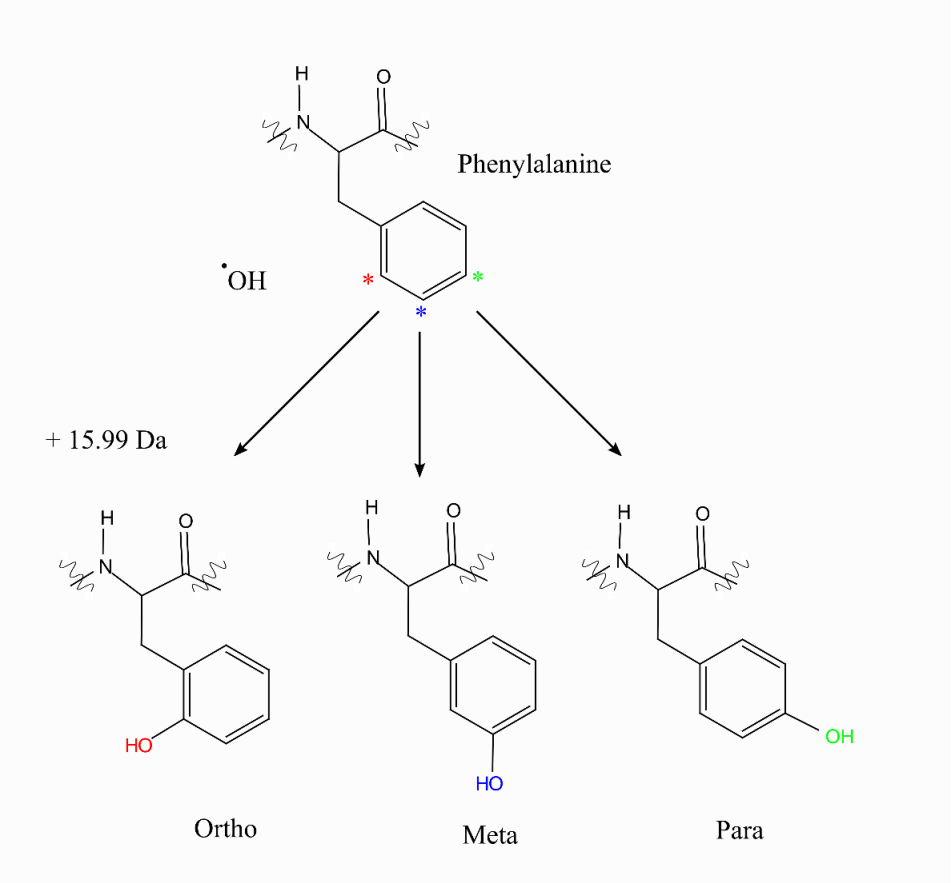
Similarly three structural isomers were observed for the oxidation of the Phe-25 residue in T4*-[1-3], as identified by identical MS/MS spectra for each corresponding peak, with para-, meta- and ortho- positions assigned based on elution order (Figure 78).

Finally inspection of the MS/MS spectra of the two isobaric T11* peaks, showed oxidation of the Met-125 residue for both sets of spectra (Figure 79 a). Unlike aromatic side chains, methionine residues lack a ring structure and hence the possibility of para-, meta- and ortho- oxidative products is removed. Hydroxyl mediated oxidation of methionine side chain occurs directly at the sulphur atom [191, 295]. The presence of multiple oxidated forms of Met-125 could suggest that there is some structural flexibility along the hydrocarbon chain which could allow multiple conformations and hence ultimately multiple hydroxyl radical oxidative targets for this methionine group. Oxidation in this instance could indeed suggest not only surface accessibility at this residue but also the capacity for conformational flexibility.

The existence of these structural isomers of oxidative products could offer the possibility of sub-residue interpretation of oxidative data [110].

Table 19. Oxidatively modified tryptic peptides (*) identified in bottom-up analysis of capillary flow FPOP labelled rhGH, amino acid oxidation site is indicated as identified by MS/MS fragment data. Due to isobaric fragment spectra, site specific oxidation at the meta-, ortho- and para-positions of the benzene ring within the Phe-residue have been assigned based on relative elution order [102, 295].

Peptide ID	**Residues**	**Sequence**	**Modification mass (Da)**	**Oxidation site**
T1*	1-8	FPTIPLSR	+ 13.98	Pro-2
T1*-1	1-8	FPTIPLSR	+ 15.99	Phe-1-*para*
T1*-2	1-8	FPTIPLSR	+ 15.99	Phe-1-*ortho*
T1*-3	1-8	FPTIPLSR	+ 15.99	Phe-1-*meta*
T2*	9-16	LFDNAMLR	+ 15.99	Met-14
T4*-1	20-38	LHQLAFDTYQEFEE AYIPK	+ 15.99	Phe-25-*para*
T4*-2	20-38	LHQLAFDTYQEFEE AYIPK	+ 15.99	Phe-25-*ortho*
T4*-3	20-38	LHQLAFDTYQEFEE AYIPK	+ 15.99	Phe-25-*meta*
T11*-1	116-127	DLEEGIQTLMGR	+ 15.99	Met-125
T11*-2	116-127	DLEEGIQTLMGR	+ 15.99	Met-125



Scheme 1. Structural isomers arising from the hydroxyl-radical mediated oxidation of phenylalanine.

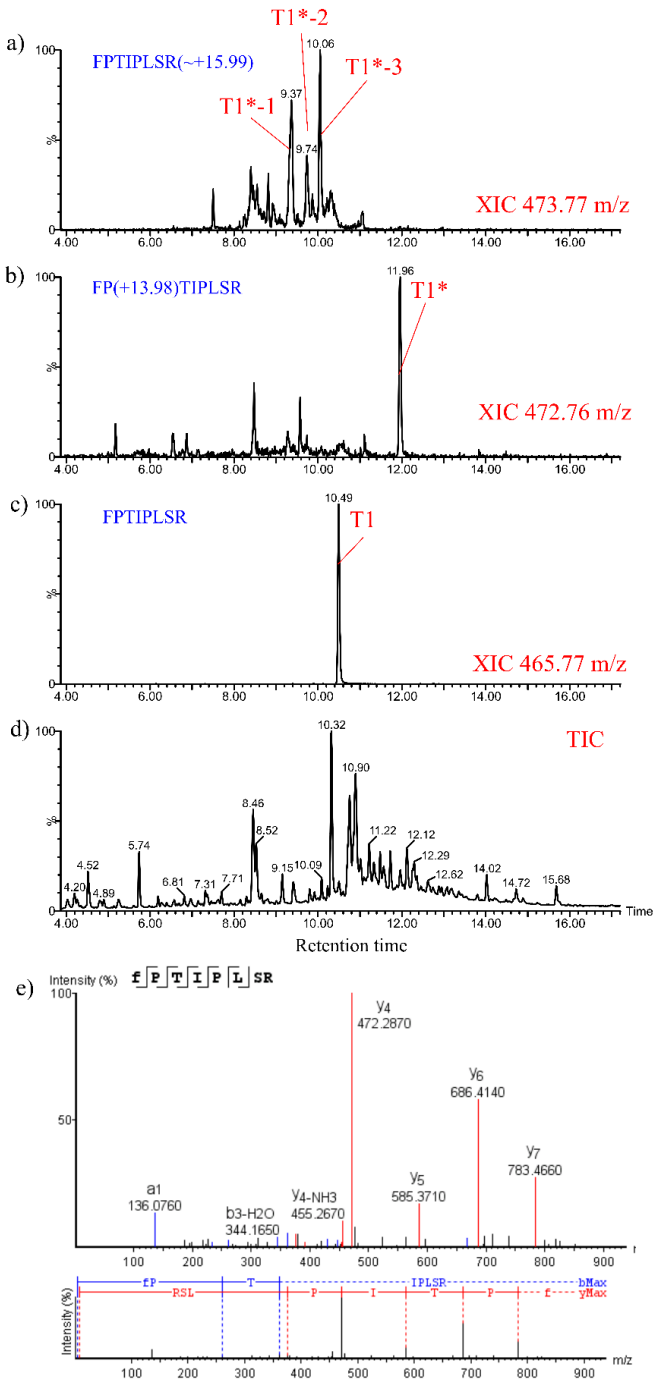


Figure 77. FPOP modification of the T1 rhGH peptide. XIC of the multiple +15.99 Da modified forms (a), the single +13.98 Da modified form (b) and the unmodified form (c) of the peptide are illustrated, d) analysis of MS/MS spectra of the T1*-1, T1*-2 and T1*-3 were identical and localised the modification to the Phe-1 residue for all three, modification was thought to have occurred at para-, meta- and ortho- positions of the peaks, respectively, based on RP-UPLC elution times.

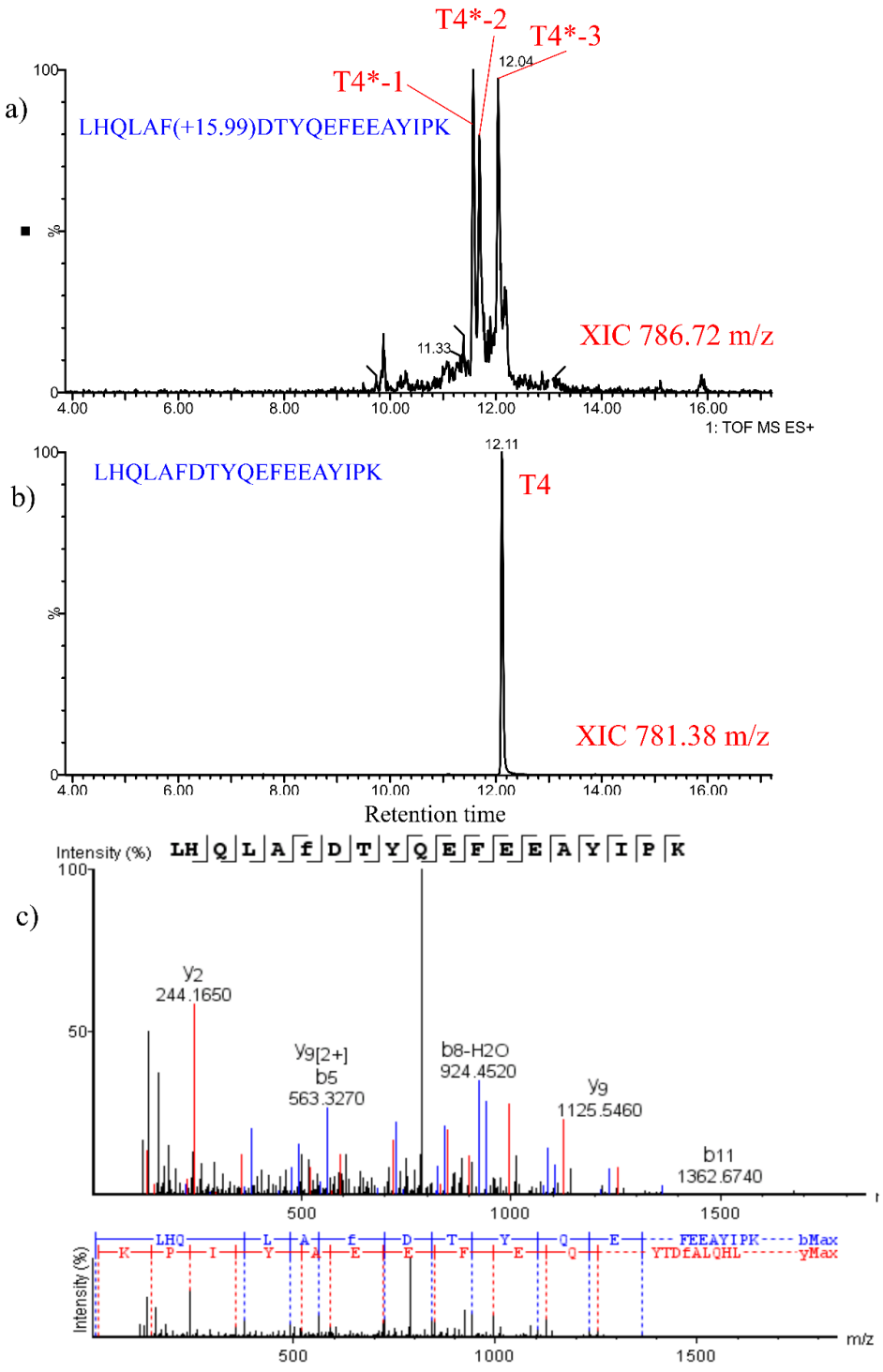


Figure 78. FPOP modification of the T4 rhGH peptide. XIC of the multiple +15.99 Da modified forms (a) and the unmodified form (b) of the peptide are illustrated, c) analysis of MS/MS spectra of the T4*-1, T4*-2 and T4*-3 were identical and localised the modification to the Phe-25 residue for all three, modification was thought to have occurred at para-, meta- and ortho- positions of the peaks respectively, based on RP-UPLC elution times.

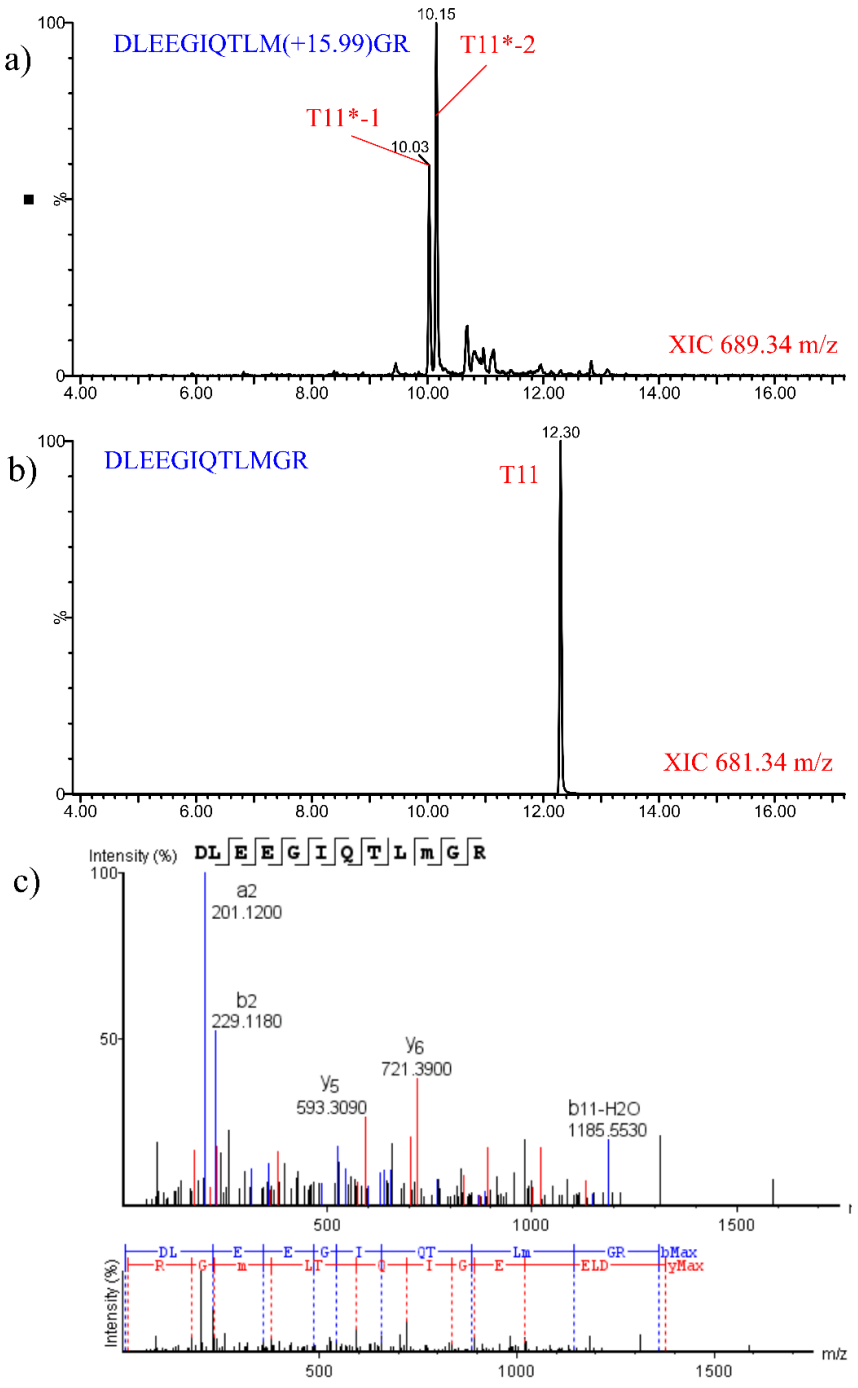


Figure 79. FPOP modification of the T11 rhGH peptide. XIC of the multiple +15.99 Da modified forms (a) and the unmodified form (b) of the peptide are illustrated, c) analyses of MS/MS spectra of the T11*-1 and T11*-2 were identical and localised the modification to the Met-125 residue for both, illustrating two structural isomers of the modified peptide.

4.3.1.3 FPOP-MS measurements to identify HOS changes induced by zinc

The optimised oxidative conditions and LC-MS/MS workflows were then used to probe the global conformational changes of the rhGH model system.

4.3.1.3.1 Quantification of global oxidation

Samples of rhGH and rhGH:Zn at three different ratios, 1:2, 1:10 and 1:20, were oxidatively labelled in triplicate using the optimised FPOP labelling conditions for the capillary flow set up. Intact protein LC-MS analysis was performed to probe for differences in global oxidation levels. Deconvolution of the resulting mass spectra revealed up to five oxidation states (~ +16, +32…. Da) across all samples, and the $\%modified_{NORM}$ for each oxidative peak combined to give a total global oxidation measurement for each sample. Figure 80, illustrates the measured average total global modification levels, oxidation levels were in the range of 58-57% and indiscriminate for rhGH, rhGH:Zn 1:2 and 1:10 samples, indicating a consistent globular structure across these samples. Total global oxidation however was lower for rhGH:Zn 1:20 samples, in the range of ~ 47%, suggesting the global structure of the rhGH-Zn complex to be a less solvent accessible area, and an overall change in global conformation relative to the rhGH. A shift towards global compaction was alluded to by IMS-MS measurements in Chapter 2.

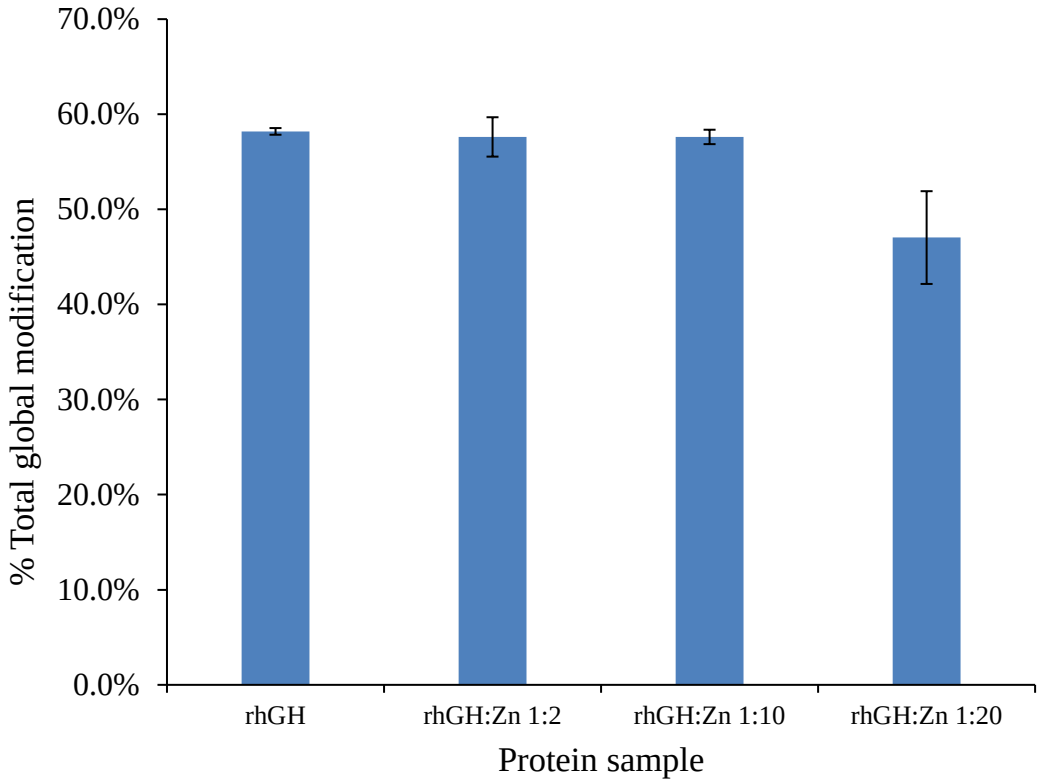


Figure 80. Total global modification (%) as measured from the deconvoluted mass spectra of FPOP labelled samples. Measurements are expressed as average values and error bars represent the standard deviation (n=3).

4.3.1.3.2 Quantification of peptide oxidation

To probe for global conformational changes between rhGH and rhGH:Zn 1:20 as indicated by intact protein analysis of capillary flow FPOP labelled samples, samples were subjected to tryptic digestion followed by bottom-up LC-MS/MS analysis and relative quantification. FPOP quantification was performed for capillary flow FPOP labelled rhGH and rhGH:Zn 1:20 samples, expressed as %modification$_{NORM}$, and averaged across triplicate labelled samples.

Analysis of a triplicate injection of a single tryptic digest of a rhGH FPOP labelled sample was shown to be repeatable with standard deviations in the range of 0.1-0.9% across all ten oxidated peptide species (Table 19). Measured peptide modification values for each rhGH sample type are displayed in Figure 81 a. Those residues (or sub-level residues) that demonstrated significant changes ($p < 0.05$) or indeed similarity in %modification as a result of the presence of zinc are summarised in Table 20.

Comparison of oxidation measurements across the two FPOP sample types revealed differences in oxidation measurements for six out of the ten oxidated peptide species, corresponding to T1*, T1*-3, T4*-1, T4*-3 T11*-1 and T11*-2 peptides, and specifically residues Phe-1-*meta* Pro-2, Phe-25-*para* and Phe-25-*ortho* and Met-125 (Table 20). All six of these peptides demonstrated higher oxidation levels for the rhGH:Zn 1:20 sample relative to the control, suggesting a trend of increased surface accessibility at these residues upon zinc binding and a change in protein HOS for rhGH due to the presence of zinc at this ratio. These data are consistent with the general trend of increased oxidation rates observed in HDX-MS measurements in the presence of zinc, Table 9.

The attributed positioning of the structural isomers for T1*-[1-3] and T4*-[1-3] + 15.99 Da species as a result of para-, meta- and ortho- ring oxidative addition are highlighted in Figure 82. For the T1* +15.99 species, the oxidation modification is increased only for Phe-1-*ortho* site in the zinc containing sample, as the case for the Pro-2 residue. This suggests that this region between the two side chain sites becomes more exposed due to the presence of zinc. The unaffected modification extent of the Phe-1 meta- and para- sites however suggested that the positioning of the rest of the side chain of Phe-1, appears to be consistent in absence and presence of zinc. This combined with the relatively small change in %modification for these oxidation sites, is consistent with a small shift in structure and not a large structural rearrangement.

Similarly oxidation rates increase for the Phe-25 side chain para- and ortho-sites (Figure 81 a and Figure 82 b), located in helix A of rhGH structure, signifying a more solvent exposed residue due to the presence of zinc. This residue is also localised near the residues involved in zinc binding of His-18, His-21 and Glu-174, though no oxidation of these residues were identified. In addition, increased oxidation rates were observed for both of the structural isomers of the oxidised Met-125 residue, located in helix C of rhGH structure (Figure 81 b). Modification levels for the Met-125 residue appear to be the most affected due to the presence of zinc, a site not located directly in the zinc binding sites, suggesting a long range structural change as a result of the presence and binding of zinc.

Of those modified residues identified, the only residue in which oxidation levels were unchanged due to the presence of zinc, was the Met-14 residue, demonstrating structural equivalence across the samples at this residue. Overall comparison of

capillary flow FPOP oxidation between rhGH and rhGH:Zn 1:20 suggest small and localised changes in structure due to the presence of zinc occurring at the N-terminus and helices A and C.

Table 20. Observed oxidations of rhGH labelled by capillary flow FPOP-MS experiments, sub-residue level information is listed where identified. Increases in oxidation levels for rhGH:Zn FPOP labelled samples relative to rhGH are indicated by (+), a blank input represents no relative change.

Amino Acid			Relative Oxidation Change
Type	Residue Number	Sub level	rhGH:Zn 1:20
Phe	1	- *para*	
Phe	1	- *meta*	
Phe	1	- *ortho*	+
Pro	2		+
Met	14		
Phe	25	- *para*	+
Phe	25	- *meta*	
Phe	25	- *ortho*	+
Met	125	*isomer 1*	+
Met	125	*isomer 2*	+

The total extent of modification was generally low across all peptides and sample type, ~ > 25%, this combined with the low number of oxidised residues suggests both forms of rhGH have a protein HOS which results in a relatively low SASA, consistent with the high levels of structure observed by the HDX-MS analysis (Chapter 2, section 2.3.1.3).

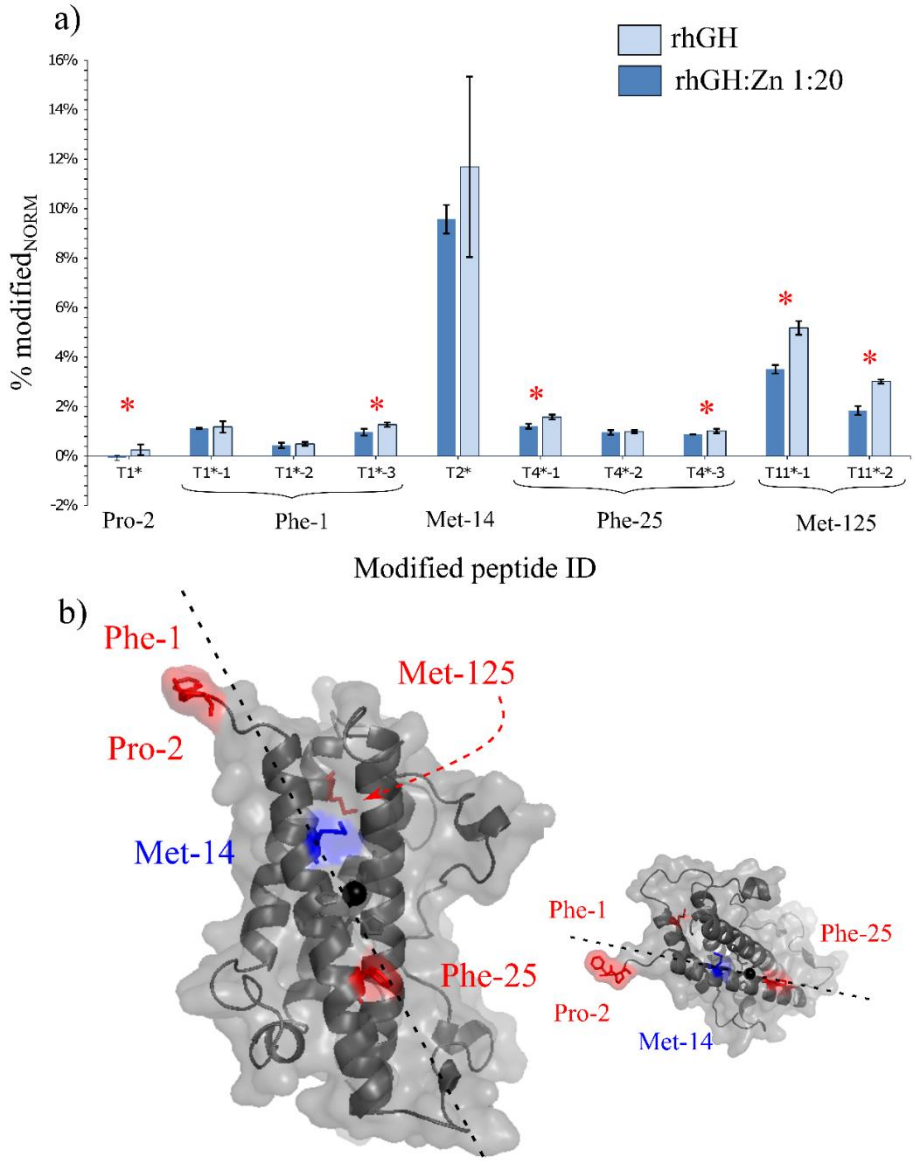


Figure 81 a) Peptide level quantification of FPOP-MS measurements of rhGH (light blue) and rhGH:Zn 1:20 (dark blue), as labelled using capillary flow FPOP set up. Error bars represent the calculated combined uncertainty. Peptides demonstrating a significant change in %modification between the rhGH and Zn bound rhGH are indicated using * . b) The XRC structure of rhGH:Zn complex, in two orientations rotated around the line (----), with the five modified side chain residues identified highlighted, residues demonstrating increased %modification due to the presence of zinc are highlighted in red (Phe-1, Pro-2, Phe-25 and Met-125) with Met14 indicated in blue where %modification levels were consistent across samples. PDB: 1BP3 [208].

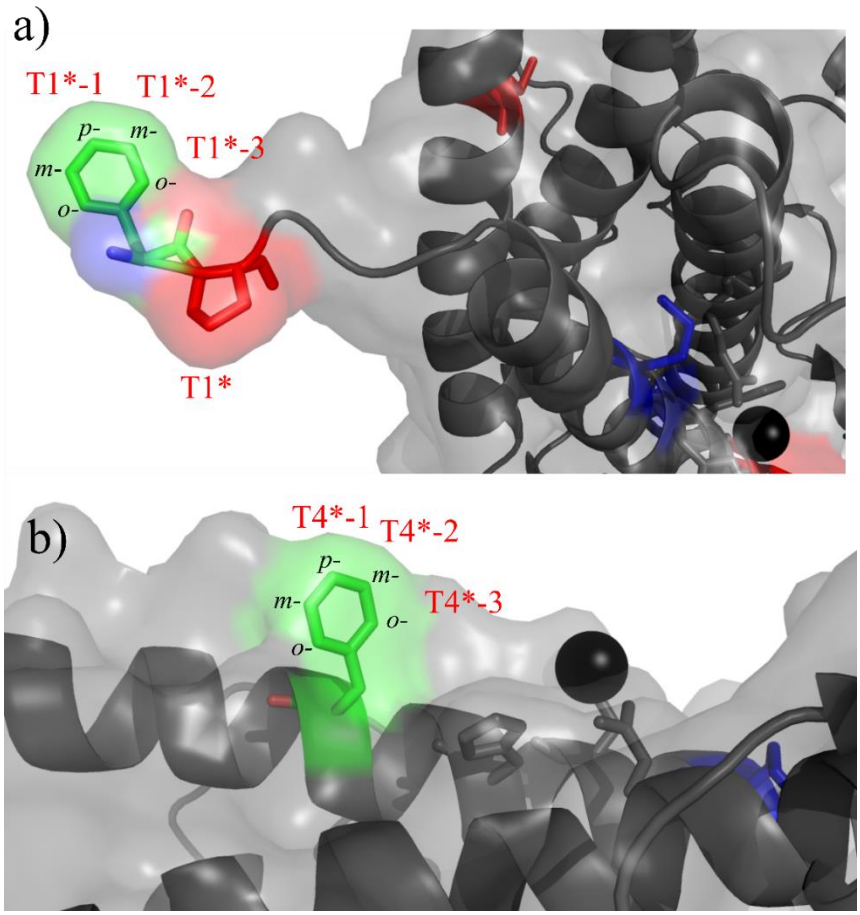


Figure 82. Representation of the para-, meta- and ortho- positions (represented as *p-*, *m- and o-*) of residues a) Phe-1 and b) Phe-25 (highlighted in green), identified as giving rise to the three isobaric oxidative species observed for the T1*-[1-3] and T4*-[1-3] modified peptides. Increased oxidation rates were observed for T1*-3, T4*-1 and T4*-3 sites due to the presence of zinc. PDB: 1BP3[208].

4.3.2 Single-shot Eppendorf FPOP labelling

During the capillary flow FPOP labelling for the rhGH samples, blockages of the capillary during flow of rhGH samples often occurred, potentially due to aggregation of the protein samples in the flow system. In these instances, re-preparation and re-labelling of the samples were required as a result of sample loss and to ensure reproducibility of measurements. It was hoped that the development and optimisation of a single-shot Eppendorf approach, such as the one described here (Figure 72), would circumvent the need for a capillary, simplifying and increasing the reproducibility of the FPOP labelling procedure for rhGH.

To explore this as a potential FPOP labelling methodology, oxidation conditions were first optimised for the single shot Eppendorf strategy, as were MS-based workflows using an alternative high resolution Orbitrap instrument to that used in section 4.3.1.2.1. Peptide level analysis was then used to identify and localise modifications, before finally applying the optimised workflow to gain structural insights of the rhGH model system. Final measurements were compared then with capillary flow measurements, IMS-MS and HDX-MS measurements.

4.3.2.1 Optimisation of oxidative conditions

It was anticipated that the laser penetration into the sample would likely be minimal leading to irradiation occurring only to those proteins close to the surface, (Figure 72 b). This would result in an unmodified portion of proteins such as occurs with the exclusion volume in traditional capillary flow FPOP experiments, typically used to avoid unwarranted secondary oxidation events. Yet, it was hypothesised that due to the tapering shape of the Eppendorf tube and resulting bias of sample volume towards the irradiation source, this unmodified portion would not be excessive. However, so as to ensure sufficient oxidation for detection, the overall sample volume (protein + scavenger + H_20_2) was reduced to 10 µL in order to compensate for anticipated reduction of laser exposure and in light of this protein sample concentration was also increased ~ 28 µM.

As for the capillary flow FPOP set up, scavenger concentration and H_20_2 were optimised to ensure sufficient labelling. Figure 83 d demonstrates the optimised FPOP conditions of rhGH irradiated in the presence of 175 µM histidine scavenger and 0.025% H_20_2, where up to four global modifications are observed. Oxidation appears slightly less in the single-shot Eppendorf FPOP experiment compared to the capillary (four vs five global modifications respectively). This could be as a result of an increased exclusion volume due to laser penetration, as discussed previously, or as a result of the increased relevant concentration of rhGH decreasing the effective ratio of rhGH molecules to hydroxyl radicals.

The single-shot Eppendorf labelling procedure however was much faster (~ 30 sec in total including quenching steps) and more straight forward compared to that of the capillary flow (~ 3-4 min in total including quenching steps) with no issues of capillary blockages.

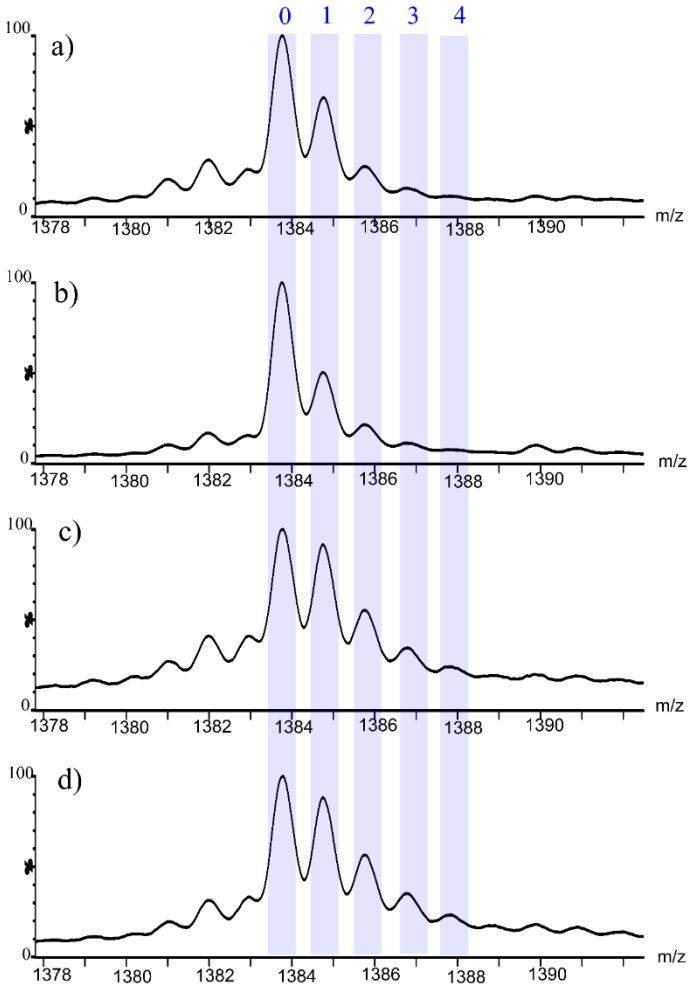


Figure 83. Intact protein LC-MS spectra of the rhGH +13 base peak of FPOP samples in presence of 0.25% H_20_2. Spectra represent an irradiated rhGH in varying histidine scavenger concentrations of a) 100 mM, b) 10 mM, c) 2.5 mM and d) 175 µM . Peaks corresponding to unlabelled protein sample are labelled 0 and oxidated peaks are labelled according to the number of oxygen molecules incorporated (1-4).

4.3.2.2 Optimisation of peptide oxidation analysis

4.3.2.2.1 LC-MS/MS method development [Q-Exactive plus]

In hopes of higher sensitivity and greater mass accuracy, bottom-up LC-MS/MS was moved to an Q-Exactive plus Orbitrap MS instrument. Chromatography flow rates were reduced to 30 µL/min and experimental conditions, such as maximum injection

time and dynamic exclusion, optimised to maximise sensitivity towards low abundant oxidative species [110].

Previous quantification research performed on this instrument within LGC, has indicated that in order to facilitate the accurate relative quantification of low abundant *m/z* species, it must be established that the responses of both the high and low abundant ion species are within the linear range of the MS detector. To do this, increasing injection volumes, 2, 4, 6, 8, 10 µL of a tryptic digest of FPOP labelled rhGH were made in duplicate and analysed by LC-MS/MS. XIC peak areas were abstracted for the three unmodified peptides T2, T4 and T11 at each injection volume. Figure 84 illustrates the linearity plots of injection volume vs signal intensity for the three unmodified peptides in their most abundant charge states. The T11 peptide appears linear across the whole range (R = 0.9969), however the T2 and T4 peptides appear to deviate in terms of linearity and reduced reproducibility at higher injection volumes, > 6 µL, due to increased signal relative to T11 peptide (~ 10^8 vs 10^7) . In light of this injection volumes of 4 µL, corresponding to 1.5 pmol of rhGH, were selected for further FPOP quantification experiments.

Finally, triplicate injections of a single rhGH FPOP labelled digest were made using the 4 µL optimised injection volume and analysis of oxidated peptides T2*, T4*-1, T11*-1 and T11*-2 performed. The repeatability of LC-MS/MS method was assessed in terms of $\%modified_{NORM}$ and RT (Table 21).

Table 21. Repeatability of LC-Orbitrap-MS/MS method, expressed as the relative standard deviation of $\%modified_{NORM}$ and RT measurements of series of oxidised peptides. Data produced from repeat injections of a single-shot Eppendorf FPOP labelled rhGH sample (n=3)

	Modified peptide ID			
	T2*	**T4*-1**	**T11*-1**	**T11*-2**
$\%modified_{NORM}$	2.1%	3.7%	5.3%	3.8%
RT	-0.4%	-0.1%	0.2%	0.0%

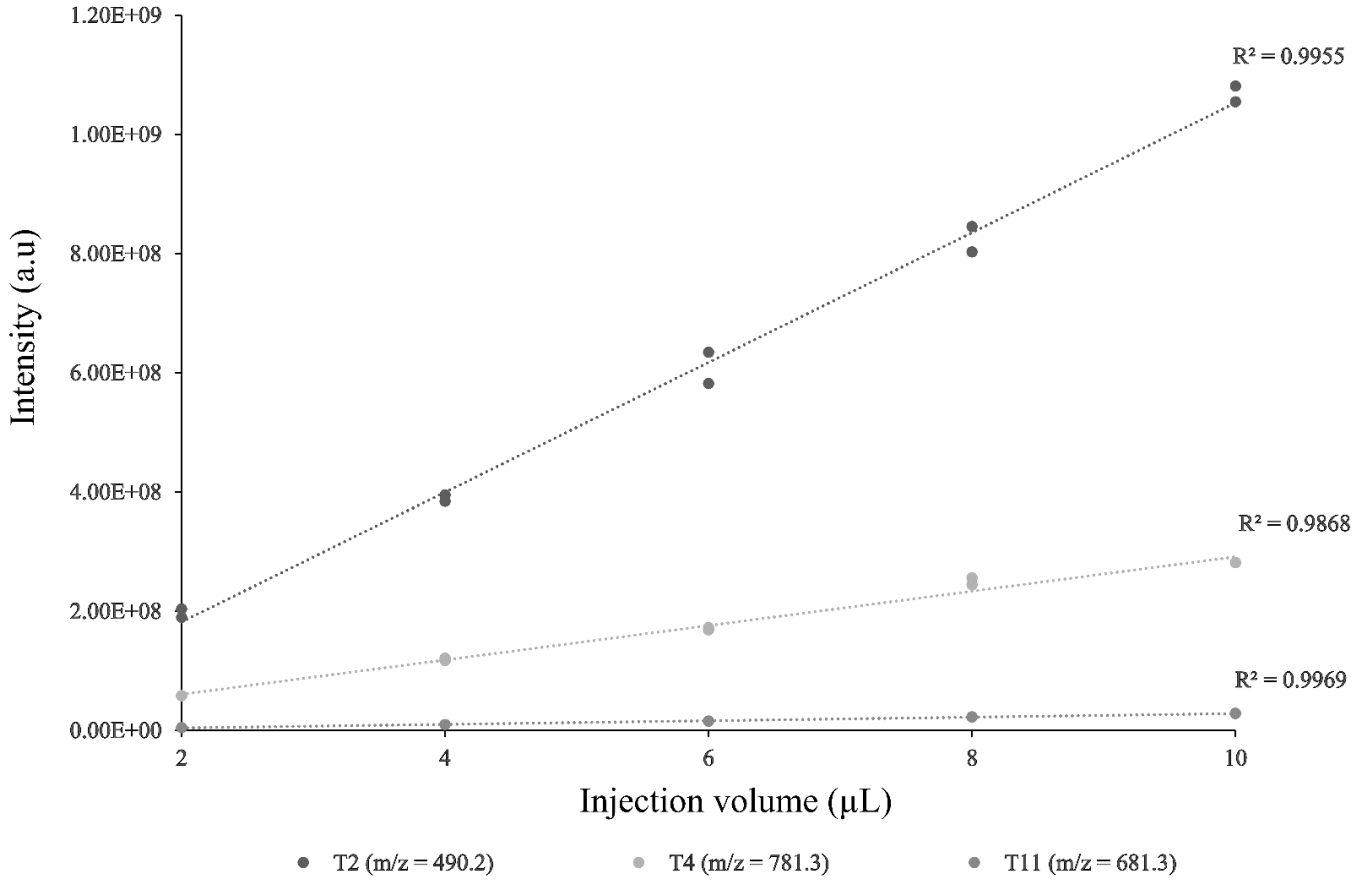


Figure 84. Linearity of the LC-Orbitrap MS/MS method. XIC peak areas of the T2, T4 and T11 peptides were extracted from a series of injection volumes (2- 10 µL) from a tryptic digest of rhGH FPOP labelled sample, injections were made in duplicate.

4.3.2.2.2 Identification and residue localisation of peptide oxidation

All the modified peptides and modification sites identified in capillary flow FPOP labelled digests, Table 20, were verified for single-shot Eppendorf labelled experiments. The lower flow rates used for the Orbitrap LC-MS/MS analysis ~ 30 µL/min, relative to 400 µL/min used in section 4.3.1, appeared to give gains in sensitivity, with increased sequence coverage across all rhGH tryptic digests in the range of 98-100%.

Two additional modification sites were also observed (Table 22), giving a total of twelve FPOP modified peptides and seven modified amino acid residues. Analysis of these modification species revealed a fourth additional isobaric +15.99 Da T4* modified peak, termed T4*-4. Analysis of MS/MS spectra localised the modification site to the His-21 side chain (Figure 85). The later elution of the histidine modified peptide (~ RT 41.95) relative to the unmodified T4 peptide (~ RT 40.25) is consistent with elution profiles observed previously for peptides with FPOP oxidated histidine [102].

Furthermore, a +31.99 Da oxidation was observed for the C-terminal T21 peptide, and the oxidation site was localised to the Cys-189 residue (Figure 86). The unmodified species T21 peptide was observed mainly in (+57.02), occurring at Cys189, as result of IAM alkylating agent used during tryptic digestion. Both this carbamidomethylated form of the unmodified T21, in addition to the XIC of non-alkylated T21 form, which was present in low abundance, was used for subsequent quantification of the T21*.

Table 22. Unique oxidatively modified (*) tryptic peptides identified in bottom-up analysis of single-shot Eppendorf FPOP labelled rhGH, amino acid oxidation site is indicated as identified by MS/MS fragment data.

Peptide ID	Residues	Sequence	Modification mass (Da)	Oxidation site
T4*-4	20-38	LHQLAFDTYQEFEEAYIPK	+ 15.99	His21
T21*	184-191	SVEGSCGF	+ 31.99	Cys189

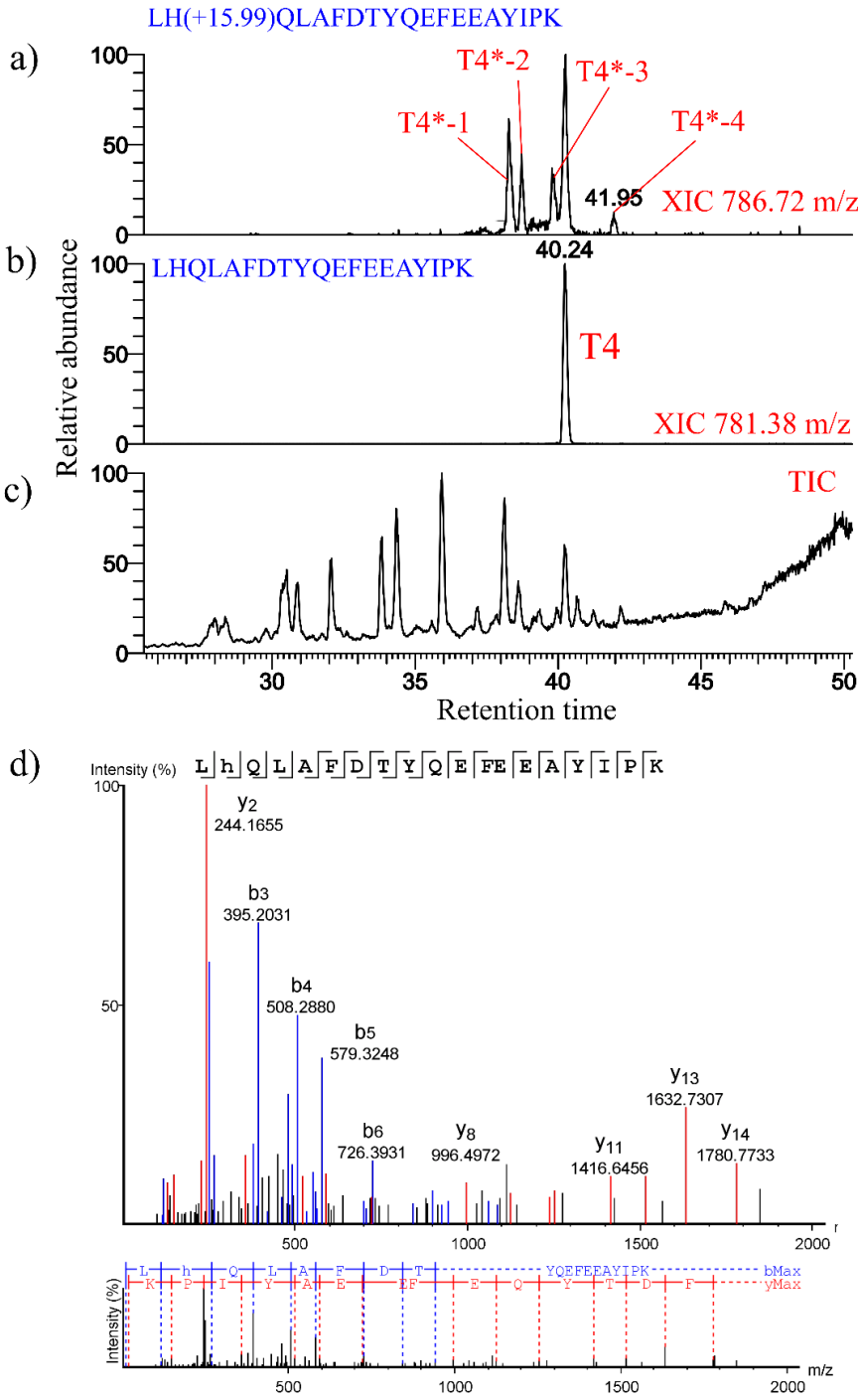


Figure 85. FPOP modification of the T4 rhGH peptide. XIC of the multiple +15.99 Da modified forms (a) and the unmodified form (b) of the peptide are illustrated, c) the TIC of bottom up analysis d) analysis of T4*-4 MS/MS spectra localised the modification to the H21 residue. T4*-[1-3] correspond to oxidation at Phe-25 residue occurring at para-, meta- and ortho- positions of the peaks respectively.

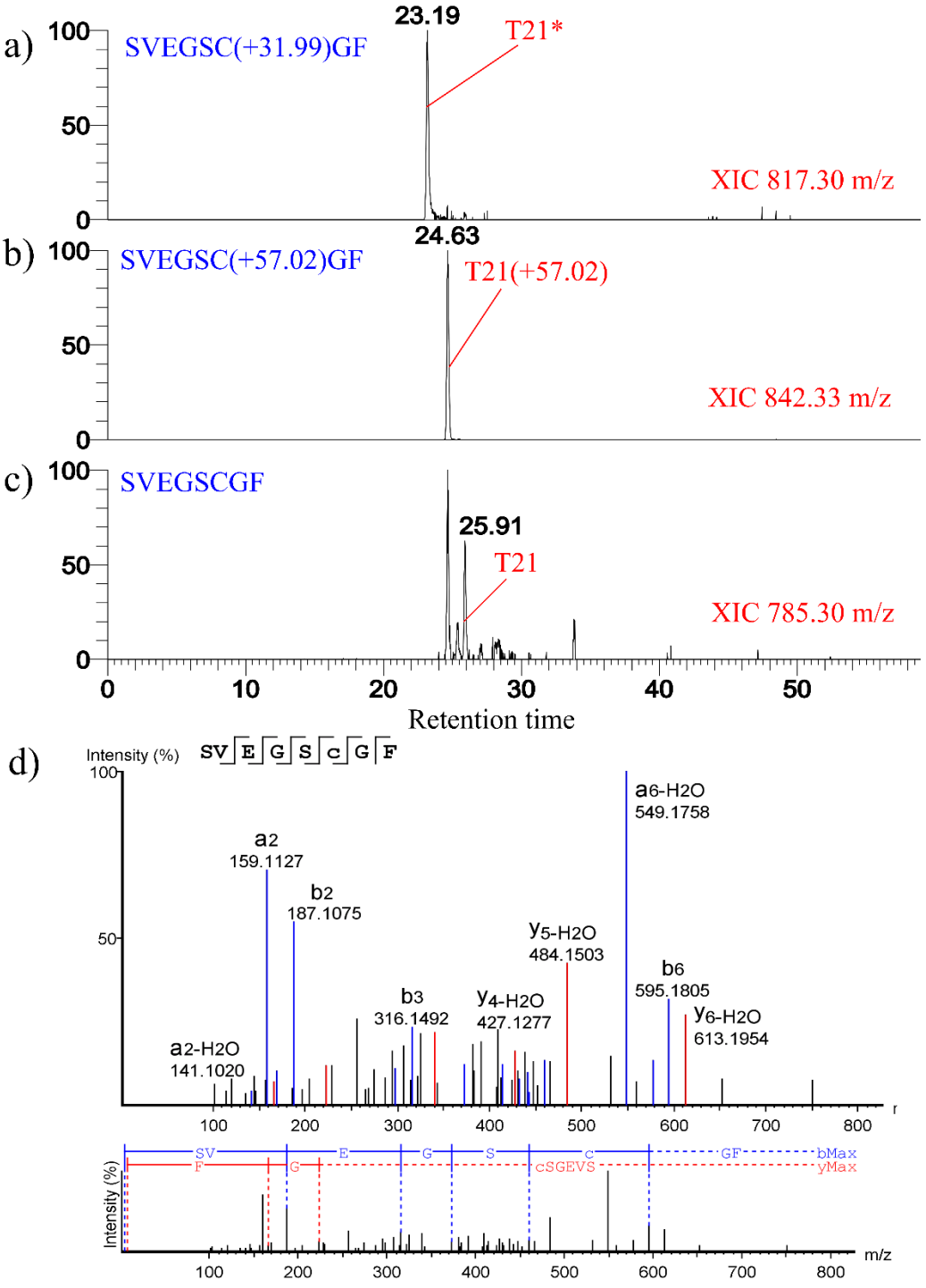


Figure 86. FPOP modification of the T21 rhGH peptide. XIC of the +31.99 Da modified forms (a) and the carbamidomethylated, +57.02 Da alkylated form of the unmodified form, and the unmodified form of the peptide are illustrated, d) analysis of T21* MS/MS spectra localised the modification to the Cys-189 residue.

4.3.2.3 FPOP-MS measurements to identify HOS changes induced by zinc

Samples of rhGH and rhGH:Zn at ratios of 1:2, 1:10 and 1:20, were labelled in triplicate using the single-shot Eppendorf FPOP labelling workflow, before being immediately frozen, defrosted and then digested using trypsin to allow for bottom-up analysis of FPOP %modification.

4.3.2.3.1 Quantification of peptide oxidation

Following the identification and localisation of the all twelve oxidated tryptic peptides and seven modified residues as listed in Table 19 and Table 22, relative FPOP quantification was performed in triplicate for the digests of rhGH and rhGH:Zn (1:2, 1:10 and 1:20) single shot Eppendorf FPOP labelled samples. Data treatment and normalisation against an irradiated control was performed as described previously.

Analysis of a triplicate injection of a single rhGH FPOP labelled sample revealed LC-MS/MS quantification was shown to be repeatable with standard deviations in the range of 0.1-0.9% across all twelve oxidated peptide species (% modification un-normalised against irradiated control).

Measured peptide modification values for each rhGH sample type are displayed in Figure 87. For each peptide, relative comparisons were made between %modification measurements of rhGH and each of the three rhGH:Zn ratios, to identify similarity or potential changes in protein HOS. Those residues (or sub-level residues) that demonstrated significant changes ($p < 0.05$) or indeed similarity in %modification as a result of the presence of zinc are summarised in Table 23.

All three rhGH:Zn FPOP labelled samples displayed increased oxidation rates for Pro-2 and Phe25-para and meta residue sites, indicating structural changes in the form of increased surface accessibility for the both N-terminal region and helix A respectively. These observations were consistent with measurements from capillary flow FPOP-MS experiments for rhGH:Zn 1:20.

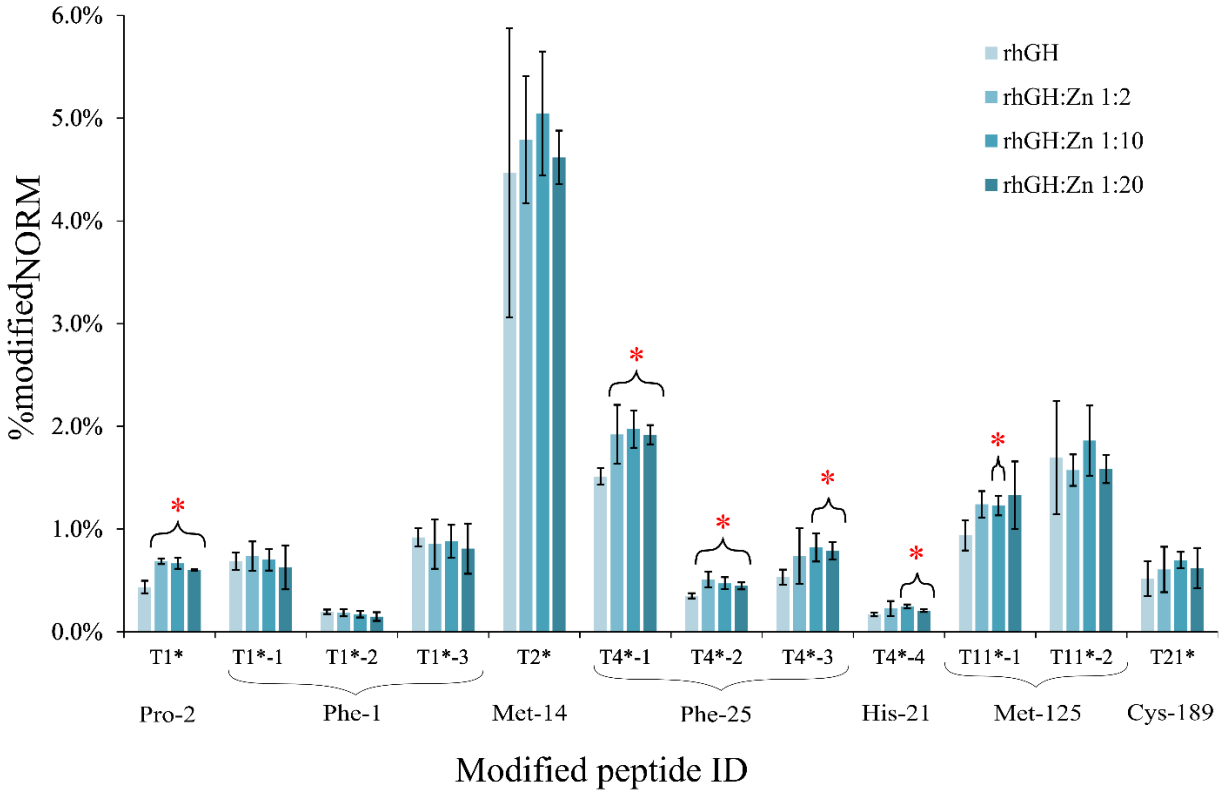


Figure 87. Peptide level quantification of FPOP-MS measurement of rhGH (light blue) and rhGH:Zn (1:2, 1:10 and 1:20) (increasingly darker blue), as labelled using single shot Eppendorf FPOP set up. Error bars represent combined uncertainty between triplicate irradiated and non-irradiated controls. Peptides demonstrating a significant change in %modification ($p < 0.05$) between the rhGH and Zn bound rhGH at each ratio are indicated using (*).

For samples incubated with rhGH:Zn ratios of 1:10 and 1:20, increased oxidation was also observed for His-21 and Phe-25-*ortho* side chains, both located in helix A, further promoting the idea of structural changes in this region, as indicated by capillary flow FPOP-MS experiments. It is of note, that the His-21 is indirectly involved in Zn^{2+} binding, via the stabilisation of the Glu-174 side chain which bonds directly to Zn^{2+} [208], as illustrated in Figure 88 b. Increased oxidation of this side chain in the presence of zinc is indicative of an increase in solvent accessibility due to the presence of the ligand, not reduction as may be expected for ligand binding areas. However, the stabilisation of the hydrogen bond between His-21 and Glu-174, appears to have increased the solvent exposure of the C2 site of the imidazole ring making it more susceptible towards hydroxyl radical attack. This would explain the increased modification at this site [295], Figure 88 b. Oxidation of this His-21 side chain was not observed/detected in capillary flow FPOP-MS experiments, demonstrating an increased sensitivity towards structural changes of this model system using this alternative workflow.

Table 23. Observed oxidations of rhGH labelled by single shot Eppendorf FPOP, sub-residue level information is listed where identified. Significant increases ($p < 0.05$) in oxidation levels for rhGH:Zn FPOP labelled samples relative to rhGH are indicated by (+), no significant change is indicated by blank input.

Amino Acid			**Relative Oxidation Change**		
Type	**Number**	**Sub level**	**rhGH:Zn 1:2**	**rhGH:Zn 1:10**	**rhGH:Zn 1:20**
Phe	1	- para			
Phe	1	- meta			
Phe	1	- ortho			
Pro	2		+	+	+
Met	14				
His	21			+	+
Phe	25	- para	+	+	+
Phe	25	- meta	+	+	+
Phe	25	- ortho		+	+
Met	125	isomer 1		+	
Met	125	isomer 2			
Cys	189				

Finally, increased oxidation is also observed for the T11*-1 isomer of Met-125 residue, only in rhGH:Zn 1:10 samples, suggesting an associated long range conformational change. Though a trend in increased oxidation is also observed for 1:20 samples, there is increased variability of measurements for the 1:20 sample relative to those of other rhGH samples (Figure 87 a), therefore it this is deemed insignificant.

For the residues, Pro-2 and Phe-25, increased oxidation is observed for all zinc containing samples relative to rhGH, however the extent of oxidation appears to be unaffected by the relative ratio of zinc, suggesting surface accessibility at these sites to be equivalent in the presence of zinc.

Similarly structures at residues Phe-1, Met-14 and Cys-189 appear to be equivalent across all samples. As with previous characterisation of the rhGH:Zn model system, presented in Chapter 2, there is a trend of increased structural changes at higher zinc concentration, in the form of increased number of FPOP modification sites showing significant changes, i.e. three vs six for rhGH:Zn 1:2 and 1:10 ratios respectively, relative to rhGH control, suggesting the structural change of the model system is proportional to the amount of zinc present.

The extent of modification was generally low across all peptides and sample type, with a mean total %modification across all peptides in the range of ~ > 14%, as with capillary flow FPOP-MS measurements suggesting high levels of protein HOS for rhGH.

There is high correlation with the structural changes identified in the two FPOP methodologies, capillary-flow (Table 20) and single-shot Eppendorf (Table 23), validating the proposed latter methodology as an alternative potential workflow for FPOP labelling experiments.

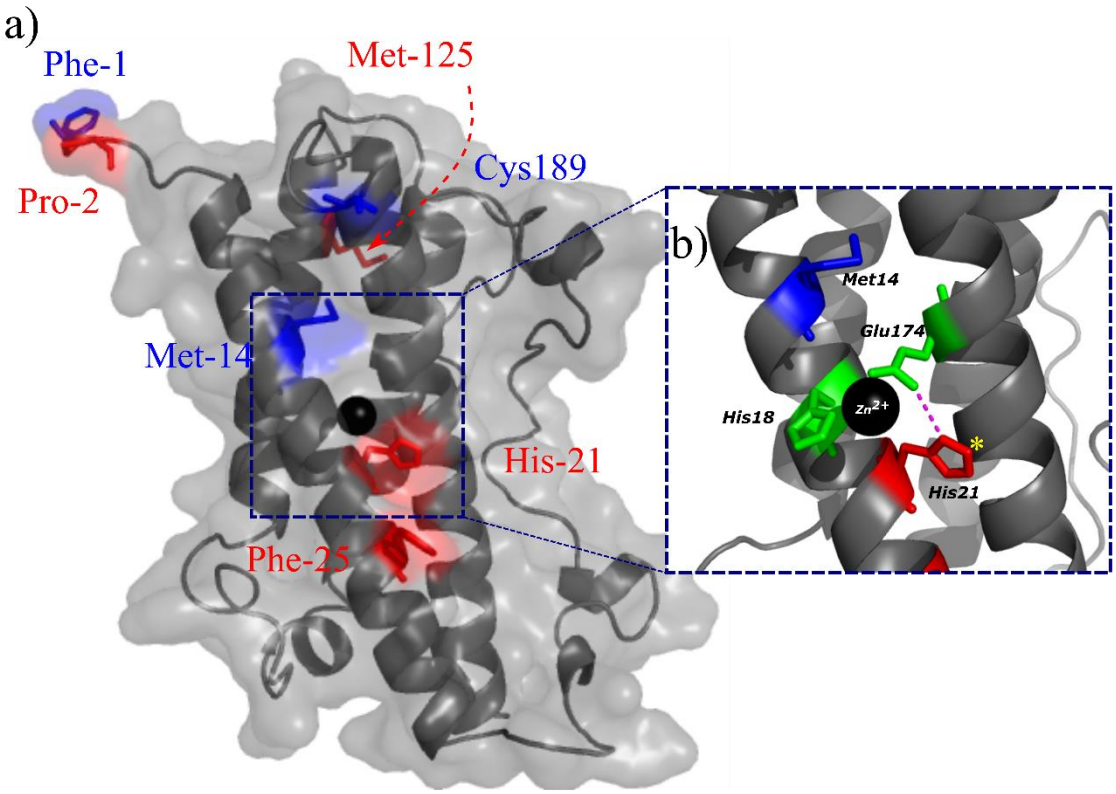


Figure 88. a) Those residues modified by single-shot FPOP experiments of rhGH are highlighted on the XRC structure of rhGH:Zn complex (PDB: 1BP3 [208]). Modified residues are highlighted, sites in which %modification remained unchanged (blue) and increased (red) for rhGH:Zn 1:10 relative to rhGH. Illustration of the b) Zn^{2+} binding site, residues His-18 and Glu-174 directly involved in Zn^{2+} binding are highlighted (green), as is the stabilising interaction of the His-21 to Glu-174. Increased oxidation occurs due to the presence of zinc at the His-21 imidazole site indicated (*), revealing an increased surface accessibility of this site upon zinc binding.

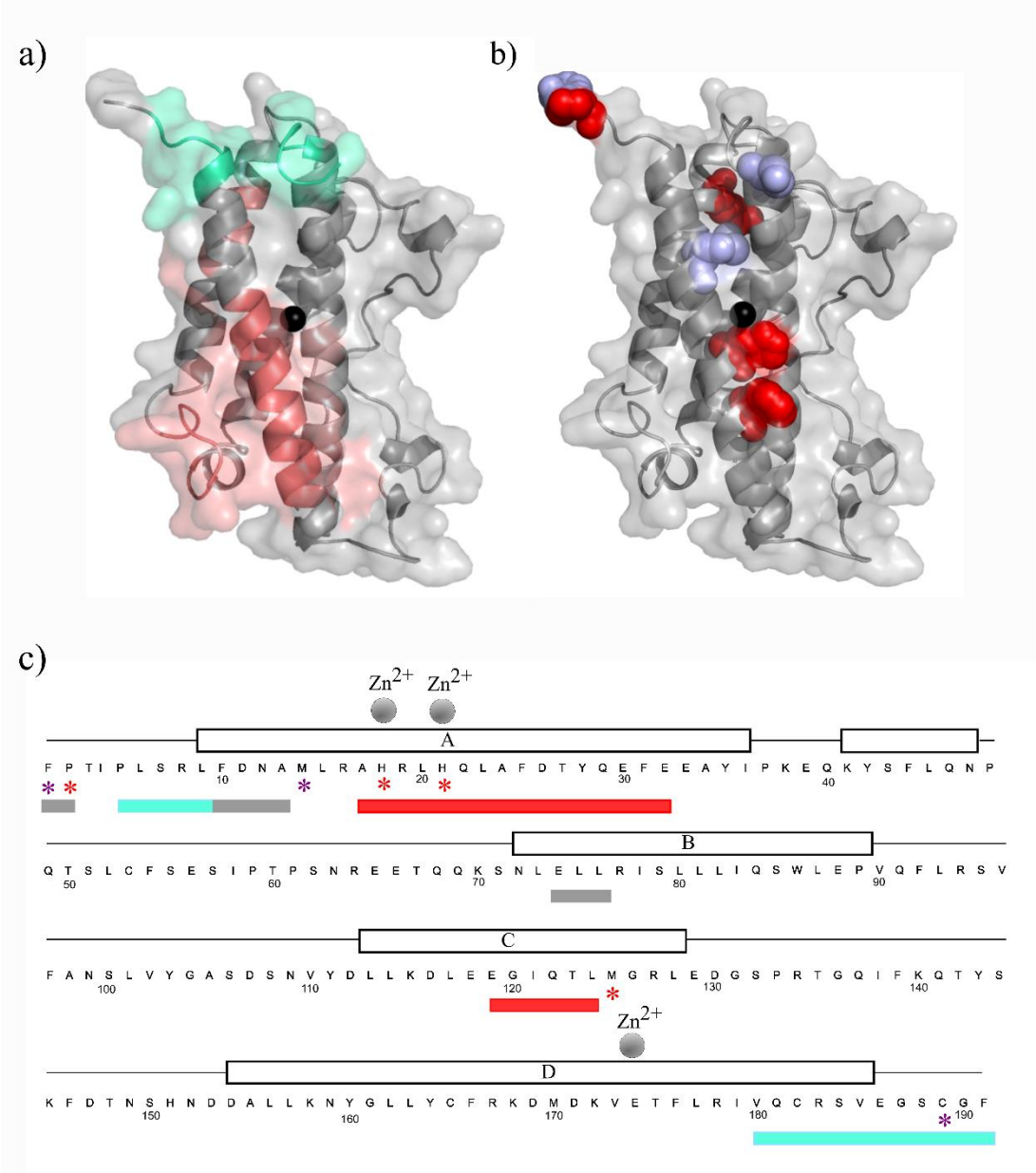


Figure 89. Structural changes of rhGH associated with the presence of zinc (rhGH:Zn 1:10) identified by a) HDX-MS and b) FPOP-MS measurements on the XRC structure of rhGH:Zn complex (PDB 1BP3, [208]); b) the primary and secondary (α-helices A-D indicated by bars) sequence of rhGH. Shaded regions illustrate those regions of increased deuterium uptake (red) and decreased uptake (cyan) in HDX-MS measurements, relative to free rhGH. Side chain residues modified by FPOP labelling experiments are modelled or indicated by a (*), both those that underwent no changes in oxidation (purple) and those that saw an increase in oxidation (red) due to the presence of zinc. Residues His-18, His-21 and Glu-174 indicated in Zn^{2+} binding are indicated. Sequence coverage missing from HDX-MS analysis are indicated by grey bars in c).

4.3.3 Comparison of FPOP-MS, HDX-MS and IMS-MS measurements

FPOP-MS and HDX-MS can be considered complementary protein footprinting strategies, with FPOP-MS providing structural insights into the surface accessibility of amino acid side chains and HDX-MS providing insights into hydrogen bonding and the solvent accessibility of backbone amides within a protein structure. To better understand the comparability of the two footprinting techniques and to characterise the rhGH model system further, the structural changes and/or dynamics identified in the between rhGH and zinc containing samples by both HDX-MS and FPOP-MS measurements are here considered. Figure 89, illustrates a summary of the structural changes identified in the two datasets for the rhGH:Zn 1:10 sample relative to rhGH.

For rhGH:Zn 1:10 samples, four regions of small but localised structural changes associated with the presence of zinc ligands were detected by HDX-MS (Chapter 2, section 2.3.1.6) measurements. Increased HDX rates were localised to residues 17-32 and 119-124, and reduced HDX rates for residues 5-9 and 180-191, demonstrating both increased and decreased structural dynamics respectively. In parallel, four amino acid side chains were identified as exhibiting increased oxidation/SASA by FPOP-MS analysis, Pro-2, His-21, Phe-25 and Met-125.

In Chapter 2, it was hypothesised that the increased HDX-MS measurements observed for residues 17-32, that include residues His-18 and His-21 involved in zinc binding, was as a result of the outwards orientation of this hydrophobic α-helix region (helix A). Should this occur, then one would expect the SASA of the correlating amino acid side chains to increase. Indeed increased FPOP modification rates were observed for both His-21 and Phe-25 in the presence of zinc. However, the third residue involved in zinc binding of rhGH, residue Glu-174 [208], revealed no significant structural change based on HDX-MS measurements. Similarly, there was an absence of modification for the Glu-174 side chain and surrounding residues, including Met-175, during FPOP-MS analysis. The very low exchange HDX rates (Figure 40) occurring in this region, suggest the burial of the residue within the highly hydrophobic central cavity of the rhGH structure. This combined with the absence of a FPOP modification site may tend to suggest this region to be a poor structural probe for HDX-MS and FPOP-MS measurements in this specific model system, highlighting some of the limitations of the techniques.

The increased HDX dynamics observed for region 119-124 (EGIQTL) in the presence of zinc, suggests destabilisation of this area. The lack of FPOP modification for these specific residues, makes commenting on structural changes on these residues impossible from the FPOP-MS measurements. However, FPOP-MS analysis did reveal increased surface accessibility for Met-125 side chain, which aligns to a very similar region of the primary sequence. When interpreting data it should be considered, that there is an increased propensity of methionine side chains to undergo hydroxyl radical oxidation (Table 3), compared to surrounding residues of EGIQTL, which may explain a preferential oxidation site at Met-125.

As for the increases in SASA observed for Phe-1 and Pro-2 in the presence of zinc, no HDX-MS data are available, as a result of the loss of N-terminal amide deuteration via back exchange [125]. FPOP-MS analysis in this context has offered structural insights unobtainable from conventional bottom up HDX-MS experiments. In contrast no FPOP modification was observed for the residues, 5-9 (PLSRL), a region which showed a decrease in structural dynamics in the presence of zinc as revealed by HDX-MS analysis. However, the increased modification of the two N-terminal residues could be an associated long range structural change as a result of changes in the 5-9 region. This demonstrates the utility of combing measurements obtained from both techniques.

Decreased dynamics for the C-terminal region, spanning residues 180-191 (VQCRSVEGSCGF) was also indicated by HDX-MS measurements. Though no specific amino acid resolution could be afforded, relative comparison of overlapping peptidic sequences in this region indicate that structural changes occur in 180-185 region as well as 185-191. Modification of the Cys189 residue was identified by FPOP-MS measurements, yet the modification was unaffected due to the presence of zinc. Combination of insights from both these datasets would indicate that the structural changes occur in 180-188 region. Finally, the structure of residue Met-14 appears to be unaffected by the presence of zinc in both FPOP-MS and HDX-MS data sets.

Similarly for rhGH:Zn 1:2 samples, comparison of structural changes as identified by both HDX- and FPOP-MS analysis, relative to free rhGH, show good correlation. Only structural changes in the 17-32 region were identified by HDX-MS measurements, in the form of increased dynamics in presence of zinc, this aligns with

the increased oxidation of Phe-25-para and -meta sites as measured by FPOP-MS. In addition, increased oxidisation was also observed for Pro-2 during FPOP-MS, though no sequence coverage was observed for this residue as discussed, previously.

Overall a very strong correlation exists between the localised regions of structural change, identified by HDX- and FPOP-MS for the rhGH model system. Both data sets indicate small localised structural changes in HOS for rhGH due to the presence of zinc, that are both specific to the binding site and alternative long range sites.

CIU-IMS-MS measurements indicate a small increase in global stability for rhGH in the presence of zinc, however localised structural insights from HDX-MS and FPOP-MS measurements showed a combination of destabilising and stabilising changes in secondary and tertiary structures. Though all technologies are showing to be effective at indicating a change in HOS, this suggests a disparity between MS HOS measurements at a global and localised peptidic level. Perhaps this indicates a necessity to characterise or monitor changes in HOS on multiple levels, to give confidence and to comply with such regulation as is indicated in pharmaceutical guidelines.

4.4 Conclusions

The rhGH:Zn model system has here been successfully applied not only to validate and illustrate comparability of the two FPOP-MS workflows described but also to assess their sensitivity towards small structural changes. For the single-shot Eppendorf FPOP workflow, structural changes were detected and characterised at an amino acid level at both the lower, 1:2, and higher, 1:10 rhGH:Zn ratio, demonstrating comparable sensitivity to that of HDX-MS. Additionally, the model system was used to explore and validate a novel FPOP-MS labelling workflow using a single-shot labelling approach, offering a potentially higher through-put and an alternative labelling strategy should protein samples be inapplicable to conventional capillary flow FPOP labelling strategies. FPOP-MS measurements performed in this thesis have been based on relative quantification between samples. As aforementioned, the potential to apply established quantitative workflows within the FPOP-MS workflow is high, and is this an avenue which moving forward should be explored.

A relatively new analytical platform compared to that of HDX-MS, FPOP-MS is less established in an industrial setting, and requires systematic validation of analysis

workflows to facilitate use in a regulatory context. The data acquired here not only demonstrate the strong correlation and complementary structural information that can be gained using multiple protein footprinting techniques, but also help to validate the FPOP-MS technology as a whole. The current data workflows for FPOP-MS data analysis are very manual and cumbersome, but should they be more streamlined and robust it is possible to achieve amino acid resolution with a less data heavy approach compared to HDX-MS workflows.

Chapter 5

Conclusions and future work

Chapter 5
Conclusions and future works

The development of protein based therapeutics has revolutionised the biopharmaceutical industry, resulting in these biotherapeutics becoming a dominant drug class. Secondary, tertiary and quaternary HOS structures of these biotherapeutics are considered CQAs for reasons of stability, efficacy and immunogenicity. For these reasons protein HOS measurements are necessary throughout the product life cycle, used to demonstrate comparability of biotherapeutics when changes in the manufacturing process occur [61] or even bio-similarity between a proposed product and an innovator product [218]. However, many of the analytical capabilities for the characterisation and sufficient monitoring of changes in the HOS and/or other CQAs of these drug candidates, have not yet been fully validated to wholly support the regulations defining such drugs. Though a wide plethora of analytical technologies exist for the analysis of protein HOS, little is known about the suitable applicability or sensitivity towards small structural changes of a particular candidate drug for each technology. In addition, regulators encourage the use of multiple analytical platforms for monitoring for changes in HOS, yet no specific guidance is provided on which techniques should be implemented. Nor is there guidance on how to provide data to regulators which establishes the sensitivity of these techniques towards small but reproducible changes in HOS.

In the clinical and diagnostic area use of RMs has long since been used for method development, validation, uncertainty definitions and indeed harmonisation of quantification assays [57, 58]. Such RMs to support the utilisation of protein HOS measurements have, however, been absent.

This thesis aimed to assess, develop and characterise two potential model systems of increasing complexity, based on protein RMs for use to support analytical method validation and define sensitivity of different MS–based analytical platforms towards small structural changes. These protein RMs were stressed to induced structural changes, simulating small changes which may occur during manufacturing and processing of a candidate drug biomolecule.

In Chapter 2, the WHO material 98/574, an international RM based on rhGH, a small, relatively simple bio-therapeutic, was first used to systematically develop analytical

methods for HOS measurements in a robust metrological manner, defining the repeatability and reproducibility of HDX-MS and IMS-MS methods developed. For both technologies, it was demonstrated that the inter-day reproducibility dominated over within-day measurements.

Once the analytical methods were developed and their variability critically assessed, the sensitivity of these methods towards small structural changes was tested using a model system based on rhGH, in which varying amount of zinc ligands were introduced so as to induce structural changes of varying magnitudes. This model system was used as a proof-of-principle to test a statistical model developed here based on HDX-MS measurements, differentiating between measurement variability and significant HOS change by considering measurement thresholds. Both HDX-MS and IMS-MS measurements revealed that as the relative amount of zinc was increased the magnitude of structural change increased, with HDX-MS providing specific localised structural changes and global insights from IMS-MS measurements. These varying magnitudes of structural changes characterised for the model system, offer a valuable tool for establishing the sensitivity of an analytical platform towards structural changes.

In Chapter 4, it's illustrated, how these data can be used as reference datasets to develop and benchmark alternative analytical methods, such as was demonstrated for both an established and a novel higher throughput FPOP-MS workflow. Small localised structural changes associated with the presence of zinc ligands as characterised by HDX-MS measurements, were confirmed by orthogonal FPOP-MS measurements. Analogous sensitivity towards small structural changes was demonstrated for the FPOP-MS and HDX-MS methods developed, with IMS-MS providing similar sensitivity albeit at a lower resolution, though more conducive to a higher throughput QC type environment. Figure 90, illustrates the extensive characterisation of all levels of protein HOS of the rhGH:Zn model system using largely MS-based measurements produced in this thesis.

Validation of HDX-MS and IMS-MS analytical methods plus the HDX-MS statistical model developed in Chapter 2, was illustrated using a model system, based on the NISTmAb RM 8761, a more complex and potentially pharmaceutically relevant molecule, in Chapter 3. In order to stress the protein HOS of the RM molecule, enzymes were used to generate differing Fc-glycan states, analogous to the type of

changes to a CQA that regulators state must be monitored for during manufacturing processes [62]. Again both HDX-MS and IMS-MS methods were demonstrated as being effective in monitoring for changes in protein HOS, though it was apparent that HDX-MS measurements demonstrated superior sensitivity towards smaller structural changes. The structural changes identified were analogous to previous studies of mAbs which show destabilising kinetics as a result of glycan structure loss [130, 279, 287], further validating the approach and methods developed.

One of the aims of this thesis was to develop and fully characterise model protein RM systems to support comparability studies to better understand analytical technologies for protein HOS measurements. Figure 90 and Figure 91 represent the information extracted from the analytical measurements produced in this thesis largely using MS-based measurements, providing reference datasets at all levels of higher order structure for these model systems which can be compared with cross and with-platform measurements performed in other laboratories.

The comparability of several MS-based methods has been demonstrated here, however more data need to be produced to better understand cross comparison measurements. One manner in which this can be achieved is by performing inter-laboratory comparisons such as that undertaken with NISTmAb HDX-MS measurements [296] though this study focuses on the comparability of absolute measurements and not the ability of methods to detect structural changes. In contrast, these model systems offer candidate systems for such studies; indeed, work has begun on collating the data from an interlaboratory comparison conducted outside the scope of this thesis based on the rhGH:Zn model system [233]. In this study analysis of the model system was performed across multiple platforms, including biophysical technologies, such as Raman, DSC, AUC, NMR, HDX-MS (multiple labs), IMS-MS (multiple labs) to mention but a few.

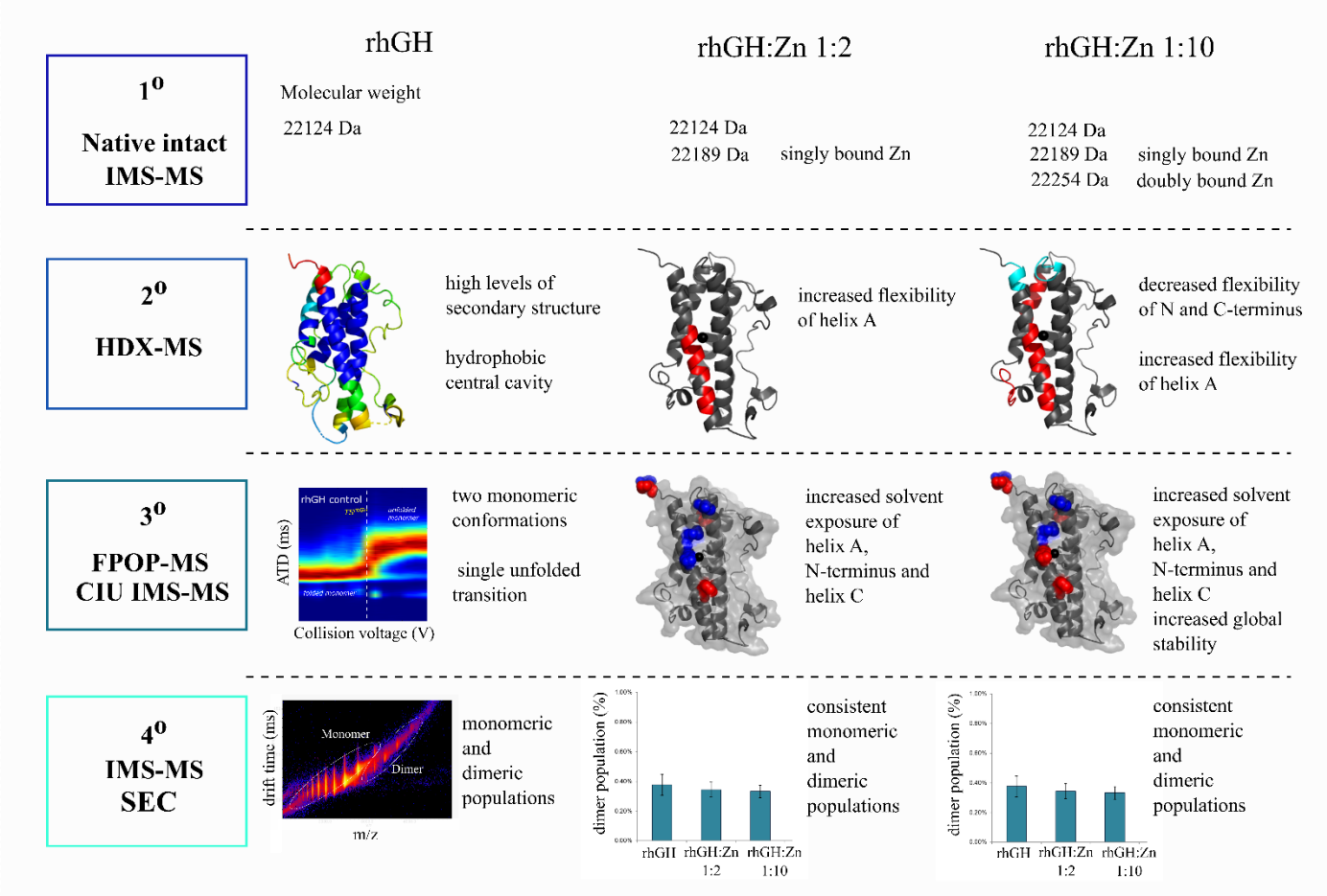


Figure 90. Higher order structure characterisation of rhGH RM and rhGH:Zn model system. Structural interpretation for rhGH:Zn samples are described relative to rhGH RM control.

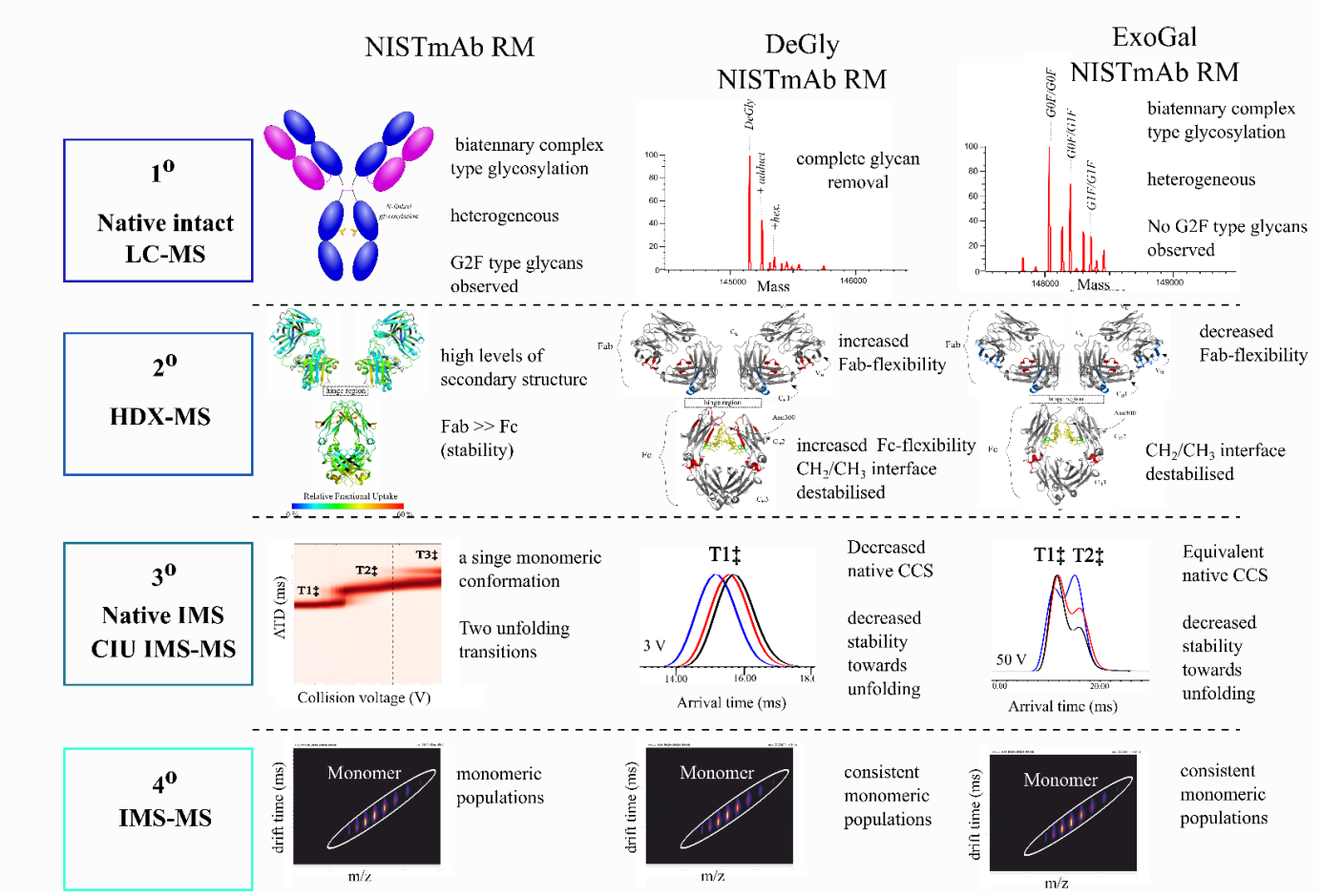


Figure 91. Higher order structure characterisation of NISTmAb RM and deglycosylated and degalactosylated Fc-glycan variants (DeGly and ExoGal respectively). Structural interpretation for Fc-glycan variants are described relative to NISTmAb RM control.

Though the statistical model developed for defining thresholds for HDX-MS measurements (Chapter 2 & 3) was proven to be effective, the measurement, data collation and analysis was relatively cumbersome and not amenable to high throughput QC analysis. It is hoped that with further analysis of the data, and complementary datasets from the inter-laboratory comparison mentioned, the sources of measurement uncertainty and their overall contribution herein can be better defined for HDX-MS measurements. This may allow for the definition of measurement thresholds for this RM model system which can be produced from a less comprehensive dataset, i.e. generation of fewer replicate datasets, in alternative laboratories. In addition by performing in depth analysis of the measurement uncertainties associated with HDX-MS measurements this is a step forwards to the standardisation of HDX-MS measurements.

Recently, the first comparisons of FPOP-MS with HDX-MS measurements [110, 113, 297] have validated and illustrated the potential value of the FPOP-MS analytical technology within the biopharmaceutical area. Yet, the lack of software and established data analysis workflows for these analysis is likely to hold the technique back within this context. As demonstrated in Chapter 4, it is possible that these model RM systems can be used to drive development in this area and move towards standardisation. For example, analysis of FPOP-MS data showed labelling reproducibility to be a limiting factor for increasing the sensitivity towards structural changes using the FPOP-MS methods here described in this thesis. The potential introduction of internal standards, may allow for the reduction of such measurement uncertainty and hence increase the sensitivity of the technique towards small structural changes.

Finally, from a metrological perspective, it is hoped that through the development and validation of protein HOS measurements a better understanding of the link between the traceable quantification of a protein based on amount (~ kg, often based on the primary structure) and the activity of a protein in an assay (of which protein HOS can be a contributing factor) can be established. Disparity between these two measurements make establishing a traceability chain from assay measurement to a SI unit limited. For example, performing such studies as epitope mapping of clinical assay targets can assist in defining the measurands of clinical assays, which in turn will assist with harmonisation of clinical measurands. Alternative protein footprinting technologies such as XL-MS, due to their applicability to a complex biological matrix

and potential for introduction of traceable standards, offer a potential pathway for the quantification of protein HOS and hence ultimately establish a link between primary sequence and activity of a protein in a clinical assay. The use of such RM model systems as those proposed here will be pivotal in establishing these links and aid future development.

As for the bio-pharmaceutical industry, structural MS platforms show great potential for structural characterisation of bio-therapeutics and are fast becoming established within the bio-analytical toolbox of these industries. Use of appropriate standards and QCs will assist the standardisation of these techniques, and for IMS-MS in particular, the potential move away from an R&D environment and into a more high throughout QC type environment. As for use within the R&D environment itself, the application of HDX-MS either on its own and in conjunction with other structural technique, such as NMR[298], XL-MS[299] and native-MS[300] has been repeatedly demonstrated as being at the cutting edge of studying protein structure and dynamics. In recent years there has been an expansion of the use of XL-MS through developments in cross-linking chemistry and software for cross-linkage identification. In their current state, XL-MS and FPOP-MS suffer from similar limitations of low abundance of modification, issues of identification and the lack of standardisation and automated downstream data workflows. However if these issues are continued to be addressed by the community these techniques show great potential for these techniques providing structural insights into bio-therapeutics both *in* and *ex vivo* and on a potentially quantitative basis, a feat that would unique to such high resolution techniques.

List of References

1. Thomson, J.J.: Bakerian Lecture: Rays of Positive Electricity. Proc. R. Soc. Lond. Ser. A. 89, 1–20 (1913)
2. Aston, F.W.: A Critical Search for a Heavier Constituent of the Atmosphere by Means of the Mass-Spectrograph. Proc. R. Soc. Lond. Ser. A. 103, 462–469 (1923)
3. Ruotolo, B.T.: Evidence for macromolecular protein rings in the absence of bulk water. Science. 310, 1658–1661 (2005). https://doi.org/10.1126/science.1120177
4. Uetrecht, C., Versluis, C., Watts, N.R., Wingfield, P.T., Steven, A.C., Heck, A.J.R.: Stability and Shape of Hepatitis B Virus Capsids In Vacuo. Angew. Chem. Int. Ed. 47, 6247–6251 (2008). https://doi.org/10.1002/anie.200802410
5. Uetrecht, C., Versluis, C., Watts, N.R., Roos, W.H., Wuite, G.J.L., Wingfield, P.T., Steven, A.C., Heck, A.J.R.: High-resolution mass spectrometry of viral assemblies: Molecular composition and stability of dimorphic hepatitis B virus capsids. Proc. Natl. Acad. Sci. 105, 9216–9220 (2008). https://doi.org/10.1073/pnas.0800406105
6. Taylor, K.E., van den Berg, C.W.: Structural and functional comparison of native pentameric, denatured monomeric and biotinylated C-reactive protein. Immunology. 120, 404–411 (2007). https://doi.org/10.1111/j.1365-2567.2006.02516.x
7. Taylor, G.I., McEwan, A.D.: The stability of a horizontal fluid interface in a vertical electric field. J. Fluid Mech. 22, 1–15 (1965). https://doi.org/10.1017/S0022112065000538
8. Taflin, D.C., Ward, T.L., Davis, E.J.: Electrified droplet fission and the Rayleigh limit. Langmuir. 5, 376–384 (1989). https://doi.org/10.1021/la00086a016
9. Iribarne, J.V.: On the evaporation of small ions from charged droplets. J. Chem. Phys. 64, 2287–2294 (1976). https://doi.org/10.1063/1.432536
10. Kebarle, P., Verkerk, U.H.: Electrospray: From ions in solution to ions in the gas phase, what we know now. Mass Spectrom. Rev. 28, 898–917 (2009). https://doi.org/10.1002/mas.20247
11. Znamenskiy, V., Marginean, I., Vertes, A.: Solvated Ion Evaporation from Charged Water Nanodroplets. J. Phys. Chem. A. 107, 7406–7412 (2003). https://doi.org/10.1021/jp034561z
12. Dole, M.: Molecular Beams of Macroions. J. Chem. Phys. 49, 2240–2249 (1968). https://doi.org/10.1063/1.1670391
13. Konermann, L., Ahadi, E., Rodriguez, A.D., Vahidi, S.: Unraveling the mechanism of electrospray ionization. Anal. Chem. 85, 2–9 (2013). https://doi.org/10.1021/ac302789c
14. Wilm, M., Mann, M.: Analytical Properties of the Nanoelectrospray Ion Source. Anal. Chem. 68, 1–8 (1996). https://doi.org/10.1021/ac9509519
15. Wilm, M.: Principles of Electrospray Ionization. Mol. Cell. Proteomics. 10, (2011). https://doi.org/10.1074/mcp.M111.009407

16. Murray, K.K., Boyd, R.K., Eberlin, M.N., Langley, G.J., Li, L., Naito, Y.: Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013). Pure Appl. Chem. 85, 1515–1609 (2013). https://doi.org/10.1351/PAC-REC-06-04-06

17. IUPAC, I.: Compendium of chemical terminology. Blackwell Scientific Publications Oxford (1997)

18. Miller, P.E., Denton, M.B.: The Quadrupole Mass Filter - Basic Operating Concepts. J. Chem. Educ. 63, 617–622 (1986)

19. Paul, W.: Electromagnetic traps for charged and neutral particles. Rev. Mod. Phys. 62, 531 (1990)

20. Dawson, P.: Quadrupole Mass Spectrometry and its Applications. Elsevier (1976)

21. March, R.E.: An Introduction to Quadrupole Ion Trap Mass Spectrometry. J. Mass Spectrom. 32, 351–369 (1997). https://doi.org/10.1002/(SICI)1096-9888(199704)32:4<351::AID-JMS512>3.0.CO;2-Y

22. de Hoffmann, E., Stroobant, V.: Mass Spectrometry: Principles and Applications, 3rd Edition | Wiley. (2013)

23. Stephens, W., E.: Proceedings of the American Physical Society. Phys. Rev. 69, 674–674 (1946). https://doi.org/10.1103/PhysRev.69.674.2

24. Wiley, W.C., McLaren, I.H.: Time-of-Flight Mass Spectrometer with Improved Resolution. Rev. Sci. Instrum. 26, 1150 (1955). https://doi.org/10.1063/1.1715212

25. Mamryin, B., Karataev, V., Schmikk, D., Zagulin, V.: The Mass-Reflectron, a new non-magnetic time of flight mass spectrometer with high resolution. Sov. Phys. JETP. 37, 45 (1973)

26. Dawson, J.H.J., Guilhaus, M.: Orthogonal-acceleration time-of-flight mass spectrometer. Rapid Commun. Mass Spectrom. 3, 155–159 (1989). https://doi.org/10.1002/rcm.1290030511

27. Kingdon, K.H.: A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures. Phys. Rev. 21, 408–418 (1923). https://doi.org/10.1103/PhysRev.21.408

28. Makarov, A.A.: Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. Anal. Chem. 72, 1156–1162 (2000)

29. Makarov, A., Denisov, E., Lange, O.: Performance Evaluation of a High-field Orbitrap Mass Analyzer. J. Am. Soc. Mass Spectrom. 20, 1391–1396 (2009). https://doi.org/10.1016/j.jasms.2009.01.005

30. Giles, K.: Travelling wave ion mobility. Int. J. Ion Mobil. Spectrom. 16, 1–3 (2013). https://doi.org/10.1007/s12127-013-0125-5

31. Waters Corporation : The Science of What's Possible : Waters, https://www.waters.com/nextgen/gb/en.html

32. 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. Mol. Cell. Proteomics. 10, M111.011015 (2011). https://doi.org/10.1074/mcp.M111.011015

33. Thermo Fisher Scientific - UK, https://www.thermofisher.com/uk/en/home.html
34. McLafferty, F.W.: Tandem mass spectrometry (MS/MS): a promising new analytical technique for specific component determination in complex mixtures. Acc. Chem. Res. 13, 33–39 (1980). https://doi.org/10.1021/ar50146a001
35. Jennings, K.R.: Collision-induced decompositions of aromatic molecular ions. Int. J. Mass Spectrom. Ion Phys. 1, 227–235 (1968). https://doi.org/10.1016/0020-7381(68)85002-8
36. Syka, J.E.P., Coon, J.J., Schroeder, M.J., Shabanowitz, J., Hunt, D.F.: Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. PNAS. 101, 9528–33 (2004)
37. Zubarev, R.A., Kelleher, N.L., McLafferty, F.W.: Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. J. Am. Chem. Soc. 120, 3265–3266 (1998). https://doi.org/10.1021/ja973478k
38. Mabud, Md.A., Dekrey, M.J., Graham Cooks, R.: Surface-induced dissociation of molecular ions. Int. J. Mass Spectrom. Ion Process. 67, 285–294 (1985). https://doi.org/10.1016/0168-1176(85)83024-X
39. Ly, T., Julian, R.R.: Ultraviolet Photodissociation: Developments towards Applications for Mass-Spectrometry-Based Proteomics. Angew. Chem. Int. Ed. 48, 7130–7137 (2009). https://doi.org/10.1002/anie.200900613
40. Roepstorff, P., Fohlman, J.: Proposal for a common nomenclature for sequence ions in mass-sepctra of peptides. Biol. Mass Spectrom. 11, 601–601 (1984). https://doi.org/10.1002/bms.1200111109
41. Brodbelt, J.S.: Ion Activation Methods for Peptides and Proteins. Anal. Chem. 88, 30–51 (2016). https://doi.org/10.1021/acs.analchem.5b04563
42. Wysocki, V.H., Tsaprailis, G., Smith, L.L., Breci, L.A.: Mobile and localized protons: a framework for understanding peptide dissociation. J. Mass Spectrom. JMS. 35, 1399–1406 (2000). https://doi.org/10.1002/1096-9888(200012)35:12<1399::AID-JMS86>3.0.CO;2-R
43. Bowers, M.T., Kemper, P.R., Helden, G. von, Koppen, P.A.M. van: Gas-Phase Ion Chromatography: Transition Metal State Selection and Carbon Cluster Formation. Science. 260, 1446–1451 (1993). https://doi.org/10.1126/science.260.5113.1446
44. Campuzano, I., Bush, M.F., Robinson, C.V., Beaumont, C., Richardson, K., Kim, H., Kim, H.I.: Structural Characterization of Drug-like Compounds by Ion Mobility Mass Spectrometry: Comparison of Theoretical and Experimentally Derived Nitrogen Collision Cross Sections. Anal. Chem. 84, 1026–1033 (2012). https://doi.org/10.1021/ac202625t
45. Valentine, S.J., Counterman, A.E., Clemmer, D.E.: A database of 660 peptide ion cross sections: Use of intrinsic size parameters for bona fide predictions of cross sections. J. Am. Soc. Mass Spectrom. 10, 1188–1211 (1999). https://doi.org/10.1016/S1044-0305(99)00079-3

46. Benesch, J.L., Ruotolo, B.T.: Mass spectrometry: come of age for structural and dynamical biology. Curr. Opin. Struct. Biol. 21, 641–649 (2011). https://doi.org/10.1016/j.sbi.2011.08.002

47. Revercomb, H.E., Mason, E.A.: Theory of plasma chromatography/gaseous electrophoresis. Review. Anal. Chem. 47, 970–983 (1975). https://doi.org/10.1021/ac60357a043

48. Mason, E.A., Schamp, H.W.: Mobility of gaseous lons in weak electric fields. Ann. Phys. 4, 233–270 (1958). https://doi.org/10.1016/0003-4916(58)90049-6

49. Ferguson, C.N., Gucinski-Ruth, A.C.: Evaluation of Ion Mobility-Mass Spectrometry for Comparative Analysis of Monoclonal Antibodies. J. Am. Soc. Mass Spectrom. 27, 822–833 (2016). https://doi.org/10.1007/s13361-016-1369-1

50. Pringle, S.D., Giles, K., Wildgoose, J.L., Williams, J.P., Slade, S.E., Thalassinos, K., Bateman, R.H., Bowers, M.T., Scrivens, J.H.: An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. Int. J. Mass Spectrom. 261, 1–12 (2007)

51. Giles, K., Pringle, S.D., Worthington, K.R., Little, D., Wildgoose, J.L., Bateman, R.H.: Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. Rapid Commun. Mass Spectrom. 18, 2401–2414 (2004). https://doi.org/10.1002/rcm.1641

52. Bush, M.F., Hall, Z., Giles, K., Hoyes, J., Robinson, C.V., Ruotolo, B.T.: Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. Anal. Chem. 82, 9557–9565 (2010). https://doi.org/10.1021/ac1022953

53. Smith, D., Knapman, T., Campuzano, I., Malham, R., Berryman, J., Radford, S., Ashcroft, A.: Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. Eur. J. Mass Spectrom. 15, 113 (2009). https://doi.org/10.1255/ejms.947

54. Wyttenbach, T., von Helden, G., Batka, J.J., Carlat, D., Bowers, M.T.: Effect of the long-range potential on ion mobility measurements. J. Am. Soc. Mass Spectrom. 8, 275–282 (1997). https://doi.org/10.1016/S1044-0305(96)00236-X

55. Shvartsburg, A.A., Jarrold, M.F.: An exact hard-spheres scattering model for the mobilities of polyatomic ions. Chem. Phys. Lett. 261, 86–91 (1996). https://doi.org/10.1016/0009-2614(96)00941-4

56. Shvartsburg, A.A., Schatz, G.C., Jarrold, M.F.: Mobilities of carbon cluster ions: Critical importance of the molecular attractive potential. J. Chem. Phys. 108, 2416–2324 (1998)

57. Pritchard, C., Quaglia, M., Mussell, C., Burkitt, W.I., Parkes, H., O'Connor, G.: Fully traceable absolute protein quantification of somatropin that allows independent comparison of somatropin standards. Clin. Chem. 55, 1984–90 (2009)

58. Torma, A., Groves, K., Biesenbruch, S., Mussell, C., Reid, A., Ellison, S., Cramer, R., Quaglia, M.: A candidate liquid chromatography mass spectrometry reference method for the quantification of the cardiac marker 1-32 B-type

natriuretic peptide. Clin. Chem. Lab. Med. 55, (2017). https://doi.org/10.1515/cclm-2016-1054

59. Urquhart, L.: Top drugs and companies by sales in 2018. Nat. Rev. Drug Discov. 18, 245–245 (2019). https://doi.org/10.1038/d41573-019-00049-0

60. European Medicines Agency: ICH Q8R2 Guideline on pharmaceutical development, http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000789.jsp&mid=WC0b01ac0580028eb2, (2005)

61. European Medicines Agency: ICH Q6B Specifications: test procedures and acceptance criteria for biotechnological/biological products, http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000883.jsp&mid=WC0b01ac058002956b, (1999)

62. European Medicines Agency: ICH Guideline Q11 on development and manufacture of drug substances (chemical entities and biotechnological/biological entities), Step 3, https://www.ema.europa.eu/en/documents/scientific-guideline/draft-ich-guideline-q11-development-manufacture-drug-substances-chemical-entities-biotechnological/biological-entities_en.pdf, (2011)

63. Shepherd, D.A., Marty, M.T., Giles, K., Baldwin, A.J., Benesch, J.L.P.: Combining tandem mass spectrometry with ion mobility separation to determine the architecture of polydisperse proteins. Int. J. Mass Spectrom. 377, 663–671 (2015). https://doi.org/10.1016/j.ijms.2014.09.007

64. Zhang, Q., Noble, K.A., Mao, Y., Young, N.L., Sathe, S.K., Roux, K.H., Marshall, A.G.: Rapid Screening for Potential Epitopes Reactive with a Polycolonal Antibody by Solution-Phase H/D Exchange Monitored by FT-ICR Mass Spectrometry. J. Am. Soc. Mass Spectrom. 24, 1016–1025 (2013). https://doi.org/10.1007/s13361-013-0644-7

65. Freeke, J., Bush, M.F., Robinson, C.V., Ruotolo, B.T.: Gas-phase protein assemblies: Unfolding landscapes and preserving native-like structures using noncovalent adducts. Chem. Phys. Lett. 524, 1–9 (2012). https://doi.org/10.1016/j.cplett.2011.11.014

66. Kaltashov, I.A., Bobst, C.E., Abzalimov, R.R.: H/D Exchange and Mass Spectrometry in the Studies of Protein Conformation and Dynamics: Is There a Need for a Top-Down Approach? Anal. Chem. 81, 7892–7899 (2009). https://doi.org/10.1021/ac901366n

67. Smith, D.P., Radford, S.E., Ashcroft, A.E.: Elongated oligomers in β2-microglobulin amyloid assembly revealed by ion mobility spectrometry-mass spectrometry. Proc. Natl. Acad. Sci. 107, 6794–6798 (2010). https://doi.org/10.1073/pnas.0913046107

68. Kaiser, P., Akerboom, T., Molnar, P., Reinauer, H.: Modified HPLC-electrospray ionization/mass spectrometry method for HbA1c based on IFCC reference measurement procedure. Clin. Chem. 54, 1018–1022 (2008). https://doi.org/10.1373/clinchem.2007.100875

69. Levine, D.M., Maine, G.T., Armbruster, D.A., Mussell, C., Buchholz, C., O'Connor, G., Tuck, V., Johnston, A., Holt, D.W.: The need for standardization

of tacrolimus assays. Clin. Chem. 57, 1739–1747 (2011). https://doi.org/10.1373/clinchem.2011.172080

70. Stokes, P., O'Connor, G.: Development of a liquid chromatography-mass spectrometry method for the high-accuracy determination of creatinine in serum. J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 794, 125–136 (2003)

71. Antonyuk, S.V., Rustage, N., Petersen, C.A., Arnst, J.L., Heyes, D.J., Sharma, R., Berry, N.G., Scrutton, N.S., Eady, R.R., Andrew, C.R., Hasnain, S.S.: Carbon monoxide poisoning is prevented by the energy costs of conformational changes in gas-binding haemproteins. Proc. Natl. Acad. Sci. 108, 15780–15785 (2011). https://doi.org/10.1073/pnas.1109051108

72. Lubkowski, J., Dauter, M., Aghaiypour, K., Wlodawer, A., Dauter, Z.: Atomic resolution structure of Erwinia chrysanthemi L-asparaginase. Acta Crystallogr. D Biol. Crystallogr. 59, 84–92 (2003). https://doi.org/10.1107/s0907444902019443

73. Costa, T.R.D., Ignatiou, A., Orlova, E.V.: Structural Analysis of Protein Complexes by Cryo Electron Microscopy. Methods Mol. Biol. Clifton NJ. 1615, 377–413 (2017). https://doi.org/10.1007/978-1-4939-7033-9_28

74. Loo, J.A.: Observation of large subunit protein complexes by electrospray ionization mass spectrometry. J. Mass Spectrom. 30, 180–183 (1995). https://doi.org/10.1002/jms.1190300127

75. Light-Wahl, K.J., Schwartz, B.L., Smith, R.D.: Observation of the Noncovalent Quaternary Associations of Proteins by Electrospray Ionization Mass Spectrometry. J. Am. Chem. Soc. 116, 5271–5278 (1994). https://doi.org/10.1021/ja00091a035

76. Loo, J.A., Loo, R.R.O., Udseth, H.R., Edmonds, C.G., Smith, R.D.: Solvent-induced conformational changes of polypeptides probed by electrospray-ionization mass spectrometry. Rapid Commun. Mass Spectrom. 5, 101–105 (1991). https://doi.org/10.1002/rcm.1290050303

77. Fernandez de la Mora, J.: Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. Anal. Chim. Acta. 406, 93–104 (2000). https://doi.org/10.1016/S0003-2670(99)00601-7

78. Rose, R.J., Damoc, E., Denisov, E., Makarov, A., Heck, A.J.R.: High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. Nat. Methods. 9, 1084–1086 (2012). https://doi.org/10.1038/nmeth.2208

79. Snijder, J., Rose, R.J., Veesler, D., Johnson, J.E., Heck, A.J.R.: Studying 18 Mega Dalton Virus Assemblies with Native Mass Spectrometry. Angew. Chem. Int. Ed Engl. 52, 4020–4023 (2013). https://doi.org/10.1002/anie.201210197

80. Young, M.M., Tang, N., Hempel, J.C., Oshiro, C.M., Taylor, E.W., Kuntz, I.D., Gibson, B.W., Dollinger, G.: High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. Proc. Natl. Acad. Sci. U. S. A. 97, 5802–5806 (2000). https://doi.org/10.1073/pnas.090099097

81. Gabelica, V., Galic, N., Rosu, F., Houssier, C., De Pauw, E.: Influence of response factors on determining equilibrium association constants of non-covalent complexes by electrospray ionization mass spectrometry. J. Mass Spectrom. JMS. 38, 491–501 (2003). https://doi.org/10.1002/jms.459

82. Chowdhury, S.K., Katta, V., Chait, B.T.: Probing conformational changes in proteins by mass spectrometry. J. Am. Chem. Soc. 112, 9012–9013 (1990). https://doi.org/10.1021/ja00180a074

83. Kaltashov, I.A., Mohimen, A.: Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. Anal. Chem. 77, 5370–5379 (2005). https://doi.org/10.1021/ac050511+

84. PDB-101: Learn: Structural Biology Highlights: Disordered Proteins, https://pdb101.rcsb.org/learn/structural-biology-highlights/disordered-proteins

85. Chernushevich, I.V., Thomson, B.A.: Collisional cooling of large ions in electrospray mass spectrometry. Anal. Chem. 76, 1754–1760 (2004). https://doi.org/10.1021/ac035406j

86. Badman, E.R., Hoaglund-Hyzer, C.S., Clemmer, D.E.: Monitoring structural changes of proteins in an ion trap over approximately 10-200 ms: unfolding transitions in cytochrome c ions. Anal. Chem. 73, 6000–6007 (2001). https://doi.org/10.1021/ac010744a

87. Beck, A., Debaene, F., Diemer, H., Wagner-Rousset, E., Colas, O., Van Dorsselaer, A., Cianférani, S.: Cutting-edge mass spectrometry characterization of originator, biosimilar and biobetter antibodies. J. Mass Spectrom. JMS. 50, 285–297 (2015). https://doi.org/10.1002/jms.3554

88. Brown, K.A., Rajendran, S., Dowd, J., Wilson, D.J.: Rapid characterization of structural and functional similarity for a candidate bevacizumab (Avastin) biosimilar using a multipronged mass-spectrometry-based approach. Drug Test. Anal. 11, 1207–1217 (2019). https://doi.org/10.1002/dta.2609

89. Gabelica, V., Shvartsburg, A.A., Afonso, C., Barran, P., Benesch, J.L.P., Bleiholder, C., Bowers, M.T., Bilbao, A., Bush, M.F., Campbell, J.L., Campuzano, I.D.G., Causon, T., Clowers, B.H., Creaser, C.S., Pauw, E.D., Far, J., Fernandez-Lima, F., Fjeldsted, J.C., Giles, K., Groessl, M., Hogan, C.J., Hann, S., Kim, H.I., Kurulugama, R.T., May, J.C., McLean, J.A., Pagel, K., Richardson, K., Ridgeway, M.E., Rosu, F., Sobott, F., Thalassinos, K., Valentine, S.J., Wyttenbach, T.: Recommendations for reporting ion mobility Mass Spectrometry measurements. Mass Spectrom. Rev. 38, 291–320 (2019). https://doi.org/10.1002/mas.21585

90. Jurneczko, E., Barran, P.E.: How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. The Analyst. 136, 20–28 (2010). https://doi.org/10.1039/c0an00373e

91. Scarff, C.A., Thalassinos, K., Hilton, G.R., Scrivens, J.H.: Travelling wave ion mobility mass spectrometry studies of protein structure: biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements. Rapid Commun. Mass Spectrom. 22, 3297–3304 (2008). https://doi.org/10.1002/rcm.3737

92. Eldrid, C., Ujma, J., Kalfas, S., Tomczyk, N., Giles, K., Morris, M., Thalassinos, K.: Gas Phase Stability of Protein Ions in a Cyclic Ion Mobility Spectrometry Traveling Wave Device. Anal. Chem. 91, 7554–7561 (2019). https://doi.org/10.1021/acs.analchem.8b05641

93. Mesleh, M.F., Hunter, J.M., Shvartsburg, A.A., Schatz, G.C., Jarrold, M.F.: Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential. J. Phys. Chem. A. 101, 968–968 (1997). https://doi.org/10.1021/jp963709u

94. Bleiholder, C., Wyttenbach, T., Bowers, M.T.: A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (I). Method. Int. J. Mass Spectrom. 308, 1–10 (2011). https://doi.org/10.1016/j.ijms.2011.06.014

95. Tian, Y., Han, L., Buckner, A.C., Ruotolo, B.T.: Collision Induced Unfolding of Intact Antibodies: Rapid Characterization of Disulfide Bonding Patterns, Glycosylation, and Structures. Anal. Chem. 87, 11509–11515 (2015). https://doi.org/10.1021/acs.analchem.5b03291

96. Freeke, J., Robinson, C.V., Ruotolo, B.T.: Residual counter ions can stabilise a large protein complex in the gas phase. Int. J. Mass Spectrom. 298, 91–98 (2010). https://doi.org/10.1016/j.ijms.2009.08.001

97. Beveridge, R., Migas, L.G., Payne, K.A.P., Scrutton, N.S., Leys, D., Barran, P.E.: Mass spectrometry locates local and allosteric conformational changes that occur on cofactor binding. Nat. Commun. 7, 12163 (2016). https://doi.org/10.1038/ncomms12163

98. Allison, T.M., Reading, E., Liko, I., Baldwin, A.J., Laganowsky, A., Robinson, C.V.: Quantifying the stabilizing effects of protein–ligand interactions in the gas phase. Nat. Commun. 6, 1–10 (2015). https://doi.org/10.1038/ncomms9551

99. Eschweiler, J.D., Rabuck-Gibbons, J.N., Tian, Y., Ruotolo, B.T.: CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions. Anal. Chem. 87, 11516–11522 (2015). https://doi.org/10.1021/acs.analchem.5b03292

100. Watanabe, Y., Vasiljevic, S., Allen, J.D., Seabright, G.E., Duyvesteyn, H.M.E., Doores, K.J., Crispin, M., Struwe, W.B.: Signature of Antibody Domain Exchange by Native Mass Spectrometry and Collision-Induced Unfolding. Anal. Chem. 90, 7325–7331 (2018). https://doi.org/10.1021/acs.analchem.8b00573

101. Hvidt, A., Linderstrøm-Lang, K.: Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. Biochim. Biophys. Acta. 14, 574–575 (1954)

102. Cornwell, O., Bond, N.J., Radford, S.E., Ashcroft, A.E.: Long-Range Conformational Changes in Monoclonal Antibodies Revealed Using FPOP-LC-MS/MS. Anal. Chem. 91, 15163–15170 (2019). https://doi.org/10.1021/acs.analchem.9b03958

103. Veesler, D., Khayat, R., Krishnamurthy, S., Snijder, J., Huang, R.K., Heck, A.J.R., Anand, G.S., Johnson, J.E.: Architecture of a dsDNA viral capsid in complex with its maturation protease. Struct. Lond. Engl. 1993. 22, 230–237 (2014). https://doi.org/10.1016/j.str.2013.11.007

104. Zhang, H., Gau, B.C., Jones, L.M., Vidavsky, I., Gross, M.L.: Fast photochemical oxidation of proteins for comparing structures of protein-ligand complexes: the calmodulin-peptide model system. Anal. Chem. 83, 311–318 (2011). https://doi.org/10.1021/ac102426d

105. Oganesyan, V.S.: EPR spectroscopy and molecular dynamics modelling: a combined approach to study liquid crystals. Liq. Cryst. 45, 2139–2157 (2018). https://doi.org/10.1080/02678292.2018.1508767

106. Calabrese, A.N., Radford, S.E.: Mass spectrometry-enabled structural biology of membrane proteins. Methods San Diego Calif. 147, 187–205 (2018). https://doi.org/10.1016/j.ymeth.2018.02.020

107. Konermann, L., Stocks, B.B., Pan, Y., Tong, X.: Mass spectrometry combined with oxidative labeling for exploring protein structure and folding. Mass Spectrom. Rev. 29, 651–667 (2010). https://doi.org/10.1002/mas.20256

108. Sinz, A.: Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. Mass Spectrom. Rev. 25, 663–682 (2006). https://doi.org/10.1002/mas.20082

109. Holding, A.N.: XL-MS: Protein cross-linking coupled with mass spectrometry. Methods. 89, 54–63 (2015). https://doi.org/10.1016/j.ymeth.2015.06.010

110. Cornwell, O., Radford, S.E., Ashcroft, A.E., Ault, J.R.: Comparing Hydrogen Deuterium Exchange and Fast Photochemical Oxidation of Proteins: a Structural Characterisation of Wild-Type and ΔN6 β2-Microglobulin. J. Am. Soc. Mass Spectrom. 29, 2413–2426 (2018). https://doi.org/10.1007/s13361-018-2067-y

111. Zhang, Q., Chen, J., Kuwajima, K., Zhang, H.-M., Xian, F., Young, N.L., Marshall, A.G.: Nucleotide-induced conformational changes of tetradecameric GroEL mapped by H/D exchange monitored by FT-ICR mass spectrometry. Sci. Rep. 3, 1–7 (2013). https://doi.org/10.1038/srep01247

112. Houde, D., Berkowitz, S.A., Engen, J.R.: The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. J. Pharm. Sci. 100, 2071–2086 (2011). https://doi.org/10.1002/jps.22432

113. Li, J., Wei, H., R. Krystek, S., Bond, D., M. Brender, T., Cohen, D., Feiner, J., Hamacher, N., Harshman, J., Huang, R., Julien, S., Lin, Z., Moore, K., Mueller, L., Noriega, C., Sejwal, P., Sheppard, P., Stevens, B., Chen, G., Schneeweis, L.: Hydrogen/Deuterium Exchange, Fast Photochemical Oxidation of Proteins (FPOP) Mass Spectrometry and Alanine Shave Mutagenesis Effectively Map the Energetic Epitope of an Antibody for Interleukin-23. Anal. Chem. 89, 2250–2258 (2017). https://doi.org/10.1021/acs.analchem.6b03058

114. Hvidt, A., Linderstrøm-Lang, K.: The kinetics of the deuterium exchange of insulin with D2O. An amendment. Biochim. Biophys. Acta. 16, 168–169 (1955). https://doi.org/10.1016/0006-3002(55)90200-6

115. Wagner, G., Wüthrich, K.: Amide protein exchange and surface conformation of the basic pancreatic trypsin inhibitor in solution. Studies with two-dimensional nuclear magnetic resonance. J. Mol. Biol. 160, 343–361 (1982). https://doi.org/10.1016/0022-2836(82)90180-2

116. Wand, A.J., Englander, S.W.: Protein complexes studied by NMR spectroscopy. Curr. Opin. Biotechnol. 7, 403–408 (1996)

117. Rosa, J.J., Richards, F.M.: An experimental procedure for increasing the structural resolution of chemical hydrogen-exchange measurements on proteins: Application to ribonuclease S peptide. J. Mol. Biol. 133, 399–416 (1979). https://doi.org/10.1016/0022-2836(79)90400-5

118. Katta, V., Chait, B.T., Carr, S.: Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. Rapid Commun. Mass Spectrom. 5, 214–217 (1991). https://doi.org/10.1002/rcm.1290050415

119. Milne, J.S., Mayne, L., Roder, H., Wand, A.J., Englander, S.W.: Determinants of protein hydrogen exchange studied in equine cytochrome c. Protein Sci. 7, 739–745 (1998). https://doi.org/10.1002/pro.5560070323

120. Bai, Y., Milne, J.S., Mayne, L., Englander, S.W.: Primary structure effects on peptide group hydrogen exchange. Proteins. 17, 75–86 (1993). https://doi.org/10.1002/prot.340170110

121. Englander, S.W., Poulsen, A.: Hydrogen-tritium exchange of the random chain polypeptide. Biopolymers. 7, 379–393 (1969). https://doi.org/10.1002/bip.1969.360070309

122. Wales, T.E., Fadgen, K.E., Eggertson, M.J., Engen, J.R.: Subzero Celsius separations in three-zone temperature controlled hydrogen deuterium exchange mass spectrometry. J. Chromatogr. A. 1523, 275–282 (2017). https://doi.org/10.1016/j.chroma.2017.05.067

123. Walters, B.T., Ricciuti, A., Mayne, L., Englander, S.W.: Minimizing back exchange in the hydrogen exchange - mass spectrometry experiment. J. Am. Soc. Mass Spectrom. 23, 2132–2139 (2012). https://doi.org/10.1007/s13361-012-0476-x

124. Saphire, E.O., Parren, P.W., Pantophlet, R., Zwick, M.B., Morris, G.M., Rudd, P.M., Dwek, R.A., Stanfield, R.L., Burton, D.R., Wilson, I.A.: Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. Science. 293, 1155–1159 (2001). https://doi.org/10.1126/science.1061692

125. Zhang, Z., Smith, D.L.: Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. Protein Sci. Publ. Protein Soc. 2, 522–531 (1993). https://doi.org/10.1002/pro.5560020404

126. Ferraro, D.M., Lazo, N.D., Robertson, A.D.: EX1 Hydrogen Exchange and Protein Folding. Biochemistry. 43, 587–594 (2004). https://doi.org/10.1021/bi035943y

127. Wales, T.E., Fadgen, K.E., Gerhardt, G.C., Engen, J.R.: High-Speed and High-Resolution UPLC Separation at Zero Degrees Celsius. Anal. Chem. 80, 6815–6820 (2008). https://doi.org/10.1021/ac8008862

128. Zhang, J., Banks, D.D., He, F., Treuheit, M.J., Becker, G.W.: Effects of sucrose and benzyl alcohol on GCSF conformational dynamics revealed by hydrogen deuterium exchange mass spectrometry. J. Pharm. Sci. 104, 1592–1600 (2015). https://doi.org/10.1002/jps.24384

129. Burkitt, W., Domann, P., O'Connor, G.: Conformational changes in oxidatively stressed monoclonal antibodies studied by hydrogen exchange mass spectrometry. Protein Sci. 19, 826–835 (2010). https://doi.org/10.1002/pro.362

130. Houde, D., Arndt, J., Domeier, W., Berkowitz, S., Engen, J.R.: Characterization of IgG1 Conformation and Conformational Dynamics by Hydrogen/Deuterium Exchange Mass Spectrometry. Anal. Chem. (2009). https://doi.org/10.1021/ac9009287

131. Bereszczak, J.Z., Rose, R.J., van Duijn, E., Watts, N.R., Wingfield, P.T., Steven, A.C., Heck, A.J.R.: Epitope-distal effects accompany the binding of two distinct antibodies to hepatitis B virus capsids. J. Am. Chem. Soc. 135, 6504–6512 (2013). https://doi.org/10.1021/ja402023x

132. Lim, X.-X., Chandramohan, A., Lim, X.Y.E., Bag, N., Sharma, K.K., Wirawan, M., Wohland, T., Lok, S.-M., Anand, G.S.: Conformational changes in intact dengue virus reveal serotype-specific expansion. Nat. Commun. 8, 14339 (2017). https://doi.org/10.1038/ncomms14339

133. Feng, L., Orlando, R., Prestegard, J.H.: Amide proton back-exchange in deuterated peptides: applications to MS and NMR analyses. Anal. Chem. 78, 6885–6892 (2006). https://doi.org/10.1021/ac060948x

134. Moorthy, B.S., Schultz, S.G., Kim, S.G., Topp, E.M.: Predicting protein aggregation during storage in lyophilized solids using solid state amide hydrogen/deuterium exchange with mass spectrometric analysis (ssHDX-MS). Mol. Pharm. 11, 1869–1879 (2014). https://doi.org/10.1021/mp500005v

135. Bobst, C.E., Abzalimov, R.R., Houde, D., Kloczewiak, M., Mhatre, R., Berkowitz, S.A., Kaltashov, I.A.: Detection and characterization of altered conformations of protein pharmaceuticals using complementary mass spectrometry-based approaches. Anal. Chem. 80, 7473–7481 (2008). https://doi.org/10.1021/ac801214x

136. Rajabi, K.: Time-Resolved Pulsed Hydrogen/Deuterium Exchange Mass Spectrometry Probes Gaseous Proteins Structural Kinetics. J. Am. Soc. Mass Spectrom. 26, 71–82 (2015). https://doi.org/10.1021/jasms.8b04868

137. Chalmers, M.J., Pascal, B.D., Willis, S., Zhang, J., Iturria, S.J., Dodge, J.A., Griffin, P.R.: Methods for the Analysis of High Precision Differential Hydrogen Deuterium Exchange Data. Int. J. Mass Spectrom. 302, 59–68 (2011). https://doi.org/10.1016/j.ijms.2010.08.002

138. Burkitt, W., O'Connor, G.: Assessment of the repeatability and reproducibility of hydrogen/deuterium exchange mass spectrometry measurements. Rapid Commun. Mass Spectrom. RCM. 22, 3893–3901 (2008). https://doi.org/10.1002/rcm.3794

139. Cummins, D.J., Espada, A., Novick, S.J., Molina-Martin, M., Stites, R.E., Espinosa, J.F., Broughton, H., Goswami, D., Pascal, B.D., Dodge, J.A., Chalmers, M.J., Griffin, P.R.: Two-Site Evaluation of the Repeatability and Precision of an Automated Dual-Column Hydrogen/Deuterium Exchange Mass Spectrometry Platform. Anal. Chem. 88, 6607–6614 (2016). https://doi.org/10.1021/acs.analchem.6b01650

140. Espada, A., Haro, R., Castañon, J., Sayago, C., Perez-Cozar, F., Cano, L., Redero, P., Molina-Martin, M., Broughton, H., Stites, R.E., Pascal, B.D., Griffin, P.R., Dodge, J.A., Chalmers, M.J.: A Decoupled Automation Platform for Hydrogen/Deuterium Exchange Mass Spectrometry Experiments. J. Am. Soc. Mass Spectrom. 30, 2580–2583 (2019). https://doi.org/10.1007/s13361-019-02331-2

141. Zhang, H.-M., Kazazic, S., Schaub, T., Tipton, J., Emmett, M., Marshall, A.: Enhanced Digestion Efficiency, Peptide Ionization Efficiency, and Sequence Resolution for Protein Hydrogen/Deuterium Exchange Monitored by Fourier

Transform Ion Cyclotron Resonance Mass Spectrometry. Anal. Chem. (2008). https://doi.org/10.1021/ac801417d

142. Ahn, J., Cao, M.-J., Yu, Y.Q., Engen, J.R.: Accessing the reproducibility and specificity of pepsin and other aspartic proteases. Biochim. Biophys. Acta - Proteins Proteomics. 1834, 1222–1229 (2013). https://doi.org/10.1016/j.bbapap.2012.10.003

143. Rey, M., Yang, M., Burns, K.M., Yu, Y., Lees-Miller, S.P., Schriemer, D.C.: Nepenthesin from Monkey Cups for Hydrogen/Deuterium Exchange Mass Spectrometry. Mol. Cell. Proteomics MCP. 12, 464–472 (2013). https://doi.org/10.1074/mcp.M112.025221

144. Ahn, J., Jung, M.C., Wyndham, K., Yu, Y.Q., Engen, J.R.: Pepsin Immobilized on High-Strength Hybrid Particles for Continuous Flow Online Digestion at 10 000 psi. Anal. Chem. 84, 7256–7262 (2012). https://doi.org/10.1021/ac301749h

145. Hamuro, Y., Coales, S.J., Molnar, K.S., Tuske, S.J., Morrow, J.A.: Specificity of immobilized porcine pepsin in H/D exchange compatible conditions. Rapid Commun. Mass Spectrom. RCM. 22, 1041–1046 (2008). https://doi.org/10.1002/rcm.3467

146. Hamuro, Y., Coales, S.J.: Optimization of Feasibility Stage for Hydrogen/Deuterium Exchange Mass Spectrometry. J. Am. Soc. Mass Spectrom. 29, 623–629 (2018). https://doi.org/10.1007/s13361-017-1860-3

147. Wang, L., Pan, H., Smith, D.L.: Hydrogen Exchange-Mass Spectrometry: Optimization of Digestion Conditions. Mol. Cell. Proteomics. 1, 132–138 (2001). https://doi.org/10.1074/mcp.M100009-MCP200

148. Busby, S., Chalmers, M., Griffin, P.: Improving digestion efficiency under H/D exchange conditions with activated pepsinogen coupled columns. Int. J. Mass Spectrom. - INT J MASS SPECTROM. 259, 130–139 (2007). https://doi.org/10.1016/j.ijms.2006.08.006

149. Ehring, H.: Hydrogen exchange/electrospray ionization mass spectrometry studies of structural features of proteins and protein/protein interactions. Anal. Biochem. 267, 252—259 (1999). https://doi.org/10.1006/abio.1998.3000

150. Majumdar, R., Middaugh, C.R., Weis, D.D., Volkin, D.B.: Hydrogen-Deuterium Exchange Mass Spectrometry as an Emerging Analytical Tool for Stabilization and Formulation Development of Therapeutic Monoclonal Antibodies. J. Pharm. Sci. epub ahead of print, (2014). https://doi.org/10.1002/jps.24224

151. Fang, X., Zhang, W.-W.: Affinity separation and enrichment methods in proteomic analysis. J. Proteomics. 71, 284–303 (2008). https://doi.org/10.1016/j.jprot.2008.06.011

152. Fang, J., Rand, K.D., Beuning, P.J., Engen, J.R.: False EX1 signatures caused by sample carryover during HX MS analyses. Int. J. Mass Spectrom. 302, 19–25 (2011). https://doi.org/10.1016/j.ijms.2010.06.039

153. Tito, P., Nettleton, E.J., Robinson, C.V.: Dissecting the hydrogen exchange properties of insulin under amyloid fibril forming conditions: a site-specific investigation by mass spectrometry. J. Mol. Biol. 303, 267–278 (2000). https://doi.org/10.1006/jmbi.2000.4142

154. Valeja, S.G., Emmett, M.R., Marshall, A.G.: Polar Aprotic Modifiers for Chromatographic Separation and Back-Exchange Reduction for Protein Hydrogen/Deuterium Exchange Monitored by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. J. Am. Soc. Mass Spectrom. 23, 699–707 (2012). https://doi.org/10.1007/s13361-011-0329-z

155. Keppel, T.R., Jacques, M.E., Young, R.W., Ratzlaff, K.L., Weis, D.D.: An efficient and inexpensive refrigerated LC system for H/D exchange mass spectrometry. J. Am. Soc. Mass Spectrom. 22, 1472–1476 (2011). https://doi.org/10.1007/s13361-011-0152-6

156. Morgan, C.R., Hebling, C.M., Rand, K.D., Stafford, D.W., Jorgenson, J.W., Engen, J.R.: Conformational transitions in the membrane scaffold protein of phospholipid bilayer nanodiscs. Mol. Cell. Proteomics MCP. 10, M111.010876 (2011). https://doi.org/10.1074/mcp.M111.010876

157. Iacob, R.E., Murphy, J.P., Engen, J.R.: Ion mobility adds an additional dimension to mass spectrometric analysis of solution-phase hydrogen/deuterium exchange. Rapid Commun. Mass Spectrom. 22, 2898–2904 (2008). https://doi.org/10.1002/rcm.3688

158. Donohoe, G.C., Arndt, J.R., Valentine, S.J.: Online Deuterium Hydrogen Exchange and Protein Digestion Coupled with Ion Mobility Spectrometry and Tandem Mass Spectrometry. Anal. Chem. 87, 5247–5254 (2015). https://doi.org/10.1021/acs.analchem.5b00277

159. Cryar, A., Groves, K., Quaglia, M.: Online Hydrogen-Deuterium Exchange Traveling Wave Ion Mobility Mass Spectrometry (HDX-IM-MS): a Systematic Evaluation. J. Am. Soc. Mass Spectrom. 28, 1192–1202 (2017). https://doi.org/10.1007/s13361-017-1633-z

160. Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G.A., Ma, B.: PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. Mol. Cell. Proteomics MCP. 11, M111.010587 (2012). https://doi.org/10.1074/mcp.M111.010587

161. Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S.: Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 20, 3551–67 (1999)

162. Bateman, R.H., Carruthers, R., Hoyes, J.B., Jones, C., Langridge, J.I., Millar, A., Vissers, J.P.C.: A novel precursor ion discovery method on a hybrid quadrupole orthogonal acceleration time-of-flight (Q-TOF) mass spectrometer for studying protein phosphorylation. J. Am. Soc. Mass Spectrom. 13, 792–803 (2002). https://doi.org/10.1016/S1044-0305(02)00420-8

163. Johnson, R.S., Krylov, D., Walsh, K.A.: Proton mobility within electrosprayed peptide ions. J. Mass Spectrom. 30, 386–387 (1995). https://doi.org/10.1002/jms.1190300224

164. Rand, K.D., Pringle, S.D., Morris, M., Engen, J.R., Brown, J.M.: ETD in a Traveling Wave Ion Guide at Tuned Z-Spray Ion Source Conditions Allows for Site-Specific Hydrogen/Deuterium Exchange Measurements. J. Am. Soc. Mass Spectrom. 22, 1784–1793 (2011). https://doi.org/10.1007/s13361-011-0196-7

165. Pan, J., Han, J., Borchers, C.H., Konermann, L.: Hydrogen/deuterium exchange mass spectrometry with top-down electron capture dissociation for

characterizing structural transitions of a 17 kDa protein. J. Am. Chem. Soc. 131, 12801–12808 (2009). https://doi.org/10.1021/ja904379w

166. Landgraf, R.R., Chalmers, M.J., Griffin, P.R.: Automated Hydrogen/Deuterium Exchange Electron Transfer Dissociation High Resolution Mass Spectrometry Measured at Single-Amide Resolution. J. Am. Soc. Mass Spectrom. 23, 301–309 (2012). https://doi.org/10.1007/s13361-011-0298-2

167. Thévenon-Emeric, G., Kozlowski, J., Zhang, Z., Smith, D.L.: Determination of amide hydrogen exchange rates in peptides by mass spectrometry. Anal. Chem. 64, 2456–2458 (1992). https://doi.org/10.1021/ac00044a027

168. Hageman, T.S., Weis, D.D.: Reliable Identification of Significant Differences in Differential Hydrogen Exchange-Mass Spectrometry Measurements Using a Hybrid Significance Testing Approach. Anal. Chem. 91, 8008–8016 (2019). https://doi.org/10.1021/acs.analchem.9b01325

169. Hageman, T.S., Weis, D.D.: A Structural Variant Approach for Establishing a Detection Limit in Differential Hydrogen Exchange-Mass Spectrometry Measurements. Anal. Chem. 91, 8017–8024 (2019). https://doi.org/10.1021/acs.analchem.9b01326

170. Chalmers, M.J., Busby, S.A., Pascal, B.D., He, Y., Hendrickson, C.L., Marshall, A.G., Griffin, P.R.: Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. Anal. Chem. 78, 1005–1014 (2006). https://doi.org/10.1021/ac051294f

171. Pascal, B.D., Chalmers, M.J., Busby, S.A., Griffin, P.R.: HD desktop: An integrated platform for the analysis and visualization of H/D exchange data. J. Am. Soc. Mass Spectrom. 20, 601–610 (2009). https://doi.org/10.1016/j.jasms.2008.11.019

172. Pascal, B.D., Willis, S., Lauer, J.L., Landgraf, R.R., West, G.M., Marciano, D., Novick, S., Goswami, D., Chalmers, M.J., Griffin, P.R.: HDX workbench: software for the analysis of H/D exchange MS data. J. Am. Soc. Mass Spectrom. 23, 1512–1521 (2012). https://doi.org/10.1007/s13361-012-0419-6

173. Kavan, D., Man, P.: MSTools—Web based application for visualization and presentation of HXMS data. Int. J. Mass Spectrom. 302, 53–58 (2011). https://doi.org/10.1016/j.ijms.2010.07.030

174. Guttman, M., Weis, D.D., Engen, J.R., Lee, K.K.: Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. J. Am. Soc. Mass Spectrom. 24, 1906–1912 (2013). https://doi.org/10.1007/s13361-013-0727-5

175. Mayne, L., Kan, Z.-Y., Chetty, P.S., Ricciuti, A., Walters, B.T., Englander, S.W.: Many overlapping peptides for protein hydrogen exchange experiments by the fragment separation-mass spectrometry method. J. Am. Soc. Mass Spectrom. 22, 1898–1905 (2011). https://doi.org/10.1007/s13361-011-0235-4

176. Iacob, R.E., Engen, J.R.: Hydrogen exchange mass spectrometry: are we out of the quicksand? J. Am. Soc. Mass Spectrom. 23, 1003–1010 (2012). https://doi.org/10.1007/s13361-012-0377-z

177. Zhang, H.-M., Li, C., Lei, M., Lundin, V., Lee, H.Y., Ninonuevo, M., Lin, K., Han, G., Sandoval, W., Lei, D., Ren, G., Zhang, J., Liu, H.: Structural and Functional Characterization of a Hole-Hole Homodimer Variant in a “Knob-

Into-Hole" Bispecific Antibody. Anal. Chem. 89, 13494–13501 (2017). https://doi.org/10.1021/acs.analchem.7b03830

178. Tullius, T.D., Dombroski, B.A.: Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to lambda repressor and Cro protein. Proc. Natl. Acad. Sci. U. S. A. 83, 5469–5473 (1986)

179. Sheshberadaran, H., Payne, L.G.: Protein antigen-monoclonal antibody contact sites investigated by limited proteolysis of monoclonal antibody-bound antigen: protein "footprinting". Proc. Natl. Acad. Sci. U. S. A. 85, 1–5 (1988)

180. Brenowitz, M., Chance, M.R., Dhavan, G., Takamoto, K.: Probing the structural dynamics of nucleic acids by quantitative time-resolved and equilibrium hydroxyl radical "footprinting." Curr. Opin. Struct. Biol. 12, 648–653 (2002)

181. Watson, C., Janik, I., Zhuang, T., Charvátová, O., Woods, R.J., Sharp, J.S.: Pulsed Electron Beam Water Radiolysis for Sub-Microsecond Hydroxyl Radical Protein Footprinting. Anal. Chem. 81, 2496–2505 (2009). https://doi.org/10.1021/ac802252y

182. Hambly, D.M., Gross, M.L.: Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. J. Am. Soc. Mass Spectrom. 16, 2057–2063 (2005). https://doi.org/10.1016/j.jasms.2005.09.008

183. Hambly, D.M., Gross, M.L.: Laser flash photochemical oxidation to locate heme binding and conformational changes in myoglobin. Presented at the (2007)

184. Vahidi, S., Stocks, B.B., Liaghati-Mobarhan, Y., Konermann, L.: Mapping pH-Induced Protein Structural Changes Under Equilibrium Conditions by Pulsed Oxidative Labeling and Mass Spectrometry. Anal. Chem. 84, 9124–9130 (2012). https://doi.org/10.1021/ac302393g

185. Zheng, X., Wintrode, P.L., Chance, M.R.: Complementary structural mass spectrometry techniques reveal local dynamics in functionally important regions of a metastable serpin. Struct. Lond. Engl. 1993. 16, 38–51 (2008). https://doi.org/10.1016/j.str.2007.10.019

186. Watkinson, T.G., Calabrese, A.N., Ault, J.R., Radford, S.E., Ashcroft, A.E.: FPOP-LC-MS/MS Suggests Differences in Interaction Sites of Amphipols and Detergents with Outer Membrane Proteins. J. Am. Soc. Mass Spectrom. 28, 50–55 (2017). https://doi.org/10.1007/s13361-016-1421-1

187. Liu, X.R., Rempel, D.L., Gross, M.L.: Composite Conformational Changes of Signaling Proteins upon Ligand Binding Revealed by a Single Approach: Calcium-Calmodulin Study. Anal. Chem. 91, 12560–12567 (2019). https://doi.org/10.1021/acs.analchem.9b03491

188. Jones, L.M., B. Sperry, J., A. Carroll, J., Gross, M.L.: Fast Photochemical Oxidation of Proteins for Epitope Mapping. Anal. Chem. 83, 7657–7661 (2011). https://doi.org/10.1021/ac2007366

189. Espino, J.A., Mali, V.S., Jones, L.M.: In Cell Footprinting Coupled with Mass Spectrometry for the Structural Analysis of Proteins in Live Cells. Anal. Chem. 87, 7971–7978 (2015). https://doi.org/10.1021/acs.analchem.5b01888

190. Takamoto, K., Chance, M.R.: Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes. Annu. Rev.

Biophys. Biomol. Struct. 35, 251–276 (2006). https://doi.org/10.1146/annurev.biophys.35.040405.102050

191. Xu, G., Chance, M.R.: Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. Chem. Rev. 107, 3514–3543 (2007). https://doi.org/10.1021/cr0682047

192. Maleknia, S.D., Brenowitz, M., Chance, M.R.: Millisecond radiolytic modification of peptides by synchrotron X-rays identified by mass spectrometry. Anal. Chem. 71, 3965–3973 (1999). https://doi.org/10.1021/ac990500e

193. Goldsmith, S.C., Guan, J.-Q., Almo, S.C., Chance, M.R.: Synchrotron Protein Footprinting: A Technique to Investigate Protein-Protein Interactions. J. Biomol. Struct. Dyn. 19, 405–418 (2001). https://doi.org/10.1080/07391102.2001.10506750

194. Xie, B., Sood, A., Woods, R.J., Sharp, J.S.: Quantitative Protein Topography Measurements by High Resolution Hydroxyl Radical Protein Footprinting Enable Accurate Molecular Model Selection. Sci. Rep. 7, 1–11 (2017). https://doi.org/10.1038/s41598-017-04689-3

195. Limpikirati, P., Liu, T., Vachet, R.W.: Covalent labeling-mass spectrometry with non-specific reagents for studying protein structure and interactions. Methods San Diego Calif. 144, 79–93 (2018). https://doi.org/10.1016/j.ymeth.2018.04.002

196. Gau, B.C., Sharp, J.S., Rempel, D.L., Gross, M.L.: Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. Anal. Chem. 81, 6563–6571 (2009). https://doi.org/10.1021/ac901054w

197. Naganathan, A.N., Muñoz, V.: Scaling of folding times with protein size. J. Am. Chem. Soc. 127, 480–481 (2005). https://doi.org/10.1021/ja044449u

198. Niu, B., Zhang, H., Giblin, D., Rempel, D.L., Gross, M.L.: Dosimetry Determines the Initial OH Radical Concentration in Fast Photochemical Oxidation of Proteins (FPOP). J. Am. Soc. Mass Spectrom. 26, 843–846 (2015). https://doi.org/10.1007/s13361-015-1087-0

199. Zhang, B., Cheng, M., Rempel, D., Gross, M.L.: Implementing fast photochemical oxidation of proteins (FPOP) as a footprinting approach to solve diverse problems in structural biology. Methods San Diego Calif. 144, 94–103 (2018). https://doi.org/10.1016/j.ymeth.2018.05.016

200. Vahidi, S., Konermann, L.: Probing the Time Scale of FPOP (Fast Photochemical Oxidation of Proteins): Radical Reactions Extend Over Tens of Milliseconds. J. Am. Soc. Mass Spectrom. 27, 1156–1164 (2016). https://doi.org/10.1007/s13361-016-1389-x

201. Zhang, Y., Rempel, D.L., Zhang, H., Gross, M.L.: An improved fast photochemical oxidation of proteins (FPOP) platform for protein therapeutics. J. Am. Soc. Mass Spectrom. 26, 526–529 (2015). https://doi.org/10.1007/s13361-014-1055-0

202. Yan, Y., Chen, G., Wei, H., Huang, R.Y.-C., Mo, J., Rempel, D.L., Tymiak, A.A., Gross, M.L.: Fast photochemical oxidation of proteins (FPOP) maps the epitope of EGFR binding to adnectin. J. Am. Soc. Mass Spectrom. 25, 2084–2092 (2014). https://doi.org/10.1007/s13361-014-0993-x

203. Hamuro, Y., Coales, S.J., Southern, M.R., Nemeth-Cawley, J.F., Stranz, D.D., Griffin, P.R.: Rapid analysis of protein structure and dynamics by hydrogen/deuterium exchange mass spectrometry. J. Biomol. Tech. JBT. 14, 171–182 (2003)

204. Seger, S.T., Breinholt, J., Faber, J.H., Andersen, M.D., Wiberg, C., Schjødt, C.B., Rand, K.D.: Probing the Conformational and Functional Consequences of Disulfide Bond Engineering in Growth Hormone by Hydrogen–Deuterium Exchange Mass Spectrometry Coupled to Electron Transfer Dissociation. Anal. Chem. 87, 5973–5980 (2015). https://doi.org/10.1021/ac504782v

205. Horn, J.R., Kraybill, B., Petro, E.J., Coales, S.J., Morrow, J.A., Hamuro, Y., Kossiakoff, A.A.: The role of protein dynamics in increasing binding affinity for an engineered protein-protein interaction established by H/D exchange mass spectrometry. Biochemistry. 45, 8488–8498 (2006). https://doi.org/10.1021/bi0604328

206. Pritchard, C., O'Connor, G., Ashcroft, A.E.: The Role of Ion Mobility Spectrometry–Mass Spectrometry in the Analysis of Protein Reference Standards. Anal. Chem. 85, 7205–7212 (2013). https://doi.org/10.1021/ac400927s

207. de Vos, A.M., Ultsch, M., Kossiakoff, A.A.: Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. Science. 255, 306–312 (1992)

208. Somers, W., Ultsch, M., De Vos, A.M., Kossiakoff, A.A.: The X-ray structure of a growth hormone-prolactin receptor complex. Nature. 372, 478–481 (1994). https://doi.org/10.1038/372478a0

209. Kasimova, M.R., Kristensen, S.M., Howe, P.W.A., Christensen, T., Matthiesen, F., Petersen, J., Sørensen, H.H., Led, J.J.: NMR studies of the backbone flexibility and structure of human growth hormone: a comparison of high and low pH conformations. J. Mol. Biol. 318, 679–695 (2002). https://doi.org/10.1016/S0022-2836(02)00137-7

210. Yang, T.H., Cleland, J.L., Lam, X., Meyer, J.D., Jones, L.S., Randolph, T.W., Manning, M.C., Carpenter, J.F.: Effect of zinc binding and precipitation on structures of recombinant human growth hormone and nerve growth factor. J. Pharm. Sci. 89, 1480–1485 (2000)

211. Khameneh, B., Jaafari, M.R., Hassanzadeh-Khayyat, M., Varasteh, A., Chamani, J., Iranshahi, M., Mohammadpanah, H., Abnous, K., Saberi, M.R.: Preparation, characterization and molecular modeling of PEGylated human growth hormone with agonist activity. Int. J. Biol. Macromol. 80, 400–409 (2015). https://doi.org/10.1016/j.ijbiomac.2015.06.037

212. Huising, M.O., Kruiswijk, C.P., Flik, G.: Phylogeny and evolution of class-I helical cytokines. J. Endocrinol. 189, 1–25 (2006). https://doi.org/10.1677/joe.1.06591

213. Cunningham, B.C., Mulkerrin, M.G., Wells, J.A.: Dimerization of human growth hormone by zinc. Science. 253, 545–548 (1991). https://doi.org/10.1126/science.1907025

214. Cunningham, B.C., Bass, S., Fuh, G., Wells, J.A.: Zinc mediation of the binding of human growth hormone to the human prolactin receptor. Science. 250, 1709–1712 (1990). https://doi.org/10.1126/science.2270485

215. Greenan, J., Lorenson, M.Y., Conconi, M.V., Walker, A.M.: Alterations in in situ prolactin secretory granule morphology and immunoactivity by thiols and divalent cations. Endocrinology. 126, 512–518 (1990). https://doi.org/10.1210/endo-126-1-512

216. Chantalat, L., Jones, N.D., Korber, F., Navaza, J., Pavlovsky, A.G.: The crystal structure of wild-type growth-hormone at 2.5 angstrom resolution. Protein PeptLett. 2, 333–340 (1995)

217. Marriott, J., O'Connor, G., Parks, H.: Study of Measurement Service and Comparison Needs for an International Measurement Infrastructure for the Biosciences and Biotechnology: Input for the BIPM Work Programme Date Assessed 29 May 2015, www.bipm.org.

218. International Conference on Harmonization: ICH Topic Q5E: Comparability of Biotechnological/Biological Products Subject to Changes in their Manufacturing Process, (2005)

219. Rey, M., Sarpe, V., Burns, K., Buse, J., Baker, C.A.H., van Dijk, M., Wordeman, L., Bonvin, A.M.J.J., Schriemer, D.C.: Mass Spec Studio for Integrative Structural Biology. Struct. Lond. Engl. 1993. 22, 1538–1548 (2014). https://doi.org/10.1016/j.str.2014.08.013

220. Hourdel, V., Volant, S., O'Brien, D.P., Chenal, A., Chamot-Rooke, J., Dillies, M.-A., Brier, S.: MEMHDX: an interactive tool to expedite the statistical validation and visualization of large HDX-MS datasets. Bioinforma. Oxf. Engl. 32, 3413–3419 (2016). https://doi.org/10.1093/bioinformatics/btw420

221. Ahn, J.: Local hydrogen deuterium exchange mass spectrometry: from pressurized online digestion to pepsin proteolysis, http://hdl.handle.net/2047/d20003044, (2013)

222. Wang, L., Smith, D.L.: Downsizing improves sensitivity 100-fold for hydrogen exchange-mass spectrometry. Anal. Biochem. 314, 46–53 (2003). https://doi.org/10.1016/S0003-2697(02)00620-6

223. Hudgens, J.W., Huang, R.Y.-C., D'Ambro, E.: Method Validation and Standards in Hydrogen Exchange Mass Spectrometry. In: Hydrogen Exchange Mass Spectrometry of Proteins. pp. 55–72. Wiley-Blackwell (2016)

224. Groves, K., Cryar, A., Cowen, S., Ashcroft, A.E., Quaglia, M.: Mass Spectrometry Characterization of Higher Order Structural Changes Associated with the Fc-glycan Structure of the NISTmAb Reference Material, RM 8761. J. Am. Soc. Mass Spectrom. (2020). https://doi.org/10.1021/jasms.9b00022

225. Wicar, S., Mulkerrin, M.G., Bathory, G., Khundkar, L.H., Karger, B.L.: Conformational changes in the reversed phase liquid chromatography of recombinant human growth hormone as a function of organic solvent: the molten globule state. Anal. Chem. 66, 3908–3915 (1994)

226. Gomez-Orellana, I., Variano, B., Miura-Fraboni, J., Milstein, S., Paton, D.R.: Thermodynamic characterization of an intermediate state of human growth hormone. Protein Sci. 7, 1352–1358 (1998). https://doi.org/10.1002/pro.5560070611

227. Shelimov, K.B., Clemmer, D.E., Hudgins, R.R., Jarrold, M.F.: Protein structure in vacuo: gas-phase conformations of BPTI and cytochrome c. J. Am. Chem. Soc. 119, 2240–2248 (1997). https://doi.org/10.1021/ja9619059

228. Valentine, S.J., Anderson, J.G., Ellington, A.D., Clemmer, D.E.: Disulfide-Intact and -Reduced Lysozyme in the Gas Phase: Conformations and Pathways of Folding and Unfolding. J. Phys. Chem. B. 101, 3891–3900 (1997). https://doi.org/10.1021/jp970217o

229. Valentine, S.J., Counterman, A.E., Clemmer, D.E.: Conformer-dependent proton-transfer reactions of ubiquitin ions. J. Am. Soc. Mass Spectrom. 8, 954–961 (1997). https://doi.org/10.1016/S1044-0305(97)00085-8

230. Richardson, K., Langridge, D., Giles, K.: Fundamentals of travelling wave ion mobility revisited: I. Smoothly moving waves. Int. J. Mass Spectrom. 428, 71–80 (2018). https://doi.org/10.1016/j.ijms.2018.03.007

231. Pan, J., Xu, K., Yang, X., Choy, W.-Y., Konermann, L.: Solution-Phase Chelators for Suppressing Nonspecific Protein−Metal Interactions in Electrospray Mass Spectrometry. Anal. Chem. 81, 5008–5015 (2009). https://doi.org/10.1021/ac900423x

232. Dixit, S.M., Polasky, D.A., Ruotolo, B.T.: Collision Induced Unfolding of Isolated Proteins in the Gas Phase: Past, Present, and Future. Curr. Opin. Chem. Biol. 42, 93–100 (2018). https://doi.org/10.1016/j.cbpa.2017.11.010

233. Groves, K., Quaglia, M., Schiel, J., Thalassinos, K., Soloview, Z., Upton, R., Barran, P., Durrant, O., Dalby, P., De Lorenzo, E., Stocks, B.B., Ashton, L.: Interlaboratory comparison of protein higher order structure measurements of human growth hormone reference material and zinc model system. Prep.

234. Köhler, G., Milstein, C.: Continuous cultures of fused cells secreting antibody of predefined specificity. Nature. 256, 495 (1975). https://doi.org/10.1038/256495a0

235. Vlug, A., de BOT, A.J., Brouwer, H., de LANGE, G.G., van LEEUWEN, A.M., de RON, E.A., van EIJK, R.V.W.: Monoclonal Antibodies Against IgG Subclasses. In: Poulik, M.D. (ed.) Protides of the Biological Fluids. pp. 39–47. Elsevier (1989)

236. Alzari, P.M., Lascombe, M.B., Poljak, R.J.: Three-Dimensional Structure of Antibodies. Annu. Rev. Immunol. 6, 555–580 (1988). https://doi.org/10.1146/annurev.iy.06.040188.003011

237. Bork, P., Holm, L., Sander, C.: The immunoglobulin fold. Structural classification, sequence patterns and common core. J. Mol. Biol. 242, 309–320 (1994). https://doi.org/10.1006/jmbi.1994.1582

238. Gallagher, D.T., Karageorgos, I., Hudgens, J.W., Galvin, C.V.: Data on crystal organization in the structure of the Fab fragment from the NIST reference antibody, RM 8671. Data Brief. 16, 29–36 (2018). https://doi.org/10.1016/j.dib.2017.11.013

239. Schiel, J.E., Davis, D.L., Borisov, O.V. eds: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 1. Monoclonal Antibody Therapeutics: Structure, Function, and Regulatory Space. American Chemical Society, Washington, DC (2014)

240. Schiel, J.E., Davis, D.L., Borisov, O.V. eds: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 2. Biopharmaceutical Characterization: The NISTmAb Case Study. American Chemical Society, Washington, DC (2015)

241. Schiel, J.E., Davis, D.L., Borisov, O.V. eds: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques. American Chemical Society, Washington, DC (2015)

242. Formolo, T., Ly, M., Levy, M., Kilpatrick, L., Lute, S., Phinney, K., Marzilli, L., Brorson, K., Boyne, M., Davis, D., Schiel, J.: Determination of the NISTmAb Primary Structure. In: Schiel, J.E., Davis, D.L., and Borisov, O.V. (eds.) ACS Symposium Series. pp. 1–62. American Chemical Society, Washington, DC (2015)

243. Borisov, O.V., Alvarez, M., Carroll, J.A., Brown, P.W.: Sequence Variants and Sequence Variant Analysis in Biotherapeutic Proteins. In: Schiel, J.E., Davis, D.L., and Borisov, O.V. (eds.) ACS Symposium Series. pp. 63–117. American Chemical Society, Washington, DC (2015)

244. Li, W., Kerwin, J.L., Schiel, J., Formolo, T., Davis, D., Mahan, A., Benchaar, S.A.: Structural Elucidation of Post-Translational Modifications in Monoclonal Antibodies. In: Schiel, J.E., Davis, D.L., and Borisov, O.V. (eds.) ACS Symposium Series. pp. 119–183. American Chemical Society, Washington, DC (2015)

245. Prien, J.M., Stöckmann, H., Albrecht, S., Martin, S.M., Varatta, M., Furtado, M., Hosselet, S., Wang, M., Formolo, T., Rudd, P.M., Schiel, J.E.: Orthogonal Technologies for NISTmAb N-Glycan Structure Elucidation and Quantitation. In: Schiel, J.E., Davis, D.L., and Borisov, O.V. (eds.) ACS Symposium Series. pp. 185–235. American Chemical Society, Washington, DC (2015)

246. Michels, D.A., Ip, A.Y., Dillon, T.M., Brorson, K., Lute, S., Chavez, B., Prentice, K.M., Brady, L.J., Miller, K.J.: Separation Methods and Orthogonal Techniques. In: Schiel, J.E., Davis, D.L., and Borisov, O.V. (eds.) ACS Symposium Series. pp. 237–284. American Chemical Society, Washington, DC (2015)

247. Gokarn, Y., Agarwal, S., Arthur, K., Bepperling, A., Day, E.S., Filoti, D., Greene, D.G., Hayes, D., Kroe-Barrett, R., Laue, T., Lin, J., McGarry, B., Razinkov, V., Singh, S., Taing, R., Venkataramani, S., Weiss, W., Yang, D., Zarraga, I.E.: Biophysical Techniques for Characterizing the Higher Order Structure and Interactions of Monoclonal Antibodies. In: Schiel, J.E., Davis, D.L., and Borisov, O.V. (eds.) ACS Symposium Series. pp. 285–327. American Chemical Society, Washington, DC (2015)

248. Saro, D., Baker, A., Hepler, R., Spencer, S., Bruce, R., LaBrenz, S., Chiu, M., Davis, D., Lang, S.E.: Developability Assessment of a Proposed NIST Monoclonal Antibody. In: Schiel, J.E., Davis, D.L., and Borisov, O.V. (eds.) ACS Symposium Series. pp. 329–355. American Chemical Society, Washington, DC (2015)

249. Ripple, D.C., Narhi, L.O.: Protein Particles (0.1 µm to 100 µm). In: Schiel, J.E., Davis, D.L., and Borisov, O.V. (eds.) ACS Symposium Series. pp. 357–386. American Chemical Society, Washington, DC (2015)

250. Drenski, M.F., Brader, M.L., Reed, W.F.: Simultaneous Multiple Sample Light Scattering (SMSLS) for Continuous Monitoring of Protein Aggregation. In: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques. pp. 159–188. American Chemical Society (2015)

251. Remmele, R.L., Bee, J.S., Phillips, J.J., Mo, W.D., Higazi, D.R., Zhang, J., Lindo, V., Kippen, A.D.: Characterization of Monoclonal Antibody Aggregates and Emerging Technologies. In: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques. pp. 113–158. American Chemical Society (2015)

252. Prakash, K., Chen, W.: Analytical Methods for the Measurement of Host Cell Proteins and Other Process-Related Impurities. In: Schiel, J.E., Davis, D.L., and Borisov, O.V. (eds.) ACS Symposium Series. pp. 387–404. American Chemical Society, Washington, DC (2015)

253. Li, W., Xu, H., Borisov, O.: Informatics for Mass Spectrometry-Based Protein Characterization. In: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques. pp. 189–225. American Chemical Society (2015)

254. Kil, Y.J., Bern, M., Crowell, K., Kletter, D., Bern, N., Tang, W., Carlson, E., Becker, C.: Towards a Comprehensive Bioinformatic Analysis of the NIST Reference mAb. In: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques. pp. 395–414. American Chemical Society (2015)

255. Arbogast, L.W., Delaglio, F., Schiel, J.E., Marino, J.P.: Multivariate Analysis of Two-Dimensional 1H, 13C Methyl NMR Spectra of Monoclonal Antibody Therapeutics To Facilitate Assessment of Higher Order Structure. Anal. Chem. 89, 11839–11845 (2017). https://doi.org/10.1021/acs.analchem.7b03571

256. Marino, J.P., Brinson, R.G., Hudgens, J.W., Ladner, J.E., Gallagher, D.T., Gallagher, E.S., Arbogast, L.W., Huang, R.Y.-C.: Emerging Technologies To Assess the Higher Order Structure of Monoclonal Antibodies. In: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques. pp. 17–43. American Chemical Society (2015)

257. Kaur, P., Kiselar, J., Shi, W., Yang, S., Chance, M.R.: Covalent Labeling Techniques for Characterizing Higher Order Structure of Monoclonal Antibodies. In: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques. pp. 45–73. American Chemical Society (2015)

258. Hao, Z., Zhang, T., Xuan, Y., Wang, H., Qian, J., Lin, S., Chen, J., Horn, D.M., Argoti, D., Beck, A., Cianférani, S., Bennett, P., Miller, K., Makarov, A.: Intact Antibody Characterization Using Orbitrap Mass Spectrometry. In: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody

Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques. pp. 289–315. American Chemical Society (2015)

259. Hodoniczky, J., Zheng, Y.Z., James, D.C.: Control of Recombinant Monoclonal Antibody Effector Functions by Fc N-Glycan Remodeling in Vitro. Biotechnol. Prog. 21, 1644–1652 (2005). https://doi.org/10.1021/bp050228w

260. Solá, R.J., Griebenow, K.: Effects of glycosylation on the stability of protein pharmaceuticals. J. Pharm. Sci. 98, 1223–1245 (2009). https://doi.org/10.1002/jps.21504

261. Natsume, A., Niwa, R., Satoh, M.: Improving effector functions of antibodies for cancer treatment: Enhancing ADCC and CDC. Drug Des. Devel. Ther. 3, 7–16 (2009)

262. Quast, I., Keller, C.W., Maurer, M.A., Giddens, J.P., Tackenberg, B., Wang, L.-X., Münz, C., Nimmerjahn, F., Dalakas, M.C., Lünemann, J.D.: Sialylation of IgG Fc domain impairs complement-dependent cytotoxicity. J. Clin. Invest. 125, 4160–4170 (2015). https://doi.org/10.1172/JCI82695

263. Goetze, A.M., Liu, Y.D., Zhang, Z., Shah, B., Lee, E., Bondarenko, P.V., Flynn, G.C.: High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. Glycobiology. 21, 949–959 (2011). https://doi.org/10.1093/glycob/cwr027

264. Higel, F., Seidl, A., Sörgel, F., Friess, W.: N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. Eur. J. Pharm. Biopharm. 100, 94–100 (2016). https://doi.org/10.1016/j.ejpb.2016.01.005

265. Nagelkerke, S.Q., Dekkers, G., Kustiawan, I., Bovenkamp, F.S. van de, Geissler, J., Plomp, R., Wuhrer, M., Vidarsson, G., Rispens, T., Berg, T.K. van den, Kuijpers, T.W.: Inhibition of FcγR-mediated phagocytosis by IVIg is independent of IgG-Fc sialylation and FcγRIIb in human macrophages. Blood. 124, 3709–3718 (2014). https://doi.org/10.1182/blood-2014-05-576835

266. Samuelsson, A., Towers, T.L., Ravetch, J.V.: Anti-inflammatory Activity of IVIG Mediated Through the Inhibitory Fc Receptor. Science. 291, 484–486 (2001). https://doi.org/10.1126/science.291.5503.484

267. Raju, T.S.: Terminal sugars of Fc glycans influence antibody effector functions of IgGs. Curr. Opin. Immunol. 20, 471–478 (2008). https://doi.org/10.1016/j.coi.2008.06.007

268. Houde, D., Peng, Y., Berkowitz, S.A., Engen, J.R.: Post-translational Modifications Differentially Affect IgG1 Conformation and Receptor Binding. Mol. Cell. Proteomics. 9, 1716–1728 (2010). https://doi.org/10.1074/mcp.M900540-MCP200

269. Krapp, S., Mimura, Y., Jefferis, R., Huber, R., Sondermann, P.: Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. J. Mol. Biol. 325, 979–989 (2003). https://doi.org/10.1016/s0022-2836(02)01250-0

270. Pan, L.Y., Salas-Solano, O., Valliere-Douglass, J.F.: Antibody Structural Integrity of Site-Specific Antibody-Drug Conjugates Investigated by Hydrogen/Deuterium Exchange Mass Spectrometry. Anal. Chem. 87, 5669–5676 (2015). https://doi.org/10.1021/acs.analchem.5b00764

271. Inouye, K., Ohnaka, S.: Pepsin digestion of a mouse monoclonal antibody of IgG1 class formed F(ab')(2) fragments in which the light chains as well as the heavy chains were truncated. J. Biochem. Biophys. Methods. 48, 23–32 (2001). https://doi.org/10.1016/s0165-022x(00)00141-x

272. Rey, M., Man, P., Brandolin, G., Forest, E., Pelosi, L.: Recombinant immobilized rhizopuspepsin as a new tool for protein digestion in hydrogen/deuterium exchange mass spectrometry. Rapid Commun. Mass Spectrom. 23, 3431–3438 (2009). https://doi.org/10.1002/rcm.4260

273. Cravello, L., Lascoux, D., Forest, E.: Use of different proteases working in acidic conditions to improve sequence coverage and resolution in hydrogen/deuterium exchange of large proteins. Rapid Commun. Mass Spectrom. 17, 2387–2393 (2003). https://doi.org/10.1002/rcm.1207

274. Man, P., Montagner, C., Vernier, G., Dublet, B., Chenal, A., Forest, E., Forge, V.: Defining the Interacting Regions between Apomyoglobin and Lipid Membrane by Hydrogen/Deuterium Exchange Coupled to Mass Spectrometry. J. Mol. Biol. 368, 464–472 (2007). https://doi.org/10.1016/j.jmb.2007.02.014

275. Marcoux, J., Thierry, E., Vivès, C., Signor, L., Fieschi, F., Forest, E.: Investigating alternative acidic proteases for H/D exchange coupled to mass spectrometry: Plasmepsin 2 but not plasmepsin 4 is active under quenching conditions. J. Am. Soc. Mass Spectrom. 21, 76–79 (2010). https://doi.org/10.1016/j.jasms.2009.09.005

276. Mazon, H., Marcillat, O., Forest, E., Vial, C.: Local dynamics measured by hydrogen/deuterium exchange and mass spectrometry of creatine kinase digested by two proteases. Biochimie. 87, 1101–1110 (2005). https://doi.org/10.1016/j.biochi.2005.05.012

277. Nirudodhi, S.N., Sperry, J.B., Rouse, J.C., Carroll, J.A.: Application of Dual Protease Column for HDX-MS Analysis of Monoclonal Antibodies. J. Pharm. Sci. 106, 530–536 (2017). https://doi.org/10.1016/j.xphs.2016.10.023

278. Matsumiya, S., Yamaguchi, Y., Saito, J., Nagano, M., Sasakawa, H., Otaki, S., Satoh, M., Kato, K.: Corrigendum to "Structural Comparison of Fucosylated and Nonfucosylated Fc Fragments of Human Immunoglobulin G1" [J. Mol. Biol. 386/3 (2007) 767–779]. J. Mol. Biol. - J MOL BIOL. 408, 1001–1001 (2011). https://doi.org/10.1016/j.jmb.2011.03.031

279. Fang, J., Richardson, J., Du, Z., Zhang, Z.: Effect of Fc-Glycan Structure on the Conformational Stability of IgG Revealed by Hydrogen/Deuterium Exchange and Limited Proteolysis. Biochemistry. 55, 860–868 (2016). https://doi.org/10.1021/acs.biochem.5b01323

280. Wu, Y., Engen, J.R., Hobbins, W.B.: Ultra performance liquid chromatography (UPLC) further improves hydrogen/deuterium exchange mass spectrometry. J. Am. Soc. Mass Spectrom. 17, 163–167 (2006). https://doi.org/10.1016/j.jasms.2005.10.009

281. Fang, J.: Pressurized Online Pepsin Digestion of mAb IgG2 for Hydrogen Deuterium Exchange Mass Spectrometry, https://www.waters.com/nextgen/gb/en/library/application-notes/2016/online-pepsin-digestion-of-mab-igg2-for-hydrogen-deuterium-exchange-ms.html

282. Yamaguchi, Y., Nishimura, M., Nagano, M., Yagi, H., Sasakawa, H., Uchida, K., Shitara, K., Kato, K.: Glycoform-dependent conformational alteration of the Fc region of human immunoglobulin G1 as revealed by NMR spectroscopy. Glycoproteomics. 1760, 693–700 (2006). https://doi.org/10.1016/j.bbagen.2005.10.002

283. Zhang, A., Hu, P., MacGregor, P., Xue, Y., Fan, H., Suchecki, P., Olszewski, L., Liu, A.: Understanding the Conformational Impact of Chemical Modifications on Monoclonal Antibodies with Diverse Sequence Variation Using Hydrogen/Deuterium Exchange Mass Spectrometry and Structural Modeling. Anal. Chem. 86, 3468–3475 (2014). https://doi.org/10.1021/ac404130a

284. Manikwar, P., Majumdar, R., Hickey, J.M., Thakkar, S.V., Samra, H.S., Sathish, H.A., Bishop, S.M., Middaugh, C.R., Weis, D.D., Volkin, D.B.: Correlating Excipient Effects on Conformational and Storage Stability of an IgG1 Monoclonal Antibody with Local Dynamics as Measured by Hydrogen/Deuterium-Exchange Mass Spectrometry. J. Pharm. Sci. 102, 2136–2151 (2013). https://doi.org/10.1002/jps.23543

285. Chennamsetty, N., Voynov, V., Kayser, V., Helk, B., Trout, B.L.: Design of therapeutic proteins with enhanced stability. Proc. Natl. Acad. Sci. 106, 11937–11942 (2009). https://doi.org/10.1073/pnas.0904191106

286. Zheng, K., Bantog, C., Bayer, R.: The impact of glycosylation on monoclonal antibody conformation and stability. mAbs. 3, 568–576 (2011). https://doi.org/10.4161/mabs.3.6.17922

287. Jensen, P.F., Larraillet, V., Schlothauer, T., Kettenberger, H., Hilger, M., Rand, K.D.: Investigating the Interaction between the Neonatal Fc Receptor and Monoclonal Antibody Variants by Hydrogen/Deuterium Exchange Mass Spectrometry. Mol. Cell. Proteomics. 14, 148–161 (2015). https://doi.org/10.1074/mcp.M114.042044

288. More, A.S., Toth, R.T., Okbazghi, S.Z., Middaugh, C.R., Joshi, S.B., Tolbert, T.J., Volkin, D.B., Weis, D.D.: Impact of Glycosylation on the Local Backbone Flexibility of Well-Defined IgG1-Fc Glycoforms Using Hydrogen Exchange-Mass Spectrometry. J. Pharm. Sci. 107, 2315–2324 (2018). https://doi.org/10.1016/j.xphs.2018.04.026

289. Shi, L., Liu, T., Gross, M.L., Huang, Y.: Recognition of Human IgG1 by Fcγ Receptors: Structural Insights from Hydrogen-Deuterium Exchange and Fast Photochemical Oxidation of Proteins Coupled with Mass Spectrometry. Biochemistry. 58, 1074–1080 (2019). https://doi.org/10.1021/acs.biochem.8b01048

290. Guttman, M., Scian, M., Lee, K.K.: Tracking Hydrogen/Deuterium Exchange at Glycan Sites in Glycoproteins by Mass Spectrometry. Anal. Chem. 83, 7492–7499 (2011). https://doi.org/10.1021/ac201729v

291. Campuzano, I.D.G., Larriba, C., Bagal, D., Schnier, P.D.: Ion Mobility and Mass Spectrometry Measurements of the Humanized IgGk NIST Monoclonal Antibody. In: State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques. pp. 75–112. American Chemical Society (2015)

292. Upton, R., G. Migas, L., J. Pacholarz, K., G. Beniston, R., Estdale, S., Firth, D., E. Barran, P.: Hybrid mass spectrometry methods reveal lot-to-lot differences and delineate the effects of glycosylation on the tertiary structure of Herceptin®. Chem. Sci. 10, 2811–2820 (2019). https://doi.org/10.1039/C8SC05029E

293. Rinas, A., Mali, V.S., Espino, J.A., Jones, L.M.: Development of a Microflow System for In-Cell Footprinting Coupled with Mass Spectrometry. Anal. Chem. 88, 10052–10058 (2016). https://doi.org/10.1021/acs.analchem.6b02357

294. Espino, J.A., Jones, L.M.: Illuminating Biological Interactions with in Vivo Protein Footprinting. Anal. Chem. 91, 6577–6584 (2019). https://doi.org/10.1021/acs.analchem.9b00244

295. Liu, X.R., Zhang, M.M., Zhang, B., Rempel, D.L., Gross, M.L.: Hydroxyl-Radical Reaction Pathways for the Fast Photochemical Oxidation of Proteins Platform As Revealed by 18O Isotopic Labeling. Anal. Chem. 91, 9238–9245 (2019). https://doi.org/10.1021/acs.analchem.9b02134

296. Hudgens, J.W., Gallagher, E.S., Karageorgos, I., Anderson, K.W., Filliben, J.J., Huang, R.Y.-C., Chen, G., Bou-Assaf, G.M., Espada, A., Chalmers, M.J., Harguindey, E., Zhang, H.-M., Walters, B.T., Zhang, J., Venable, J., Steckler, C., Park, I., Brock, A., Lu, X., Pandey, R., Chandramohan, A., Anand, G.S., Nirudodhi, S.N., Sperry, J.B., Rouse, J.C., Carroll, J.A., Rand, K.D., Leurs, U., Weis, D.D., Al-Naqshabandi, M.A., Hageman, T.S., Deredge, D., Wintrode, P.L., Papanastasiou, M., Lambris, J.D., Li, S., Urata, S.: Interlaboratory Comparison of Hydrogen–Deuterium Exchange Mass Spectrometry Measurements of the Fab Fragment of NISTmAb. Anal. Chem. 91, 7336–7345 (2019). https://doi.org/10.1021/acs.analchem.9b01100

297. Houde, D.J., Bou-Assaf, G.M., Berkowitz, S.A.: Deciphering the Biophysical Effects of Oxidizing Sulfur-Containing Amino Acids in Interferon-beta-1a using MS and HDX-MS. J. Am. Soc. Mass Spectrom. 28, 840–849 (2017). https://doi.org/10.1021/jasms.8b05526

298. Georgescauld, F., Popova, K., Gupta, A.J., Bracher, A., Engen, J.R., Hayer-Hartl, M., Hartl, F.U.: GroEL/ES Chaperonin Modulates the Mechanism and Accelerates the Rate of TIM-Barrel Domain Folding. Cell. 157, 922–934 (2014). https://doi.org/10.1016/j.cell.2014.03.038

299. Yang, B., Stjepanovic, G., Shen, Q., Martin, A., Hurley, J.H.: Vps4 disassembles an ESCRT-III filament by global unfolding and processive translocation. Nat. Struct. Mol. Biol. 22, 492–498 (2015). https://doi.org/10.1038/nsmb.3015

300. Snijder, J., Burnley, R.J., Wiegard, A., Melquiond, A.S.J., Bonvin, A.M.J.J., Axmann, I.M., Heck, A.J.R.: Insight into cyanobacterial circadian timing from structural details of the KaiB–KaiC interaction. Proc. Natl. Acad. Sci. 111, 1379–1384 (2014). https://doi.org/10.1073/pnas.1314326111

List of Abbreviations

analytical ultracentrifugation (AUC)

atmospherics-pressure chemical ionisation (APCI)

arrival time distribution (ATD)

chain ejection model (CEM)

charge residue model (CRM)

charge state distribution (CSD)

circular dichroism (CD)

coefficient of variance (CVs)

collisional cross section (CCS)

collisional induced dissociation (CID)

collisional induced unfolding (CIU)

complementary-determining regions (CDR)

complement-dependent cytotoxicity (CDC)

critical quality attribute (CQA)

cross-linking (XL)

cryogenic electron microscopy (cryo-EM)

deglycosylated (DeGly)

differential mobility spectrometry (DMS)

differential scanning calorimetry (DSC)

digestion factor (DF)

direct current (DC)

drift-time ion mobility spectrometry (DT-IMS)

electron capture dissociation (ECD)

electron ionisation (EI)

electron multiplier (EM)

electron transfer dissociation (ETD)

electrospray ionisation (ESI)

exo-degalactosylated (ExoGal)

extracted ion chromatogram (XIC)

fast photochemical oxidation of proteins (FPOP)

Fourier transform (FT)

fourier transform infrared spectroscopy (FTIR)

full width half maximum (FWHM)

heavy chain (HC)

high performance liquid chromatography (HPLC)

higher order structure (HOS)

hydrogen deuterium exchange (HDX)

hydroxyl radical protein footprinting (HRPF)

immunoglobulins (Igs)

intermediate measurement precision (IMP)

International Bureau of Weights and Measures (BIPM)

ion evaporation model (IEM)

ion mobility spectrometry (IMS)

isothermal titration calorimetry (ITC)

light chain (LC)

matrix-assisted laser desorption ionisation (MALDI)

mass spectrometry (MS)

mass-to-charge ratio (*m/z*)

microchannel plate (MCP)

monoclonal antibody (mAb)

nano electrospray ionisation (nESI)

National Institute for Standards and Technology (NIST)

nuclear magnetic resonance (NMR)

orthogonal-acceleration time of flight (oa-TOF)

peak intensity ratio (PIR)

peak width at 10% maximal intensity (FWTH)

post translational modifications (PTMs)

quadrupole (Q)

quality control (QC)

radio-frequency (RF)

recombinant human growth hormone (rhGH)

reference material (RM)

relative deuterium uptake (rDU)

relative standard deviation (RSD)

research and development (R&D)

retention time (RT)

size exclusion chromatography (SEC)

solvent accessible surface area (SASA)

stacked ring ion guide (SRIG)

standard deviation (SD)

standard error of the mean (SEM)

standard operating procedures (SOP)

surface induced dissociation (SID)

System of International Units (SI)

tandem MS (MS/MS)

time of flight (TOF)

travelling wave ion mobility spectrometry (TWIMS)

ultra performance liquid chromatography (UPLC)

ultra violet photo dissociation (UVPD)

x-ray crystallography (XRC)

Appendix

Table A-24. PLGS workflow and processing parameters used for peptide identification from MS^E experiments

Workflow parameters	
Peptide tolerance	20 ppm
Fragment tolerance	10 ppm
Min fragment ion matches per protein	2
Min fragment ion matches per peptide	7
Min peptide matches per protein	1
Maximum protein mass	250,000
Primary digestion reagents	Non-specific
Missed Cleavages	1
Variable modified reagents	Oxidation, M
False discovery rate	2
Processing parameters	
Chromatographic peak width	Automatic
MS TOF resolution	Automatic
Lock mass for charge 2	785.8426 Da
Lock mass window	0.25 Da
Low energy threshold	500 counts
Elevated energy threshold	100 counts
Intensity threshold	750 counts

Table A-25. Raw differential HDX-MS measurements, S (i), between rhGH and rhGH:Zn 1:2 and 1:10 samples. Measurement thresholds (S_L) established for rhGH vial to vial comparability study are also stated. For comparison with rhGH:Zn samples those measurements highlighted in red are measurements which show a significant structural change relative to established thresholds. All measurements were performed in triplicate and the error of each measurement is represented by the standard deviation (SD S) of the averaged measurements.

Identifier		**rhGH:Zn 1:2**		**rhGH:Zn 1:10**		**Vial-to-vial rhGH**		**Threshold S_L**	
Peptide, i	**Peptide ID**	**S (i) , Da**	**SD S**	**S (i) , Da**	**SD S**	**S (i) , Da**	**SD S**	**S_L +**	**S_L -**
1	1-9-	0.16	0.10	0.51	0.12	0.28	0.14	0.42	-0.42
2	2-9-	-0.04	0.15	0.51	0.16	0.19	0.02	0.21	-0.21
3	3-9-	0.18	0.12	0.33	0.10	0.24	0.09	0.33	-0.33
4	14-25-	-0.19	0.08	-0.62	0.10	0.11	0.03	0.14	-0.14
5	15-25-	-0.33	0.09	-0.81	0.08	0.08	0.03	0.11	-0.11
6	16-25-	-0.16	0.08	-0.78	0.07	0.12	0.07	0.19	-0.19
7	26-32-	-0.93	0.06	-1.70	0.06	0.24	0.14	0.37	-0.37
8	31-44-	-0.17	0.24	-0.17	0.25	0.16	0.06	0.22	-0.22
9	32-44-	-0.09	0.28	-0.01	0.25	0.21	0.04	0.25	-0.25
10	33-44-	-0.39	0.21	0.15	0.26	0.31	0.06	0.37	-0.37
11	34-44-	-0.09	0.15	-0.15	0.21	0.18	0.03	0.21	-0.21
12	35-44-	-0.13	0.22	0.03	0.24	0.20	0.04	0.24	-0.24

13	45-51-	-0.05	0.11	-0.02	0.10	0.22	0.05	0.27	-0.27
14	45-52-	-0.13	0.11	0.03	0.10	0.18	0.07	0.25	-0.25
15	45-63-	-0.81	0.62	-0.18	0.52	0.79	0.38	1.18	-1.18
16	53-63-	0.04	0.17	-0.19	0.17	0.21	0.07	0.28	-0.28
17	53-73-	-0.03	0.43	-0.44	0.32	0.38	0.06	0.43	-0.43
18	54-63-	-0.07	0.16	-0.22	0.15	0.27	0.10	0.37	-0.37
19	54-73-	-0.32	0.35	-0.26	0.34	0.71	0.19	0.90	-0.90
20	55-63-	-0.15	0.12	-0.08	0.11	0.21	0.04	0.25	-0.25
21	55-73-	-0.16	0.31	-0.14	0.31	0.53	0.27	0.80	-0.80
22	64-73-	0.11	0.19	-0.09	0.16	0.57	0.19	0.76	-0.76
23	77-83-	-0.05	0.05	0.09	0.08	0.17	0.08	0.24	-0.24
24	81-87-	-	-	-	-	0.11	0.04	0.14	-0.14
25	86-92-	-0.56	0.34	-0.36	0.41	0.15	0.06	0.21	-0.21
26	93-100-	-0.22	0.10	-0.37	0.14	0.21	0.04	0.25	-0.25
27	98-106-	-0.05	0.12	-0.21	0.15	0.21	0.08	0.29	-0.29
28	98-110-	-0.07	0.18	-0.08	0.81	0.24	0.10	0.34	-0.34
29	101-110-	-0.17	0.18	-0.37	0.16	0.31	0.14	0.45	-0.45
30	102-110-	-0.02	0.10	0.16	0.12	0.14	0.03	0.17	-0.17
31	107-113-	-0.01	0.05	0.05	0.06	0.09	0.01	0.10	-0.10

32	111-119-	0.12	0.12	0.05	0.07	0.13	0.02	0.15	-0.15
33	111-120-	-	-	-	-	0.14	0.03	0.17	-0.17
34	114-124-	-0.15	0.07	-0.48	0.05	0.16	0.06	0.21	-0.21
35	115-124-	-0.12	0.06	-0.40	0.03	0.10	0.02	0.12	-0.12
36	117-124-	-0.19	0.02	-0.45	0.09	0.15	0.06	0.20	-0.20
37	125-138-	0.00	0.19	0.25	0.21	0.40	0.07	0.47	-0.47
38	125-139-	-0.15	0.32	0.28	0.29	0.30	0.08	0.38	-0.38
39	125-146-	-0.03	0.43	0.31	0.49	0.49	0.18	0.67	-0.67
40	126-146-	-0.07	0.84	-	-	0.44	0.22	0.66	-0.66
41	140-146-	-0.04	0.09	0.03	0.96	0.24	0.03	0.27	-0.27
42	140-157-	-	-	-0.01	0.56	0.34	0.15	0.49	-0.49
43	147-156-	-0.02	0.09	0.08	0.09	0.10	0.03	0.14	-0.14
44	147-157-	0.00	0.12	0.05	0.08	0.09	0.03	0.11	-0.11
45	155-163-	-0.09	0.11	-0.04	0.09	0.21	0.08	0.29	-0.29
46	156-163-	0.05	0.14	0.05	0.10	0.24	0.10	0.34	-0.34
47	157-163-	-0.09	0.14	-0.02	0.13	0.19	0.08	0.27	-0.27
48	164-175-	-0.09	0.25	-0.18	0.10	0.09	0.02	0.11	-0.11
49	167-176-	-0.11	0.15	-0.18	0.08	0.16	0.03	0.19	-0.19
50	176-182-	0.02	0.07	0.08	0.08	0.10	0.03	0.13	-0.13

51	178-191-	-	-	0.88	0.18	0.36	0.10	0.46	-0.46
52	182-191-	0.18	0.14	0.70	0.18	0.19	0.03	0.22	-0.22
53	183-191-	-0.03	0.13	0.51	0.16	0.16	0.05	0.21	-0.21

Table A-26. Raw differential HDX-MS measurements, S(i), between NISTmAb RM and Fc glycan variants. Measurement thresholds (S_L) established for NISTmAb RM vial to vial comparability study are stated. For comparison with Fc-glycan variants those measurements highlighted in red are measurements which show a significant structural change relative to established thresholds. All measurements were performed in triplicate and the error of each measurement is represented by the standard deviation (SD S) of the averaged measurements.

Identifier		**Degly NISTmAb**		**Exogal NISTmAb**		**Vial-to-vial Gly NISTmAb**		**Threshold S_L**	
Peptide, i	**Peptide ID**	**S (i) , Da**	**SD S**	**S (i) , Da**	**SD S**	**S (i) , Da**	**SD S**	**S_L +**	**S_L -**
1	Heavy Chain-5-17-	-0.17	0.08	0.44	0.11	0.10	0.05	0.15	-0.15
2	Heavy Chain-5-18-	-	-	-	-	0.12	0.05	0.17	-0.17
3	Heavy Chain-5-19-	-0.27	0.08	0.40	0.14	0.14	0.04	0.19	-0.19
4	Heavy Chain-5-20-	-0.26	0.09	0.38	0.09	0.16	0.04	0.20	-0.20
5	Heavy Chain-14-34-	-	-	-	-	0.05	0.03	0.09	-0.09
6	Heavy Chain-21-27-	-0.11	0.04	0.05	0.07	0.06	0.02	0.07	-0.07
7	Heavy Chain-35-46-	-	-	-	-	0.05	0.04	0.09	-0.09
8	Heavy Chain-35-50-	-0.15	0.16	0.18	0.17	0.10	0.06	0.15	-0.15

9	Heavy Chain-36-50-	-0.12	0.16	0.17	0.19	0.07	0.05	0.12	-0.12
10	Heavy Chain-37-50-	-	-	-	-	0.07	0.02	0.09	-0.09
11	Heavy Chain-38-50-	-	-	-	-	0.16	0.15	0.31	-0.31
12	Heavy Chain-39-48-	-0.07	0.04	0.16	0.07	0.07	0.01	0.08	-0.08
13	Heavy Chain-39-50-	-	-	-	-	0.06	0.01	0.07	-0.07
14	Heavy Chain-66-82-	-0.14	0.07	0.38	0.11	0.07	0.03	0.09	-0.09
15	Heavy Chain-70-82-	-0.05	0.06	0.20	0.07	0.08	0.03	0.10	-0.10
16	Heavy Chain-72-82-	-	-	-	-	0.12	0.03	0.15	-0.15
17	Heavy Chain-83-94-	0.00	0.12	0.06	0.11	0.05	0.07	0.13	-0.13
18	Heavy Chain-106-115-	-0.13	0.06	0.14	0.04	0.11	0.09	0.20	-0.20
19	Heavy Chain-106-117-	-0.09	0.10	0.09	0.09	0.07	0.06	0.13	-0.13
20	Heavy Chain-107-117-	0.02	0.07	0.14	0.07	0.07	0.07	0.14	-0.14
21	Heavy Chain-118-145-	-0.10	0.17	0.09	0.11	0.28	0.16	0.44	-0.44
22	Heavy Chain-118-147-	-0.27	0.23	0.00	0.15	0.43	0.22	0.66	-0.66
23	Heavy Chain-149-158-	-0.07	0.12	0.18	0.13	0.08	0.05	0.12	-0.12
24	Heavy Chain-159-166-	-0.34	0.06	0.30	0.13	0.09	0.02	0.11	-0.11
25	Heavy Chain-159-170-	-0.35	0.09	0.30	0.16	0.18	0.03	0.21	-0.21
26	Heavy Chain-159-177-	-0.45	0.09	0.30	0.19	0.13	0.10	0.23	-0.23
27	Heavy Chain-162-177-	-0.33	0.09	0.13	0.13	0.13	0.04	0.17	-0.17

28	Heavy Chain-165-177-	-0.18	0.09	0.12	0.08	0.08	0.06	0.15	-0.15
29	Heavy Chain-166-177-	-0.12	0.04	0.14	0.08	0.07	0.03	0.09	-0.09
30	Heavy Chain-167-177-	-0.24	0.09	0.14	0.09	0.35	0.35	0.70	-0.70
31	Heavy Chain-171-177-	-0.10	0.03	0.07	0.04	0.05	0.02	0.07	-0.07
32	Heavy Chain-183-200-	-0.18	0.14	0.06	0.11	0.29	0.22	0.50	-0.50
33	Heavy Chain-188-196-	-0.07	0.03	0.04	0.06	0.05	0.03	0.08	-0.08
34	Heavy Chain-188-200-	-0.12	0.08	0.01	0.09	0.13	0.01	0.14	-0.14
35	Heavy Chain-189-196-	-0.11	0.04	0.10	0.06	0.05	0.01	0.06	-0.06
36	Heavy Chain-189-200-	-0.12	0.10	0.11	0.12	0.10	0.02	0.12	-0.12
37	Heavy Chain-238-244-	-0.79	0.03	0.03	0.06	0.02	0.01	0.04	-0.04
38	Heavy Chain-238-255-	-	-	-	-	0.13	0.09	0.22	-0.22
39	Heavy Chain-244-254-	-0.89	0.04	-0.05	0.09	0.04	0.01	0.05	-0.05
40	Heavy Chain-245-254-	-0.77	0.22	-0.08	0.20	0.02	0.02	0.04	-0.04
41	Heavy Chain-245-255-	-1.08	0.08	-0.01	0.11	0.03	0.02	0.05	-0.05
42	Heavy Chain-246-255-	-0.85	0.05	-0.28	0.06	0.07	0.07	0.14	-0.14
43	Heavy Chain-247-254-	-0.60	0.03	-0.16	0.05	0.04	0.04	0.09	-0.09
44	Heavy Chain-247-255-	-0.75	0.03	-0.27	0.06	0.06	0.00	0.06	-0.06
45	Heavy Chain-255-264-	-	-	-	-	0.11	0.08	0.19	-0.19
46	Heavy Chain-256-263-	-0.08	0.03	0.06	0.03	0.04	0.01	0.05	-0.05

47	Heavy Chain-256-265-	-0.02	0.03	0.06	0.04	0.08	0.06	0.14	-0.14
48	Heavy Chain-265-278-	-1.33	0.06	0.11	0.08	0.10	0.07	0.17	-0.17
49	Heavy Chain-265-280-	-1.35	0.10	0.16	0.15	0.09	0.03	0.12	-0.12
50	Heavy Chain-266-280-	-1.28	0.05	0.15	0.11	0.13	0.08	0.22	-0.22
51	Heavy Chain-267-280-	-1.02	0.10	0.06	0.09	0.09	0.04	0.13	-0.13
52	Heavy Chain-269-280-	-0.25	0.15	0.12	0.13	0.06	0.01	0.07	-0.07
53	Heavy Chain-310-321-	-0.18	0.12	0.08	0.13	0.09	0.05	0.14	-0.14
54	Heavy Chain-310-336-	-0.42	0.07	0.27	0.16	0.14	0.07	0.21	-0.21
55	Heavy Chain-310-359-	-	-	-	-	0.34	0.44	0.78	-0.78
56	Heavy Chain-312-321-	-0.27	0.08	0.06	0.08	0.06	0.05	0.12	-0.12
57	Heavy Chain-322-336-	-	-	-	-	0.10	0.04	0.14	-0.14
58	Heavy Chain-337-351-	-0.04	0.06	0.27	0.14	0.10	0.03	0.13	-0.13
59	Heavy Chain-337-359-	-0.16	0.49	0.41	0.49	0.13	0.07	0.21	-0.21
60	Heavy Chain-337-368-	-	-	-	-	0.26	0.09	0.35	-0.35
61	Heavy Chain-352-359-	0.00	0.07	0.03	0.07	0.03	0.01	0.04	-0.04
62	Heavy Chain-352-368-	-0.24	0.13	0.18	0.12	0.10	0.03	0.13	-0.13
63	Heavy Chain-355-368-	0.02	0.10	0.19	0.08	0.23	0.13	0.35	-0.35
64	Heavy Chain-369-379-	-0.03	0.06	-0.04	0.10	0.11	0.07	0.18	-0.18
65	Heavy Chain-369-383-	-0.02	0.14	0.08	0.11	0.11	0.01	0.12	-0.12

66	Heavy Chain-371-383-	0.09	0.08	0.14	0.10	0.07	0.05	0.11	-0.11
67	Heavy Chain-372-379-	0.07	0.04	0.06	0.03	0.04	0.02	0.06	-0.06
68	Heavy Chain-372-381-	0.09	0.05	0.03	0.13	0.13	0.16	0.29	-0.29
69	Heavy Chain-372-382-	-	-	-	-	0.10	0.08	0.17	-0.17
70	Heavy Chain-372-383-	0.10	0.07	0.14	0.07	0.04	0.01	0.05	-0.05
71	Heavy Chain-380-393-	-	-	-	-	0.16	0.05	0.22	-0.22
72	Heavy Chain-380-401-	-0.11	0.15	0.23	0.18	0.15	0.13	0.28	-0.28
73	Heavy Chain-382-401-	-0.05	0.14	0.34	0.17	0.18	0.11	0.29	-0.29
74	Heavy Chain-383-393-	-	-	-	-	0.13	0.05	0.18	-0.18
75	Heavy Chain-384-393-	-0.05	0.07	0.11	0.15	0.10	0.04	0.13	-0.13
76	Heavy Chain-384-401-	-0.13	0.09	0.24	0.15	0.08	0.08	0.16	-0.16
77	Heavy Chain-385-401-	-	-	-	-	0.13	0.13	0.26	-0.26
78	Heavy Chain-394-401-	0.02	0.03	0.11	0.06	0.03	0.04	0.07	-0.07
79	Heavy Chain-395-401-	0.03	0.03	-0.02	0.05	0.05	0.07	0.12	-0.12
80	Heavy Chain-402-408-	0.04	0.05	-0.02	0.07	0.08	0.09	0.17	-0.17
81	Heavy Chain-414-425-	-	-	-	-	0.13	0.05	0.18	-0.18
82	Heavy Chain-414-426-	-0.19	0.08	0.33	0.14	0.07	0.02	0.10	-0.10
83	Heavy Chain-427-449-	-0.04	0.47	0.28	0.48	0.06	0.03	0.10	-0.10
84	Heavy Chain-429-449-	-0.02	0.69	0.28	0.74	0.09	0.02	0.11	-0.11

85	Heavy Chain-430-449-	-0.21	0.06	0.18	0.10	0.07	0.01	0.08	-0.08
86	Heavy Chain-432-449-	-0.11	0.15	0.19	0.18	0.06	0.02	0.08	-0.08
87	Heavy Chain-432-450-	-0.08	0.46	0.25	0.48	0.11	0.08	0.19	-0.19
88	Heavy Chain-434-449-	-0.16	0.15	0.13	0.17	0.09	0.07	0.16	-0.16
89	Heavy Chain-436-449-	0.18	0.08	0.29	0.15	0.05	0.03	0.08	-0.08
1	Light Chain-5-11-	0.03	0.07	0.12	0.07	0.09	0.07	0.16	-0.16
2	Light Chain-35-45-	-0.24	0.09	0.15	0.05	0.12	0.03	0.15	-0.15
3	Light Chain-36-45-	-	-	-	-	0.13	0.03	0.15	-0.15
4	Light Chain-39-45-	-	-	-	-	0.04	0.05	0.08	-0.08
5	Light Chain-46-53-	-0.02	0.06	0.02	0.10	0.08	0.05	0.13	-0.13
6	Light Chain-54-70-	0.00	0.23	0.24	0.20	0.31	0.15	0.46	-0.46
7	Light Chain-71-81-	0.08	0.09	-0.02	0.12	0.06	0.02	0.08	-0.08
8	Light Chain-71-82-	-0.03	0.04	0.08	0.09	0.10	0.03	0.13	-0.13
9	Light Chain-72-82-	0.00	0.13	-	-	0.19	0.06	0.25	-0.25
10	Light Chain-74-82-	-0.02	0.07	-	-	0.12	0.04	0.16	-0.16
11	Light Chain-89-115-	-0.03	0.28	-	-	0.16	0.03	0.19	-0.19
12	Light Chain-115-121-	-0.03	0.02	0.02	0.05	0.05	0.01	0.06	-0.06
13	Light Chain-115-122-	-0.07	0.03	0.04	0.05	0.09	0.02	0.11	-0.11
14	Light Chain-115-124-	-0.11	0.10	0.22	0.19	0.08	0.05	0.14	-0.14

15	Light Chain-116-122-	-	-	-	-	0.06	0.04	0.10	-0.10
16	Light Chain-122-131-	-0.28	0.05	0.23	0.09	0.08	0.04	0.12	-0.12
17	Light Chain-123-131-	0.00	0.11	0.01	0.07	0.04	0.01	0.05	-0.05
18	Light Chain-125-133-	-	-	-	-	0.07	0.03	0.10	-0.10
19	Light Chain-131-142-	-0.10	0.16	0.04	0.12	0.10	0.04	0.14	-0.14
20	Light Chain-135-160-	-0.49	0.31	0.24	0.31	0.27	0.15	0.42	-0.42
21	Light Chain-143-160-	-0.23	0.22	0.21	0.30	0.23	0.06	0.29	-0.29
22	Light Chain-148-160-	0.03	0.11	0.14	0.20	0.12	0.06	0.17	-0.17
23	Light Chain-161-171-	-	-	-	-	0.08	0.07	0.15	-0.15
24	Light Chain-161-174-	-0.07	0.09	0.16	0.07	0.05	0.04	0.09	-0.09
25	Light Chain-161-177-	-0.14	0.15	0.13	0.09	0.11	0.02	0.13	-0.13
26	Light Chain-161-178-	-0.16	0.16	0.03	0.09	0.14	0.02	0.16	-0.16
27	Light Chain-172-178-	-0.07	0.05	0.07	0.06	0.04	0.03	0.07	-0.07
28	Light Chain-172-180-	-0.06	0.08	0.11	0.06	0.09	0.05	0.14	-0.14
29	Light Chain-178-191-	-	-	-	-	0.04	0.02	0.07	-0.07
30	Light Chain-178-194-	-0.07	0.21	0.14	0.21	0.06	0.02	0.08	-0.08
31	Light Chain-179-194-	0.01	0.12	0.18	0.10	0.05	0.03	0.08	-0.08
32	Light Chain-181-194-	0.07	0.26	0.02	0.22	0.04	0.02	0.06	-0.06
33	Light Chain-195-213-	0.00	0.34	0.17	0.35	0.07	0.05	0.12	-0.12

34	Light Chain-201-213-	-0.05	0.07	0.09	0.11	0.05	0.04	0.09	-0.09

Table A-27. Tryptic peptides from an in-silico digest of rhGH.

Peptide ID	Residues	Sequence
T1	1-8	FPTIPLSR
T2	9-16	LFDNAMLR
T3	17-19	AHR
T4	20-38	LHQLAFDTYQEFEEAYIPK
T5	39-41	EQK
T6	42-64	YSFLQNPQTSLCFSESIPTP SNR
T7	65-70	EETQQK
T8	71-77	SNLELLR
T9	78-94	ISLLLIQSWLEPVQFLR
T10	95-115	SVFANSLVYGASDSNVYDLL K
T11	116-127	DLEEGIQTLMGR
T12	128-134	LEDGSPR
T13	135-140	TGQIFK
T14	141-145	QTYSK
T15	146-158	FDTNSHNDDALLK
T16	159-167	NYGLLYCFR
T17	168	K
T18	169-172	DMDK
T19	173-178	VETFLR
T20	179-183	IVQCR
T21	184-191	SVEGSCGF