Analysis of Fragmented and Unfragmented Human and Porcine Immunoglobulin G (IgG) *N*-Glycoforms by Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry and Ultra Performance Liquid Chromatography Mass Spectrometry

by

Claudia A. Nelson

A thesis submitted to the Faculty of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry

Faculty of Science

University of Manitoba

Winnipeg, Manitoba

Table of Contents

List of tables

List of Figures

List of Acronyms

1). Antibodies	(Abs)
2). Immunoglobulin	(IgG)
3). Antigen-binding fragments	(Fab), (Fab'), $(F(ab'))_2$
4). Fragments crystallizable	(Fc)
5). Matrix-assisted laser desorption/ionization-mass spectrometry	(MALDI-MS)
6). Ultra performance liquid chromatography	(UPLC)
7). Electrospray ionization-mass spectrometry	(ESI-MS)
8). High performance liquid chromatography in the size exclusion	(HPLC- SEC)
9). Sodium dodecyl sulfate polyacrylamide gel electrophoresis	(SDS- PAGE)
10). Post source decay	(PSD)
11). *N*-Acetylglucosamine	(GlcNAc)
12). TOF/TOF technology	(LIFT™)
13). Post-translational modification	(PTM)
14). Hydrophobic Interaction	(HIC)
15). Acetonitrile	(ACN)
16). Dalton	(Da)
17). Kilo dalton	(kDa)

18). Kinetic Energy (KE)

19). Molar (M)

20). Millimolar (mM)

21). Microliter (µl)

22). Triflouroacetic acid (TFA)

23). Mass spectrometry (MS)

24). Mass-to-charge ratio (*m/z)*

25). Tandem Mass Spectrometry (MS/MS)

26). Time of Flight (TOF)

27). Tandem Time of flight (TOF/TOF)

28). *N*- acetylneuraminic acid (NeuAc)

29). Micromolar (µM)

30). Liquid Chromatography (LC)

31). Collision induced dissociation CID

32). Laser induced dissociation (LID)

Word Index

1)**. Fragmentation:** selectively cleave the Ig molecule into fragments that have discrete characteristics.

2)**. Flow Rate:** The volume of mobile phase passing through the column in unit time. In HPLC systems, the flow rate is set by the controller for the solvent delivery system [pump]. Flow rate accuracy can be checked by timed collection and measurement of the effluent at the column outlet.

3). **Mobility:** is a function of the length, conformation and charge of the molecule.

4). **Gradient:** The change over time in the relative concentrations of two [or more] miscible solvent components that form a mobile phase of increasing elution strength.

5). **Isocratic Elution:** A procedure in which the composition of the mobile phase remains constant during the elution process.

6). **Peak:** The portion of a differential chromatogram recording the detector response while a single component is eluted from the column.

7). **Retention Time:** The time between the start of elution [typically, in HPLC, the moment of injection or sample introduction] and the emergence of the peak maximum.

8). **UPLC Technology:** The use of a high-efficiency LC system holistically designed to accommodate sub-2 μM particles and very high operating pressure is termed ultra-performance liquid chromatography.

9). **Chromatogram**: A chromatogram is a representation of the separation that has chemically occurred in the LC system. A series of peaks rising from a baseline is drawn on a time axis. Each peak represents the detector response for a different compound.

10). **Antibodies (Abs**): most of which are immunoglobulins G (IgGs), are large Y-shaped protein complexes which function to identify and help remove foreign antigens or targets such as viruses and bacteria.

11). **IdeS Protease (eg. FaBRICATOR)**: is immunoglobulin G (IgG)-degrading enzymes that are valuable tools for the characterization of therapeutic antibodies, Fc fusion proteins and antibody-drug conjugates.

12). **TOF/TOF technology (LIFT™**): enables the use of various MS/MS techniques (e.g. LID, high energy CID) for fast and sensitive MS/MS experiments.

Acknowledgements

I would like to thank my supervisor, Dr. Hélène Perreault, for all her support, advice, effort with this research, as well as her heartfelt kindness. I would also like to thank my committee members, Dr. Gregg Tomy (Chemistry), and Dr. Werner Ens (Physics), for their effort and time throughout this process.

I would also like to thank the following people for their much appreciated technical assistance: Emy Komatsu (Chemistry), and Raymond Bacala (Chemistry).

I would also like to thank Genovis (Cambridge, MA, USA) for supplying the Fabulous™ enzyme.

I would also like to thank my family for their continued support, love and encouragement while completing this program, namely my mom, Angella Rowe (1961-2017), brother (Levol Williams), Aunt (Janet Rowe-Pingue), and uncle (Ervin Pingue).

Abstract

A developed workflow for detailed structural characterization of antibodies from different sources, using methods which involve MS is discussed in this thesis. The aim of this thesis is to emphasize the usefulness of detailed characterization of glycoforms of human and pig immunoglobulin (IgG), by comparing a conventional method of (1) whole antibody digestion with trypsin with (2) a more detailed and extensive workflow method of fragmenting the antibody by papain, Fabricator™ and Fabulous™. In (2) antigen-binding fragments (Fab) and crystallisable fragments (Fc) then obtained from enzymatic fragmentation are digested in-gel with trypsin and analyzed by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) and ultra performance liquid chromatography (UPLC-MS). Results from Fab-Fc fragmentation are compared to those obtained using a more conventional approach where whole antibodies are trypsinised into peptides and glycopeptides (1), to indicate which method gives a more detailed and accurate determination of structure. The fragmentation of human and pig IgGs into their Fc and Fab portions is followed by fragment collection and separation by high performance liquid chromatography in the size exclusion mode (HPLC SEC) and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE). This procedure with papain, HPLC SEC and SDS PAGE(2) results in a more detailed analysis of antibodies than the more conventional method(1), as it allows for detailed analysis of individual Fab and Fc without interference of these fragments with each other. The use of UPLC-MS was complementary to MALDI-MS in the determination of glycosylation of the Fc portions of the Abs.

Chapter 1

1.1 Introduction and background

1.1.1 Understanding structure and function of immunoglobulins. Antibodies (Abs) need to be analyzed for glycosylation, following production by bioengineering methods or after extraction from biological tissue or fluid [1], thus the importance of the various analyses done in this thesis on human and pig immunoglobulins G (IgGs). Antibodies (Abs), most of which are IgGs, are large Y-shaped protein complexes which function to identify and help remove foreign antigens or targets such as viruses and bacteria. Every different Ab recognizes a specific foreign antigen [2]. Abs are produced by the immune system in response to the presence of an antigen. Antigens are large molecules, usually proteins, on the surface of cells, viruses, fungi, bacteria, and some non-living substances such as toxins, chemicals, and foreign particles. Any substance capable of triggering an immune response is called an antigen [2]. The IgG Ab subclass is the most abundant serum Ig of the immune system. It is secreted by B cells and is found in blood and extracellular fluids and provides protection from infections caused by bacteria, fungi and viruses [3]. IgG antibodies are large molecules of about 150 kDa made of four peptide chains. They contain two identical class γ heavy chains of about 50 kDa and two identical light chains of about 25 kDa, thus forming a tetrameric quaternary structure. The two heavy chains are linked to each other and to a light chain by disulfide bonds. The resulting tetramer has two identical halves, which together form the Y-like shape (see Figure 1.1.1).

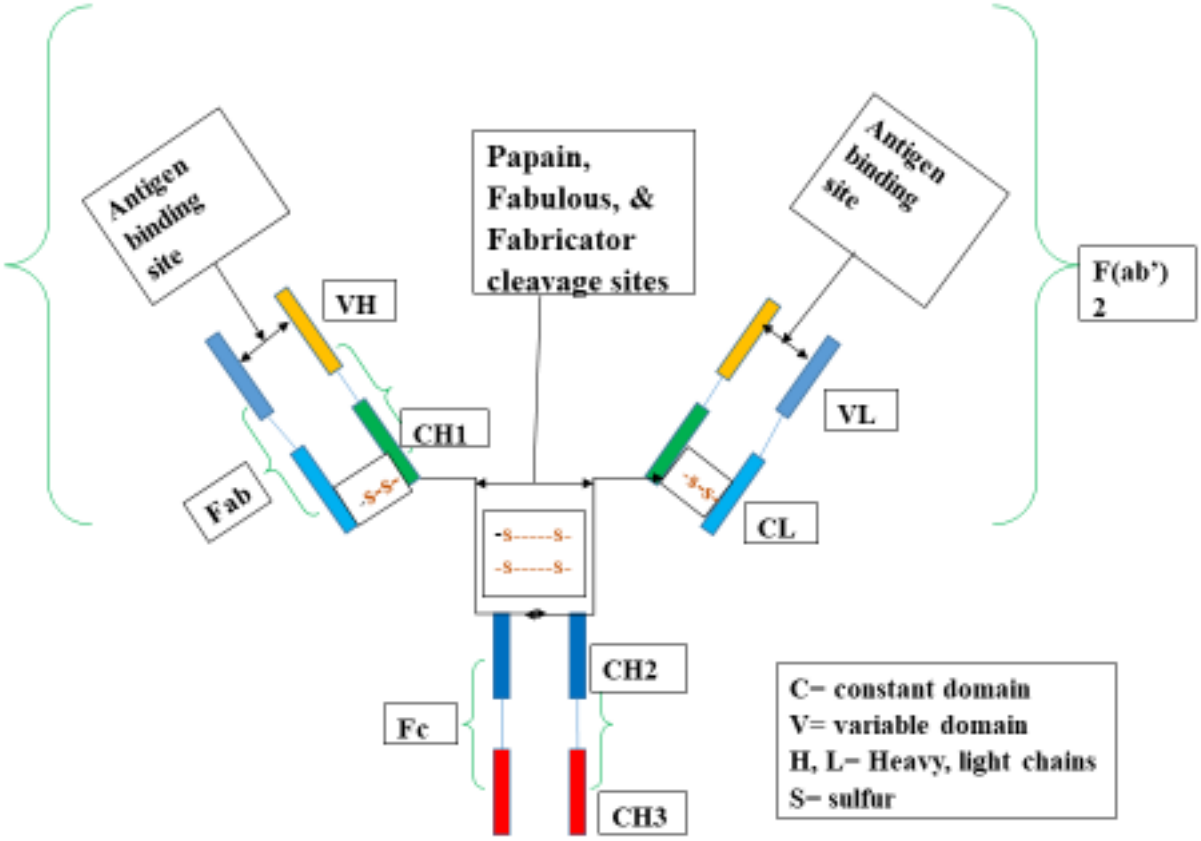

Figure 1.1.1: Structure of Immunoglobulin G (IgG), with enzyme cleavage sites indicated. *Adapted from ref [8]*.

The Fc regions of IgGs bear a highly conserved *N*-glycosylation site. The *N*-glycans attached to this site are predominantly core-fucosylated biantennary structures of the complex type (see Figure 1.1.2). In addition, small amounts of these *N*-glycans bear bisecting *N*-acetylglucosamine (GlcNAc) and sialic acid (*N*-acetylneuraminic or *N*-glycolylneuraminic acid, NeuAc or NeuGc) residues [4].

1.1.2 Understanding glycoforms of IgG. Glycosylation is the attachment of sugar moieties to proteins, lipids, and other chemicals. It is a post-translational modification (PTM) that provides great proteomic diversity. Glycosylation is critical for a wide range of biological processes, including cell attachment to extracellular matrix and protein ligand interactions in the cell [5]. N-linked glycosylation, is the attachment of the sugar molecule oligosaccharide known as glycan to a nitrogen atom (amide nitrogen of asparagine (Asn) residue of a protein), in a process called *N*-glycosylation. This type of linkage is important for both the structure and function of some eukaryotic proteins. The N-linked glycosylation process occurs in eukaryotes and widely in archaea, but very rarely in bacteria. The nature of *N*-linked glycans attached to a glycoprotein is determined by the protein and the cell in which it is expressed. It also varies across species. Different species synthesize different types of *N*-linked glycan. Glycoforms are any of several forms of a glycoprotein (or other biological glycoside) having different saccharides attached, or having a different structure. Glycoforms of an IgG can be purified, detected and analyzed by different strategies, including glycan staining and visualization, glycan crosslinking to agarose or magnetic resin for labeling or for purification, or proteomic analysis by mass spectrometry [5]. IgGs are proteins that comprise glycans, glycans make up a small proportion of the molecular weight of the IgGs. IgGs can be analyzed as whole glycoprotein assemblies or as individual fragmented glycoprotein/protein components. Also, glycans may be removed and analyzed separately. Glycoproteomics represents the global analysis of glycosylated proteins and integrates glycoprotein enrichment and proteomic analysis for the systematic identification and quantitation of glycoproteins in complex systems [5]. This subset of proteomics differs from glycomics, which is restricted to all glycans in a system, released from glycoconjugates or free.

In one example of glycoproteomic analysis workflow, glycoproteins are first digested into glycopeptides (usually by trypsin protease) and either analyzed directly by liquid chromatography and tandem mass spectrometry (LC-MS/MS) or deglycosylated prior to analysis. An enzyme that reveals data on *N*- linked glycosylation is trypsin. Trypsin is a serine protease, found in the digestive system of many vertebrates, where it hydrolyses proteins [6]. It cleaves peptide chains mainly at the carboxyl side of the amino acids lysine or arginine, except when either is followed by proline. It is used for numerous biotechnological processes [6]. The process is commonly referred to as trypsin proteolysis or trypsinization, and proteins that have been digested/treated with trypsin are said to have been trypsinized. Another enzyme, PNgase F, is at times used to reveal details on glycan structure of an antibody. PNGase F is the most effective enzymatic method for removing almost all *N*-linked oligosaccharides from glycoproteins. PNGase F is an amidase, which cleaves between the innermost GlcNAc and asparagine residues of high mannose, hybrid, and complex oligosaccharides. It can be used under native or denaturing conditions, and optimized for deglycosylation of glycoproteins. It leaves *N*-glycan core oligosaccharides intact and suitable for further analysis [7].

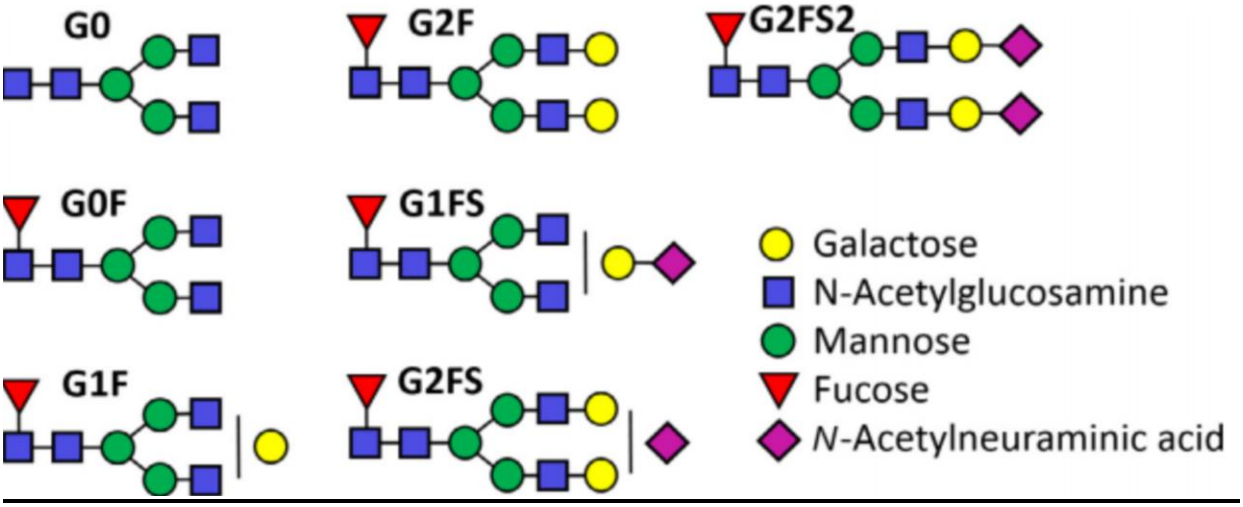

Figure 1.1.2: Main *N*- linked glycans observed in human-derived antibodies. *Adapted from ref [1]*.

1.1.3 Antibody Fragmentation. Due to their interaction with several types of molecules, Abs are powerful tools for protein and molecular detection and purification. Although whole Abs (usually IgG or IgM) are ideal for most immunoassay applications and analytical purposes [8], the performances of certain procedures are enhanced by using Ab fragments, such as Fab and $F(ab')_2$. Sometimes it is useful to study or make use of the activity of one portion of an Ig without interference from other portions of the molecule [8]. It is possible to selectively cleave the Ig molecule into fragments that have discrete characteristics. Antibody fragmentation is accomplished using reducing agents and proteases that digest or cleave certain portions of the Ig protein structure [8]. Although fragmentation of all Ig classes is possible, only procedures for fragmentation of mouse, rabbit, and human IgGs and IgMs have been well characterized [8]. The two groups of antibody fragments of primary interest are (a) antigen-binding fragments such as Fab and (b) class-defining fragments such as Fc (fragment crystallizable) that do not bind

antigens [8]. Several types of antigen-binding fragments are possible, but each contains at least the variable regions of both heavy and light immunoglobulin chains (VH and VL, respectively) held together (by disulfide bonds) so as to preserve the antigen-binding site. Fc fragments consist of the heavy chain constant region (Fc region) of an immunoglobulin and mediate cellular effector functions (see Figure 1.1.1) [8]. Ab fragmentation is somewhat laborious, requires optimization of enzyme-mediated digestion of the protein and necessitates an ample supply of Ab (e.g., 10 mg) to make it reasonably efficient. For these reasons, fragmentation is usually performed only when the antibody of interest is available in large quantity and the particular application demands it [8].

The hinge region of an immunoglobulin monomer (IgG) is readily accessible to proteolytic attack by enzymes. Cleavage at this point produces F(ab')2 or Fab fragments and the Fc fragment [8]. Fab' fragments can be formed by the reduction of F(ab')2 fragments [8]. The Fab' fragment contains a free sulfhydryl group that may be alkylated or utilized in conjugation with an enzyme, toxin or other protein of interest. Fc (50,000 daltons) fragment contain the CH2 and CH3 regions and part of the hinge region held together by one or more disulfides and noncovalent interactions. Fc and Fc5μ fragments are produced from fragmentation of IgG and IgM, respectively. The term Fc is derived from the ability of these antibody fragments to crystallize. Fc fragments are generated entirely from the heavy chain constant region of an immunoglobulin [8]. The Fc fragment cannot bind antigens, but it is responsible for the effector functions of antibodies, such as complement fixation. The Fc fragment may remain intact or become further degraded, depending upon the enzyme and conditions used (see Figure 1.1.1). Traditionally, proteolysis was accomplished in solution using free enzyme. Currently newly developed immobilized enzyme products that enable better control of digestion and efficient

separation of reaction-products from the protease are used [8]. Papain is a nonspecific, thiol-endopeptidase that has a sulfhydryl group in the active site, which must be in the reduced form for activity. When IgG molecules are incubated with papain in the presence of a reducing agent, one or more peptide bonds in the hinge region are split, producing three fragments of similar size: two Fab fragment and one Fc fragment [8]. When Fc fragments are of interest, papain is the enzyme of choice because it yields an intact 50,000-dalton Fc fragment. The Fc may remain intact based on conditions and enzyme used. The optimal pH for papain is 6.5, although it can be effective between 4 to 9.5. Papain is primarily used to generate Fc and Fab fragments, but it also can be used to generate $F(ab')_2$ fragments. To prepare $F(ab')_2$ fragments, the papain is first activated with 10 mM cysteine. The excess cysteine is then removed by gel filtration. If no cysteine is present during papain digestion, F(ab')2 fragments can be generated. These fragments are often inconsistent, and reproducibility can be a problem [8]. If the cysteine is not completely removed, overdigestion of disulfide bonds can be a problem.

Crystalline papain is often used for the digestion of IgG; however, it is prone to autodigestion. Mercuripapain, which is less prone to auto digestion than crystalline papain, can be used; however, both of these non-immobilized enzymes require an oxidant to terminate digestion. Immobilized papain (i.e., papain bound to agarose resin) is the preferred reagent because it allows for easy control of the digestion reaction and quick removal of enzyme from the digestion products following incubation [8]. The use of immobilized papain also prevents formation of antibody-enzyme adducts, which can occur when using the soluble form of sulfhydryl proteases. These adducts can be detrimental to fragments in the presence of reductants. Immobilization also increases stability of the enzyme against heat denaturation and autolysis and results in longer maintenance of activity. Regeneration and reuse of papain resin

are possible, which decreases costs. Cleavage can be regulated by digestion time or flow rate through a column, yielding reproducible digests.

Fabulous™ (SpeB) enzyme is a recombinantly produced cysteine protease that under reduced conditions digests in the hinge region of antibodies from many species and subclasses, including human, mouse, rat and goat, yielding Fab and Fc fragments (see Figure 1.1.1). As reducing agent is present during digestion reaction, it is likely that interchain thiols will be reduced [9]. Fabricator™ is a partially modified and his-tagged IdeS enzyme cleaves IgG in the hinge region leaving an intact F(ab′)2 and two residual Fc fragments. Fabricator™ is an enzyme for preparation of (Fab)'$_2$. Fab and F(ab′)$_2$ antibody fragments are used in applications or assays where the presence of the Fc region may cause problems [10]. Fab and F(ab′)$_2$ fragments are also desirable for staining or binding to cell preparations in the presence of plasma since they cannot bind to complement, which would lyse the cells [11]. The divalency of the F(ab′)$_2$ fragment enables it to cross-link antigens and hence makes it useful for resetting, cellular aggregation and precipitation assays. F(ab′)$_2$ fragments are generally produced using pepsin.

Here, a novel method for antibody fragmentation using the enzyme Fabricator™ is introduced. This enzyme is a partially modified his-tagged IdeS enzyme [12]. Fabricator™ cleaves IgG in the hinge region leaving an intact F(ab′)$_2$ and two residual Fc fragments. The enzymes papain, Fabulous™ and Fabricator™ mentioned above were used for the generation of Fab, Fab, F(ab′)$_2$ fragments respectively and Fc fragments of human and pig IgGs.

1.1.4 Separation methods of fragments HPLC is a highly improved form of column chromatography. Instead of a solvent being allowed to drip through a column by gravity, it is

forced through under high pressures. This makes separation much faster and more efficient. It also allows usage of very much smaller particle sizes for the column packing material, which gives a much greater surface area for interactions between the stationary phase and the molecules flowing past it. This allows a much better separation of the components of the mixture [13]. HPLC is now one of the most powerful tools in analytical chemistry. It has the ability to separate, identify, and quantify compounds present in most samples that can be dissolved in a liquid, using different types of detectors. Today, compounds in trace concentrations as low as parts per trillion (ppt) may easily be identified. HPLC can be, and has been, applied to just about any sample, such as pharmaceuticals, food, nutraceuticals, cosmetics, environmental matrices, forensic and industrial samples [14]. There are two main variants in use for HPLC, depending on the relative polarity of the solvent and the stationary phase: normal and reversed phase HPLC.

Normal phase HPLC: In this mode the column is filled with polar particles, either silica or amine coated silica, and the solvent is non-polar. Polar compounds in the mixture being passed through the column will adhere longer to the polar stationary phase than non-polar compounds will. The non-polar compounds will therefore pass more quickly through the column.

Reversed phase HPLC: In this case, the silica is modified to make it non-polar by attaching long hydrocarbon chains to its surface - typically with either 8 or 18 carbon atoms in them. A polar solvent is used - for example, a mixture of water and an alcohol such as methanol. In this case, there will be a strong attraction between the polar solvent and polar molecules in the mixture being passed through the column. There will not be as much attraction between the hydrocarbon chains attached to the silica (the stationary phase) and the polar molecules in the solution. Polar molecules in the mixture will therefore elute faster with the solvent [13]. Non-

polar compounds in the mixture will tend to form attractions with the hydrocarbon groups because of van der Waals dispersion forces. They have less affinity for the solvent and this will slow them down on their way through the column. Reversed phase HPLC is the most commonly used form of HPLC [13].

In HPLC, a reservoir holds the solvent (called the mobile phase). A high-pressure pump (solvent delivery system) is used to generate and meter a specified flow rate of mobile phase, typically milliliters per minute for analytical mode HPLC. An injector (manual or autosampler) is able to introduce (inject) the sample into the continuously flowing mobile phase stream that carries the sample into the HPLC column. The column contains the chromatographic packing material needed to achieve the separation. This packing material is called the stationary phase. A detector is needed to monitor the separated compounds as they elute from the HPLC column. The mobile phase exits the detector and can be sent to waste, or collected, as desired. When the mobile phase contains a separated compound band, HPLC provides the ability to collect this fraction of the eluate containing that purified compound for further study [14]. The detector is connected to the computer data station, the HPLC system component that records the electrical signal needed to generate the chromatogram on its display and to identify and quantify the concentration of the sample constituents. As sample compound characteristics can be very different, several types of detectors have been developed. For example, if a compound can absorb ultraviolet light, a UV-absorbance detector is used. If the compound fluoresces, a fluorescence detector is used. If the compound does not have either of these characteristics, a more universal type of detector is used, such as an evaporative-light-scattering detector (ELSD) can be used. The most powerful approach is the use of multiple detectors in series [14]. The

output will be recorded as a series of peaks - each one representing a compound in the mixture passing through the detectors [13].

Two main elution modes are used in HPLC. The first is called isocratic elution. In this mode, the mobile phase, either a pure solvent or a mixture, remains the same composition throughout the run. The second type is called gradient elution, wherein, as its name implies, the mobile phase composition changes during the separation. This mode is useful for samples that contain compounds that span a wide range of polarity. As the separation proceeds, the elution strength of the mobile phase is increased to elute the more strongly retained sample components [14].

HPLC column hardware: A column tube and fittings must contain the chromatographic packing material (stationary phase) that is used to effect a separation. In general, three primary characteristics of chemical compounds can be used to create HPLC separations. They are: polarity, electrical charge and molecular size. Polarity has been previously described and the two main separation modes that exploit this characteristic are normal phase and reversed-phase chromatography.

HIC (hydrophobic interaction chromatography) is a type of reversed-phase chromatography that is used to separate large biomolecules, such as proteins. As part of the experiments described here, it was more convenient to use SEC (size exclusion chromatography) chromatography. It is usually desirable to maintain these molecules (IgG and IgG fragments) intact in an aqueous solution, avoiding contact with organic solvents or surfaces that might denature them. HIC takes advantage of the hydrophobic interaction of large molecules with a moderately hydrophobic stationary phase, *e.g.*, butyl-bonded (C4), rather than octadecyl-bonded (C18), silica. Initially, higher salt concentrations in water will encourage the proteins to be

retained on the packing. Gradient separations are typically run by decreasing salt concentration. In this way, biomolecules are eluted in order of increasing hydrophobicity [14].

In the 1950s, Porath and Flodin discovered that biomolecules could be separated based on their size, rather than on their charge or polarity, by passing, or *filtering*, them through a controlled-porosity, hydrophilic dextran polymer. This process was termed gel filtration [14]. Later, an analogous scheme was used to separate synthetic oligomers and polymers using organic-polymer packings with specific pore-size ranges. This process was called gel-permeation chromatography (GPC). Similar separations done using controlled-porosity silica packings were called size-exclusion chromatography (SEC).

For this project, SEC was selected to be used to separate IgG and IgG fragments. SEC has been applied successfully in the past to separate different sizes of proteins under native conditions [15]. It has been routinely used for the characterization and quality control of monoclonal antibody therapeutics in the pharmaceutical industry [15]. However, with normal SEC salt-containing mobile phases, direct MS analysis of antibody fragments is not feasible [15]. This makes relevant the case of using a no-salt mobile phase for this project. Aggregates, monomers and degradation products of mAbs can be separated on SEC columns based on their molecular size under native conditions [15].

SEC separations are typically done on stationary phases that have been synthesized with a pore-size distribution over a range that permits the analytes of interest to enter, or to be excluded from, more or less of the pore volume of the packing. Smaller-sized molecules penetrate more of the pores on their passage through the bed. Larger-sized molecules may only penetrate pores above a certain size so they spend less time in the bed and may be totally excluded from pores and pass only between the particles, eluting very quickly in a small volume.

Mobile phases are chosen according to two criteria. First, they must be good solvents for the analytes; second, they must prevent any interactions based on polarity or charge between the analytes and the stationary phase surface. In this way, the larger-sized molecules elute first, while the smaller-sized molecules travel slower because they move into and out of more of the pores and elute later, in decreasing order of their size in solution [14].

In gel electrophoresis, the term electrophoresis refers to the movement of charged molecules in response to an electric field, resulting in their separation. In an electric field charged proteins move toward an electrode of opposite charge. The rate at which they move (the mobility of proteins) is governed by a complex relationship between the physical characteristics of both the electrophoresis system and the proteins. Factors affecting gel electrophoresis include the strength of the electric field, the temperature of the system, the pH, ion type, concentration of the buffer, size, shape, and charge of the proteins[16]. Protein electrophoresis can be performed in either liquid or gel based media [16], in capillary and slab gel format. For this project, the slab gel based media was chosen, in the form of Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

Polyacrylamide gel electrophoresis (PAGE), describes a technique widely used in many scientific fields to separate biological macromolecules, usually proteins, according to their electrophoretic mobility. As with all forms of gel electrophoresis, molecules may be analyzed in their native state, preserving the molecules' higher-order structures. This method is called native-PAGE. Alternatively, a chemical denaturant may be added to remove these structures and turn the molecules into denatured proteins, whose mobilities depend only on length and mass-to-charge ratio. This procedure is called SDS-PAGE. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a method of separating molecules based on the difference of

their molecular weight. The SDS molecules are negatively charged and bind to proteins in a set ratio, approximately one molecule of SDS for every 2 amino acids [17]. The sample to analyze is optionally mixed with a chemical denaturant if so desired, usually SDS for proteins or urea for nucleic acids. SDS is an anionic detergent that denatures secondary and non–disulfide–linked tertiary structures, and additionally applies a negative charge to each protein in proportion to its mass. Heating the samples to at least 60 °C further promotes denaturation [18]. In addition to SDS, proteins may optionally be briefly heated to near boiling in the presence of a reducing agent, such as dithiothreitol (DTT) or 2-mercaptoethanol (beta-mercaptoethanol, BME), which further denatures the proteins by reducing disulfide linkages, thus overcoming some forms of tertiary protein folding, and breaking up quaternary protein structure (oligomeric subunits). This is known as reducing SDS-PAGE. The gels typically consist of acrylamide, bisacrylamide, the optional denaturant (SDS or urea), and a buffer with an adjusted pH. The solution may be degassed under vacuum to prevent the formation of air bubbles during polymerization. Alternatively, butanol may be added to the resolving gel (for proteins) after it is poured, as butanol removes bubbles and makes the surface smooth. The polymerization reaction creates a gel because of the added bisacrylamide, which can form cross-links between two acrylamide molecules. Gels are usually polymerized between two glass plates in a gel caster, with a comb inserted at the top to create the sample wells. After the gel is polymerized the comb can be removed and the gel is ready for electrophoresis [18]. Following electrophoresis, the gel is stained with imperial protein stain, allowing visualization of the separated proteins, or processed further (e.g. Western blot) [18]. After staining, different species of biomolecules appear as distinct bands within the gel. It is common to run molecular weight size markers of known molecular weight in a separate lane in the gel to calibrate the gel and determine the approximate

molecular mass of unknown biomolecules by comparing the distance traveled relative to the marker. PAGE may also be used as a preparative technique for the purification of proteins [19].

1.1.5 Mass Spectrometry Mass spectrometry (MS) is an analytical technique that ionizes chemical species and sorts the ions based on their mass-to-charge ratio. MS to identify and characterize biological molecules is a fundamental technology in protein biochemistry and proteomic analysis. MS can be performed on mixtures and MS/MS analyses are performed on the individual peptides, and the information is then analyzed to reveal the protein identity and/or its structural characteristics (e.g., co-and post-translational modifications (PTMs) or isoforms) [5]. Key steps in this strategy include the preparation of the protein sample for digestion, and cleanup or desalting of the final peptide mixture prior to MS analysis by either MALDI-TOF-MS or LC-ESI-MS and further characterization by MS/MS. Careful sample handling of the protein and peptide levels is the key for successful analysis by MS. Therefore it is necessary to take great care, and in some instances optimize each step involved, to obtain the maximum amino acid sequence coverage of the protein(s) of interest and on PTMs [5] (refers to the covalent and generally enzymatic modification of proteins following protein biosynthesis). Though MS was discovered in the early 1900s, its scope was limited to the chemical sciences. However, the development of electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) in 1980s increased the applicability of MS to large biological molecules like proteins. In both ESI and MALDI, peptides are converted into ions by either addition or loss of one or more protons. Both are examples of "soft ionization" methods where ion formation does not lead to a significant loss of sample integrity (sample contamination and degradation). MALDI-TOF MS has certain advantages over ESI-MS viz. (i) MALDI-TOF MS produces singly charged ions,

minimizing spectral complexity, thus interpretation of data is easier relative to ESI-MS, (ii) for analysis by ESI-MS, prior separation by chromatography is required which is not needed for MALDI-TOF MS analysis. Some of the advantages of ESI-MS over MALDI-MS are (i) fast data acquisition rate; (ii) stable resolution across m/z range; relatively high resolution and mass accuracy; and (iii) ESI-MS can efficiently be interfaced with separation techniques prevents sample loss, thus enhancing resolution.

MALDI-MS, first introduced in 1988 by Hillenkamp and Karas, [20] and has become a widespread analytical tool for peptides, proteins, and most other biomolecules (oligonucleotides, carbohydrates, natural products, and lipids). The efficient and directed energy transfer during a matrix-assisted laser-induced desorption event provides high ion yields of the intact analyte, and allows for mass determination of compounds with high accuracy and subpicomole sensitivity [21]. MALDI provides for the non-destructive vaporization and ionization of both large and small biomolecules. In MALDI analysis, the analyte is first co-crystallized with a large molar excess of a matrix compound, usually a UV-absorbing weak organic acid, after which laser radiation of this analyte–matrix mixture results in the vaporization of the matrix which carries the analyte with it. The matrix therefore plays a key role by strongly absorbing the laser light energy and causing, indirectly, the analyte to vaporize. The matrix also serves as a proton donor and receptor, acting to ionize the analyte in both positive and negative ionization modes, respectively [22]. It is generally thought that ionization occurs through proton transfer or cationization. The ionization depends critically on the matrix–analyte combination, but is not critically dependent on the number of acidic or basic groups of the analyte [23]. This suggests that a more complex interaction of analyte and matrix, rather than simple acid–base chemistry, is responsible for ionization.

There are four types of mass analyzers typically used with the MALDI ionization source: a TOF reflector, a fourier transform mass analyzer, hybrid quadrupole-TOF [24] and orbitrap instruments. The linear TOF mass analyzer is the simplest of the three devices and has enjoyed a renaissance with the invention of MALDI. TOF analysis is based on measuring the time needed for a set of ions accelerated using the same kinetic energy to reach a detector. Because the ions have the same energy, yet a different mass, the ions reach the detector at different times. The smaller ions reach the detector first because of their greater velocity while the larger ions take longer owing to their larger mass and low velocity. Hence, the analyzer is called TOF because the mass is determined from the ions' time of flight. The arrival time at the detector is dependent upon the mass, charge, and kinetic energy (KE) of the ion. However, ions are accelerated with a distribution of KE values, contributing to peak widening in the linear MALDI-TOF mode. To circumvent this problem, the TOF reflector combines TOF technology with an electrostatic analyzer, the reflector [24]. The reflector serves to increase or decrease the amount of time that ions need to reach the detector while reducing the temporal distribution at the detector and, increasing resolution. This increased resolution, however, often comes at the expense of sensitivity (measure only of signal magnitude for peak heights) and reduced mass range (detection limit) [24].

A MALDI-based approach that provides amino acid sequence information on peptides relies on the fragmentation of the ions in the field-free drift region (between the ion source and the detector) of the TOF mass analyzer. This fragmentation, referred to as post source decay (PSD), is due to collisions between analyte ions and neutral matrix molecules or residual gas molecules during the desorption and acceleration stage [24]. Initially these fragments have the same velocities as their precursor ion. Consequently, in a linear TOF mass analyzer, these

fragments arrive at the detector at the same time as their precursor ion and are therefore not represented in the mass spectrum. In a reflector TOF mass analyzer, however, these fragments which have lower KE value than the precursor require lowering the reflector voltages to be properly reflected towards the detector. By varying the reflector voltage as a function of the fragment mass, a full MS/MS spectrum can be obtained and valuable structural information is obtained. Amino acid sequencing by PSD has been a popular method for sequencing peptides and is now still used in an automated fashion on most commercial MALDI reflector instruments [24]. All MS analyses in this thesis were performed on an UltrafleXtreme™ mass spectrometer (Bruker Daltonics) equipped with LID-LIFT™ technology for tandem MS experiments. This flexible instrument has linear only mode, high resolution reflector mode and the latest TOF/TOF technology (LIFT™) which enables the use of various MS/MS techniques (e.g. LID, high energy CID) for fast and sensitive MS/MS experiments.

An LC-ESI-MS is a liquid chromatography (LC) system coupled with a tandem mass analyzer. Electrospray ionization (ESI) is an ionization technique where a sample solution is sprayed into a strong electric field, resulting in highly charged droplets in the presence of nitrogen to assist desolvation. The charged droplets formed evaporate in a region under atmospheric pressure, and are then directed to a zone maintained under vacuum causing solvent evaporation and the charge density to increase on the droplets. The multiply charged ions then enter the analyzer [25]. The most obvious feature of an ESI spectrum of a protein is that the ions carry multiple charges, which reduces their mass-to-charge ratio compared to a singly charged species. This allows mass spectra to be obtained for large molecules at relatively low *m/z* ranges. When analyzing with LC/MS/MS, the LC separates compounds by chromatography on a column. After the compounds elute from the column, they enter the ESI source, where ionization

takes place. Ions are filtered on the basis of their mass-to-charge ratio (*m/z*). Ions below and above a certain *m/z* value will be filtered out depending on the ratio of the direct current (DC) and alternating current (AC) voltages, applied to the quadrupole. By ramping the voltages on each set of poles, a complete range of masses can be channeled through the collision cell in MS/MS (tandem mass spectrometry) [25]. The Acquity UPLC console contains four major parts: binary solvent manager, sample manager, column manager options, and detector options. In the binary solvent manager (solvent mixer), there are two independent pump systems, A and B. Each pump delivers precise flow of a single solvent. The two pump systems combine their two solvents at a filter/tee mixer. From there, the solvent mixture flows to the sample manager. A gradient elution program is commonly used so that the eluent composition (and strength) is steadily changed during the analysis. This increases separation efficiency, decreases the retention time and improves peak shape by minimizing tailing. Temperature used 50 °C. Reverse phase columns are more commonly used in UPLC [25].

The Acquity UPLC can be configured with different detectors (e.g. UV, light scattering) but here, the ESI-MS/MS analyzer was used and is explained in more details as follows. The G2 synapt is a high performance tandem quadrupole mass spectrometer designed for routine LC-MS/MS operation. After a sample is introduced at atmospheric pressure, into the ionization source and ionized, ions are filtered by *m/z* at the first quadrupole, and the mass-separated ions undergo collision-induced decomposition in the collision cell, then the fragment ions are analyzed by *m/z* ratio at the second quadrupole. The MS1 mode is directly analogous to a single quadrupole MS. The MS2 mode is MS\MS used with collision gas. The instrument can quickly switch between MS and MS/MS operation during data acquisition, which allows MS1 and MS2 results to be obtained in one experiment [25].

Tandem mass spectrometry, also known as MS/MS, involves multiple steps of mass spectrometry selection, with some form of fragmentation occurring in between the stages [26]. In a tandem mass spectrometer, ions are formed in the ion source and separated by mass-to-charge ratio in the first stage of mass spectrometry (MS1). Ions of a particular mass-to-charge ratio (precursor ions) are selected and fragment ions (product ions) are created by collision-induced dissociation, ion-molecule reaction, photodissociation, or other process. The resulting ions are then separated and detected in a second stage of mass spectrometry (MS2).

When tandem MS is performed with a quadrupole mass analyzer as is incorporated in the ESI-MS instrument used for this research, the instrument must operate in one of a variety of modes. There are a number of different tandem MS/MS experimental setups and each mode has its own applications and provides different information. Mentioned here are four main scan experiments possible using MS/MS: precursor ion scan, product ion scan, neutral loss scan, and selected reaction monitoring. For a precursor ion scan, the product ion is selected in the second mass analyzer, and the precursor masses are scanned in the first mass analyzer. In a product ion scan, a precursor ion is selected in the first stage, allowed to fragment and then all resultant masses are scanned in the second mass analyzer and detected in the detector that is positioned after the second mass analyzer. This experiment is commonly performed to identify transitions used for quantification by tandem MS. In a neutral loss scan, the first mass analyzer scans all the masses. The second mass analyzer also scans, but at a set offset from the first mass analyzer [27]. This offset corresponds to a neutral loss that is commonly observed for the class of compounds. In a constant-neutral-loss scan, all precursors that undergo the loss of a specified common neutral are monitored. To obtain this information, both mass analyzers are scanned simultaneously, but with a mass offset that correlates with the mass of the specified neutral.

Similar to the precursor-ion scan, this technique is also useful in the selective identification of closely related class of compounds in a mixture. In selected reaction monitoring, both mass analyzers are set to a selected ion mass in the first stage of a tandem mass spectrometer and an ion product of a fragmentation reaction of the precursor ion is selected in the second mass spectrometer stage for detection. This mode is analogous to selected ion monitoring for MS experiments. A selective analysis mode, which can increase sensitivity [28].

Tandem mass spectrometry can be used for protein sequencing [29]. When intact proteins are introduced to a mass analyzer, this is called "top-down proteomics" and when proteins are digested into smaller peptides and subsequently introduced into the mass spectrometer, this is called "bottom-up proteomics". Shotgun proteomics is a variant of bottom up proteomics in which proteins in a mixture are digested prior to separation and tandem mass spectrometry. Tandem mass spectrometry can produce a peptide sequence tag that can be used to identify a peptide in a protein database [30-32]. A notation has been developed for indicating peptide fragments that arise from a tandem mass spectrum [33]. Peptide fragment ions are indicated by a, b, or c if the charge is retained on the N-terminus and by x, y or z if the charge is maintained on the C-terminus. The subscript indicates the number of amino acid residues in the fragment. Although peptide backbone cleavage is the most useful for sequencing and peptide identification other fragment ions may be observed under high energy dissociation conditions. These include the side chain loss ions d, v, w and ammonium ions [34-35] and additional sequence-specific fragment ions associated with particular amino acid residues [36]. Oligosaccharides may be sequenced using tandem mass spectrometry in a similar manner to peptide sequencing [37]. Fragmentation generally occurs on either side of the glycosidic bond (b, c, y and z ions) but also under more energetic conditions through the sugar ring structure in a cross-ring cleavage (x

ions). Again trailing subscripts are used to indicate position of the cleavage along the chain. For cross ring cleavage ions the nature of the cross ring cleavage is indicated by preceding superscripts [38], [39].

Fragmentation of gas-phase ions is essential to tandem mass spectrometry and occurs between different stages of mass analysis. There are many methods used to fragment the ions and these can result in different types of fragmentation and thus different information about the structure and composition of the molecule. Energy induced fragmented is most often what is being used in a tandem mass spectrometry experiment. The energy required for dissociation can be added by photon absorption, resulting in ion photodissociation. When ultraviolet lasers are used the process is termed laser induced dissociation (LID), which leads to a large amount of fragmentation of biomolecules [40]. Energy can also be added to the ions, which are usually already vibrationally excited, through post-source collisions with neutral atoms or molecules, the absorption of radiation, or the transfer or capture of an electron by a multiply charged ion. Collision-induced dissociation (CID), also called collisionally activated dissociation (CAD), involves the collision of an ion with a neutral atom or molecule in the gas phase and subsequent dissociation of the ion [41-42].

1.2 Objectives/aims of this M.Sc. thesis:

The global objective is to develop a workflow for detailed structural characterization of Abs from different sources, using methods which involve MS.

1) More specifically, it is very difficult to obtain glycosylation and amino acid sequence information on the Fab portion of IgGs, especially within the variable portion. For this, the objective is to fragment antibodies into smaller fragments (Fab and Fc), by papain, Fabulous™ and Fabricator™ digestion, thus allowing further tryptic digestion and analysis of those fragments individually without interference from each other, so that their specific *N*-glycosylation can be determined. This will be developed with human and porcine IgG samples (collaboration with Université de Nantes).

2) Results from (1) will be compared to those obtained using a more conventional approach where whole antibodies are trypsinized into peptides and glycopeptides for determination of peptide sequences and PTMs. This will be applied to human and pig IgG samples as in (1), but also to modified bioengineered mAbs prepared for pharmaceutical purposes (collaboration with the University of Toronto).

1.3 Significance:

Human IgG

Several diseases and forms of cancers have been shown to have an impact on protein glycosylation. The development of protocols leading to possible biomarker diagnostic tests has been a rapidly evolving research field. In this research, antibodies from the serum of patients

affected with monoclonal gammopathy, a condition often linked to myeloma, will be analyzed. The more complete the characterization of these antibodies, the greater likelihood of being able to develop an accurate test (collaboration with Université de Nantes).

Porcine IgG

There have been reports on the mass spectrometric analysis of pig IgG in relationship with its use in a xenotransplantation context [27, 28]. It is important that the modifications of such porcine IgG glycoforms be analyzed extensively, to ensure that glycan structures are present or absent as desired. This is necessary in order to prevent any negative immune responses, as specific glycoforms influence antibody function and efficiency [1]. For this reason fragmented antibody glycosylation analyses are relevant. Although fragmentation of all Ig classes is possible, only procedures for fragmentation of mouse, rabbit, and human IgGs and IgMs have been well characterized [8]. For the first time, this study attempts to characterize porcine IgGs (wild-type and genetically modified for use in xenotransplantation) using enzymatic fragmentation techniques.

Bioengineered monoclonal antibodies

In production of mAbs, (which play a major role in the biotherapeutics and also have diagnostic and analytical applications), downstream glycan modification methods may be used to study particular glycan-receptor interactions. Ready-made fragments of antibodies are seldom offered commercially, because there is limited demand for any given item. For this reason, except for custom antibody production, fragmentation is an activity for each individual laboratory to perform for its specific needs. The Fab fragment retains the binding specificity of the parent antibody, and its separation from the Fc fragment is often desired for analytical,

diagnostic, or therapeutic purposes. For analytical purposes, monovalent Fab fragments simplify the study of binding properties compared with the divalent whole antibody, and Fab fragments are more easily crystallized than whole antibodies for structure determination [8]. The smaller Fab fragment (~50 kDa compared with the 150-kDa parent antibody) has greater penetration into tissues and a shorter half-life, making it a useful in vivo reagent for diagnostic imaging or delivery of cytotoxic conjugates [8].

Because of their smaller size as functional components of the whole molecule, antibody fragments offer several advantages over intact antibodies for use in certain immunochemical techniques and experimental applications [8]: Reduced nonspecific binding from Fc interactions (many cells have receptors that bind the Fc region); Ability to control Fc-binding to Protein A or Protein G in experiments involving immunoprecipitation and Western blotting; More efficient penetration of tissue sections, resulting in improved staining in immunohistochemistry (IHC)

Chapter 2

2.1 Materials

Human immunoglobulin G polyclonal antibody was purchased from Sigma-Aldrich (Ontario, Canada). Modified bioengineered mAbs prepared for pharmaceutical purposes were obtained from the University of Toronto (Dr. James Rini Lab). Pig IgGs were obtained from the Université de Nantes (Dr. Jean- Paul Soulillou Lab). Trypsin Ultra™ was purchased from Promega (Wisconsin, USA). Fabulous™ (SpeB) and Fabricator™ enzymes, were donated by and purchased respectively from Genovis (Cambridge, MA, USA). Ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide (IAA), trifluoroacetic acid (TFA) were purchased from Sigma-Aldrich. Mini-Protean TGX precast gels (4-15%), Precision Plus™ protein standard, 2-mercapto-ethanol and 4x Laemmli sample buffer were obtained from Bio-Rad (Hercules, CA). Imperial protein stain was purchased from Thermo-Scientific (Rockford, IL). Acetonitrile (ACN) was purchased from EMD Millipore (Darmstadt, Germany). Strata-X C-18 cartridges were purchased from Phenomenex (Torrance, CA). Sodium phosphate dibasic anhydrous was purchased from McArthur Chemical Co. Ltd. (Montreal, QC, Canada). L- Cysteine was purchased from Sigma-Aldrich (St. Louis, MO). Hydrochloric acid was obtained from Anachemia (Vancouver, BC, Canada). Tris base (2-amino-2-(hydroxymethyl)-1,3-propanediol) was purchased from Fisher Scientific (Bridgewater NJ). Water was obtained from a Millipore filtration system from MilliporeSigma (Loveland, CO). Immobilized papain-cross linked, 6% in beaded agarose supplied as 50% glycerol in sodium acetate pH 4.5 was purchased from Thermo Scientific (Rockford, IL).

2.2 Experimental

2.2.1 In-solution tryptic digestion (oven incubation)

Whole antibodies were subject to tryptic digestion to produce peptides and glycopeptides. IgGs (50μg) were first evaporated to dryness and reconstituted in 100 uL of 50 mM ammonium bicarbonate solution. Trypsin (1 μg) was added and the mixture was vortexed, then samples were placed in an oven for ~18h at 37ºC for digestion to occur. Digestion was stopped by placing samples in a freezer for ~10 minutes.

2.2.2 Fractionation of glycopeptides using C18 reverse phase cartridge

Each Strata-XTM C18 cartridge was first washed with 1:1 ACN–water and then conditioned with 0.1% TFA in ACN and then 0.1% TFA in water. Each sample was applied onto the column and washed with 0.1% TFA in water. The following elution steps were then performed: 5x 200 μL of 0.1% TFA in 10:90 ACN–water, 5x 200 μL of 0.1% TFA in 15:85 ACN– water, 1x 1 mL of 0.1% TFA in 20:80 ACN–water, and 1x 1 mL of 0.1% TFA in 30:70 ACN–water. All eluates were dried to be further analyzed using MALDI-TOF-MS on a Bruker UltraFleXtreme™ (see instrument description in Section 1.1.5).

2.2.3 Papain digestion (oven incubation) for both human and pig IgG

A. Preparation for papain digestion of human IgG

Just before use, the digestion buffer was prepared by adding 10 mM cysteine and HCl to a final concentration of 20 mM to the 20 mM sodium phosphate buffer and the pH was adjusted to 7.0. The immobilized papain was mixed by inversion or gentle shaking to obtain an even suspension. The immobilized papain slurry (10 µL, 50%) was added to an Eppendorf™ tube. A pipette tip was cut to ensure proper gel slurry dispensing. The gel slurry was equilibrated by addition of 80 µL of digestion buffer to the gel slurry. The gel was separated from the buffer by centrifugation. Washing was repeated once with another 80 µL of buffer. Both washes were discarded. The gel was resuspended in 10 µL of digestion buffer. A specified amount of the immobilized papain slurry (0.5 mL, 50%) was required for 10 mg of antibody. It is important to note that in this thesis, amounts were much smaller than 10 mg, i.e. on the order of 200 µg, and that all proportions needed readjusting. The use of immobilized papain with such small amounts has not been reported in the literature so far [8].

Preparation for papain digestion of pig IgG

Just before use, 20 mM sodium phosphate was prepared by adding 10 mM cysteine and HCl to a final concentration of 20 mM to the sample buffer and the pH was adjusted to 7.0. The immobilized papain was mixed by inversion or gentle shaking to obtain an even suspension. The immobilized papain slurry (20 µL, 50%) was added to an Eppendorf™ tube. A pipette tip was cut to ensure proper gel slurry dispensing. The gel slurry was equilibrated by addition of 160 µL of digestion buffer. The gel was separated from the buffer by centrifugation. The wash was repeated once with another 160 µL of buffer. Both washes were discarded. The gel was resuspended in digestion buffer.

B. Generation of fragments for human and pig IgG

Pure lyophilized human or non-lyophilized IgG (200 µg) was dissolved in digestion buffer. This solution was added to the tube that contained the washed immobilized papain and digestion buffer. The sample was incubated for ~24 h at 37°C. Constant mixing of the slurry gel was pursued during incubation by manually pipetting the sample every 30 min for at least 6 h for human IgG. For pig IgG, constant mixing was done by vortexing in the oven for ~24 h at 37°C. The immobilized enzyme was separated from the digest by centrifugation, and 20 µL of 10 mM Tris•HCl, pH 7.5 were added to the digest before centrifugation. The supernatant, which contained the IgG fragments, was then removed, and section 2.3.2 was carried out.

2.2.4 Fabricator™ digestion for human IgG only

Pure lyophilized human IgG (200µg) samples were dissolved in 200 µL of 50mM ammonium bicarbonate buffer at pH 6.8. Antibody sample solution was added to 200 units of Fabricator™ enzyme. Each sample was vortexed and placed for incubation in an oven at 37°C for 2 h, after which Section 2.3.2 was carried out.

2.3.1 Fabulous™ digestion for pig IgG only

Each pig IgG sample (200 µg) was added to 200 units of Fabulous ™ enzyme in 200 µL of 10 mM Tris, 50 mM cysteine buffer. Each sample was vortexed and incubated in an oven for 1 h, then Section 2.3.2 was carried out

2.3.2 Fractionation of IgG fragments by HPLC using a SEC column

The digestion mixture was injected into a preconditioned SEC- 300 4.6x300 mm silica based column (Sigma-Aldrich, MO). The mixture was eluted with a mobile phase of 0.1% TFA, 0.1% formic acid in 20% ACN at a flow rate of 0.3 mL/min. The HPLC system used was a Waters 1525 binary pump equipped with a Waters 2707 autosampler and a Waters 2998 photodiode array detector (Waters, Milford, MA). Fractions were collected, and dried, for MALDI-MS analysis.

2.2.3 Separation of IgG fragments by SDS polyacrylamide gel electrophoresis

Fab and Fc components were separated on a Mini-PROTEAN® Tetra Cell system (Bio-Rad). The Bio-Rad TGX^{TM} gels used were 10-well, gradient (4%–15%). Wells were washed individually four times with running buffer (10 × tris–glycine–SDS buffer diluted 1 × with water) prior to the loading samples. Each sample fraction containing Fab, Fc or both Fab and Fc had its own lane on the gel. Each gel was loaded with 15 μL of each fraction (in water) in 11.3 μL of 4 × Laemmli sample buffer and no 2-mercaptoethanol was added. Well 1 was loaded with 10 μL of Precision Plus Protein ^{TM} Kaleidoscope ^{TM} standard. Intact reduced IgG (15 μL, ~14 μg) was loaded into well 2. IgG fragments (~14 μg) were loaded on the gel. Running buffer was poured in the cell system and the voltage was set at 150 V. Samples were allowed to migrate for 40 min, until the dye front reached the bottom of the gel. The gel was removed from the cell and rinsed four times with Millipore water and sufficient Imperial^{TM} protein stain was added. IgG fragments absorbed the stain overnight and the stain was decanted and replaced with water until gel bands became visible.

2.3.4 In gel tryptic digestion of IgG Fab and Fc fragments

Reduced heavy chain, reduced light chain, non-reduced Fab and Fc bands were then excised from the gel. Tryptic digestion was done on each single cut out band prior to digestion. The bands were chopped into smaller pieces and placed into 1.5 mL Eppendorf tubes. The digestion buffer was 50 mM ammonium bicarbonate in water. The wash buffer was 50% of 50 mM ammonium bicarbonate in ethanol. Each tube contained one (1) lane's worth of gel. Gel pieces were washed with wash buffer (1:1 digestion buffer–ethanol) until all protein stain was removed and gels were incubated in absolute ethanol for 10 min. Gel pieces were then washed with digestion buffer for 20 min and then incubated in absolute ethanol for 20 min, which was removed from the gel by vacuum centrifugation. Trypsin solution was added and the tubes were placed on ice where the gel was allowed to swell. Thereafter, excess trypsin solution was discarded. Gel pieces was covered with digestion buffer and incubated at 37 °C for ~18h. Glycopeptides and peptides were extracted with buffer A: 0.5% acetic acid, extraction buffer B: 5:3 30% ACN – 0.5% acetic acid, and extraction buffer C 100% ACN. Samples were dried down for further analysis.

2.4.1 PNGase F removal of glycans from trypsin diegested Fab and Fc fragments

PNGase F solution (4 µL, 10 units/µL) was added to Ab solution of tryptic digested glycopeptides. The sample was vortexed and set at 37^0C for ~18 h.

C18 cartridge to isolate glycans from the deglycosylated protein.

Directly after the digest, the glycans were separated from the deglycosylated protein on a C18 cartridge. The cartridge was conditioned with 5 x 1 mL (ACN + 0.1 % TFA) and then 5 x 1

mL (H_2O + 0.1 % TFA). The sample was added and allowed to equilibrate for about 1-2 min. The glycans were eluted with 3 mL H_2O + 0.1 % TFA and collected in two fractions. If desired the protein could be eluted with 1.5 mL of 50:50 ACN:H2O + 0.1% TFA. The solvent was evaporated from the fractions. In preparation for MALDI analysis, the glycan samples were resuspended in 20 uL of Millipore water. A solution of 0.5 uL of DHB matrix was spotted on the target followed right away by addition of 0.5 uL of sample solution and mixing on target. Experiments were conducted in reflector positive mode MALDI-MS on a Bruker UltraFleXtreme instrument (Bruker, Billerica, MA).

2.4.2 Preparation of samples for MALDI MS analysis

Glycopeptides fractions were mixed directly with 5 μL of DHB in 30:70 ACN: water. The samples (1 μL) were then spotted onto the stainless steel target and allowed to dry, for reflector positive mode MALDI-MS. For the Fab and Fc fragments from HPLC fractions, 5 μL of 0.1% TFA and 5 μL of sinapinic acid in 0.1% TFA 30:70 ACN: water were added and 1 μL was spotted onto the target already pre-spotted with 0.5 μL of sinapinic acid in ethanol. Spots were then allowed to dry for linear positive mode MALDI-MS.

2.4.3 MALDI-MS analysis

All MS analyses were performed on an UltrafleXtreme^{TM} mass spectrometer (Bruker Daltonics, Billerica, MA) equipped with LID-LIFT^{TM} technology for tandem MS experiments. For positive ionization mode, a nine-peptide calibration mixture with masses ranging from 500 to

5000 Da was used. In linear positive mode, the tryptic peptides of cytochrome C calibration mixture, with mass range from 10000 Da to 160000 Da was used.

2.4.4 Preparation of samples for UPLC ESI-MS analysis

For the Fab and Fc tryptic digests fragments, 100 µL of water was added to the combined fractions of Fab and Fc respectively and then they were sonicated, for UPLC MS analysis.

2.5 UPLC ESI-MS analysis

The digestion mixture (50 µL) was injected into a preconditioned Acquity BEH C18 (1.7 µm, 2.1 X 50 m) silica based, reverse phase column (Waters, Mississauga, ON). The mixture was eluted with a mobile phase gradient, composed of: A: 0.1% formic acid (v/v) in water, B: 0.1% FA (v/v) in 50% acetonitrile (ACN), at a flow rate of 0.25 mL/min. Fab and Fc fragment fractions were analyzed on a Waters Acquity UPLC with a ACQUITY BEH C18 C18(1.7 µm, 2.1 X 50 m) column with MS detection done on a Waters G2 synapt ESI-MS. This instrument is a tandem quadrupole model. Positive ionization mode was used (ESI+), and the ESI capillary voltage was 3.00 kV. The cone voltage was set at 25 V. Progenesis^{TM} software handled all UPLC ESI-MS analyses.

Chapter 3

Results and discussion

3.1 Analysis of tryptic digests of unfragmented human and porcine IgG for *N*-glycosylation by MALDI-MS

3.1.1 General analysis of antibody *N*-glycosylation with trypsin (glycoproteomics approach) for human monoclonal antibody constructs.

In solution tryptic digestion was carried out on intact antibodies. Two monoclonal recombinant antibodies (human IgG1) were obtained from Dr. James Rini's laboratory, University of Toronto. These constructs were used as controls to verify the efficacy of the method used. Figure 3.1.1 shows positive ion, reflector mode MALDI-TOF-MS spectra of tryptic glycopeptide fractions of constructs of Her2f and Her2FZ human based IgGs showing glycoforms G0F, G1F, and G2F as seen in Fig. 1.1.2. In Figure 3.1.1 a) and b) two peptide sequences are observed, 1) (EEQYNSTYR) with the three main glycoforms G0F, G1F, and G2F (*m/z* 2634, 2796, 2958 respectively), and 2) (TKPREEQYNSTYR), resulting from a missed tryptic cleavage, with the same glycoforms observed. These glycoforms are at *m/z* 3116, 3278, and 3440 respectively. These spectra suggest that the tryptic digestion and C_{18} cartridge processes are rather satisfactory. Although the C18 process was meant to enrich glycopeptides,

other tryptic peptides from both Fc and Fab portions are also observed as indicated in Table 3.1 below.

Table 3.1. Human Fab and Fc determined from MS analyses. Values of *m/z* correspond to monoisotopic $[M+H]^+$ ions.

Human Fab peptides, measured *m/z*	**Theoretical *m/z***	**Mass Error(ppm)**	**Analytical technique**	**Amino acid sequence**
788.3882	N/A	N/A	UPLC	From heavy chain
1186.645	1186.646	-0.842	UPLC-MS	GPSVFPLAPSSK
1690.252	N/A	N/A	MALDI-MS	From heavy chain
1740.992	N/A	N/A	MALDI-MS	From heavy chain
1818.437	N/A	N/A	MALDI-MS	From heavy chain
2135.9698	2135.9680	0.8427	MALDI-MS, UPLC-MS	VDNALQSGNSQESVTEQDSK
Human Fc peptides, measured *m/z*	**Theoretical *m/z***	**Mass Error(ppm)**	**Analytical technique**	**Amino acid Sequence**
575.340	575.339	1.7381	UPLC-MS	LTVDK
632.3716	632.372	-0.6325	UPLC-MS	TVERK
838.4992	838.503	-4.5322	UPLC-MS, MALDI-MS	ALPAPIEK
1104.6088	1104.608	0.7242	MALDI-MS, UPLC-MS	NQVSLTCLVK
1267.9912	N/A	N/A	MALDI-MS	N/A
1298.0102	N/A	N/A	MALDI-MS	N/A
1677.8028	1677.801	1.0728	MALDI-MS, UPLC-MS	FNWYVDGVEVHNAK
1690.241	N/A	N/A	MALDI-MS	From heavy chain
1793.434	N/A	N/A	MALDI-MS	From heavy chain
1808.4582	N/A	N/A	MALDI-MS	From heavy chain
1873.424	N/A	N/A	MALDI-MS	Form heavy chain

2601.052	2601.0488	1.2302	UPLC-MS, MALDI-MS	EEQFNSTYR + G0F
2634.0496	2634.0469	1.0250	UPLC-MS, MALDI-MS	EEQYNSTYR +G0F
2763.106	2763.1016	1.5924	UPLC-MS, MALDI-MS	EEQFNSTYR + G1F
2796.0928	2796.0997	-2.4677	UPLC-MS, MALDI-MS	EEQYNSTYR + G1F
2925.147	2925.1544	-2.5297	UPLC-MS, MALDI-MS	EEQFNSTYR + G2F
2958.1528	2958.1525	0.1014	UPLC-MS, MALDI-MS	EEQYNSTYR + G2F
3054.202	3054.1970	1.6370	UPLC-MS	EEQFNSTYR + G1FS1
3216.249	3216.2498	-0.2487	UPLC-MS	EEQFNSTYR + G2FS1
3249.253	3249.2479	1.5695	MALDI-MS, UPLC-MS	EEQYNSTYR + G2FS1

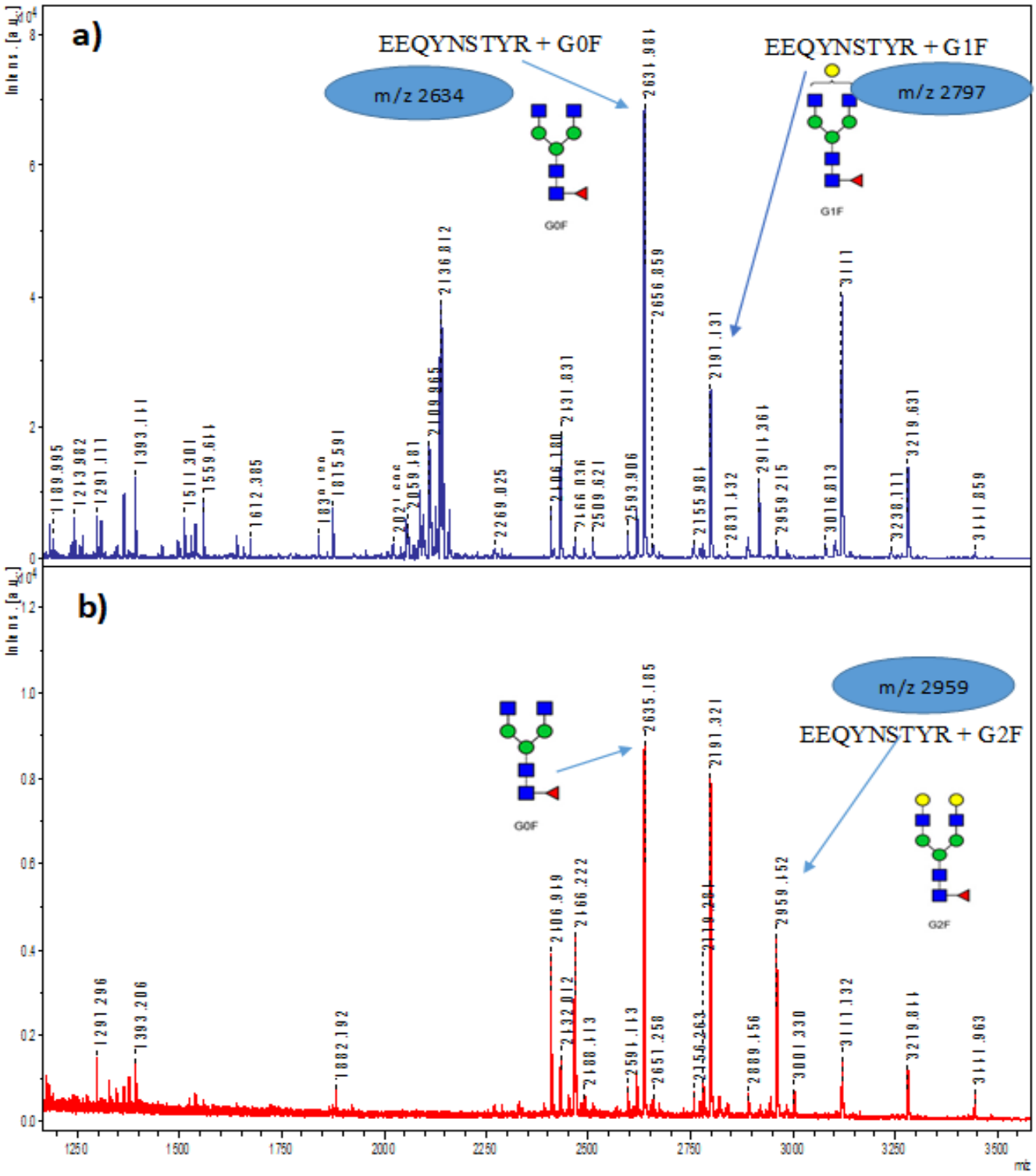

Figure 3.1.1. Reflector positive mode MALDI-TOF spectra of the tryptic digestion products of intact a) Her2F and b) Her2FZ mAb constructs based on human IgG1 (see Table 3.1).

3.1.2 General analysis of antibody *N*-glycosylation with trypsin (glycoproteomics approach) for pig IgG

The general analysis seen from the whole tryptic digestion of the pig IgG shows peptides and glycopeptides of whole IgG, and no distinguishing info on where in the IgG glycoslation is occurring or on the origin of peptides within the antibody. Figure 3.1.2 shows positive ion, reflector mode MALDI-ToF-MS spectra of tryptic glycopeptide fractions of pig IgG showing glycoforms G0F, G1F, and G2F as seen in Fig. 1.1.2. In Figure 3.1.2, two peptide sequences are observed, 1) (EEQFNSTYR) with the three main glycoforms G0F, G1F, and G2F (see Fig. 1.1.2 for nomenclature) (*m/z* 2618, 2780, 2942 respectively), and 2) (EAQFNSTYR), in which G0F, G1F and G2F are observed as *m/z* 2560, 2722, and 2884 respectively. These spectra suggest that the digestion process is satisfactory. As for human IgG the C18 cartridge fractionation process was aimed at enriching glycopeptides, although tryptic peptides from both Fc and Fab portions were also observed, and cannot be clearly identified by this method. Table 3.2 identifies some of peptides observed in the spectra.

Table 3.2. Pig Fab and Fc determined from MS analyses. Values of *m/z* correspond to monoisotopic $[M+H]^+$ ions.

Pig Fab peptides, measured *m/z*	**Theoretical *m/z***	**Mass Error(ppm)**	**Analytical technique**	**Amino acid sequence**
517.2838	N/A	N/A	UPLC-MS	Unknown (from heavy chain)
806.3448, 807.556	N/A	N/A	UPLC-MS, MALDI-MS	Unknown (from light chain)

816.4430	N/A	N/A	UPLC-MS	Unknown (from heavy chain)
827.5109	N/A	N/A	UPLC-MS	Unknown (from light chain)
1242.806	N/A	N/A	MALDI-MS	Unknown (from light chain)
1302.609	N/A	N/A	UPLC-MS, MALDI-MS	Unknown (from light chain)
1340.548	N/A	N/A	UPLC-MS, MALDI-MS	Unknown (from light chain)
1687.123	N/A	N/A	MALDI-MS	Unknown (from light chain)
1801.950	N/A	N/A	UPLC-MS, MALDI-MS	Unknown (from light chain)
1910.241	N/A	N/A	MALDI-MS	Unknown (from light chain)
2063.373	N/A	N/A	MALDI-MS	Unknown (from light chain)
2092.889	N/A	N/A	UPLC-MS	Unknown (from light chain)
2338.645	N/A	N/A	MALDI-MS	Unknown (from light chain)
2377.626	N/A	N/A	MALDI-MS	Unknown (from light chain)
2508.905	N/A	N/A	MALDI-MS	Unknown (from light chain)
Pig Fc peptides, measured *m/z*	**Theoretical *m/z***	**Mass Error(ppm)**	**Analytical technique**	**Amino acid sequence**
882.4988	882.504	-5.8923	UPLC-MS, MALDI-MS	DLPAPITR
1347.5948	1347.592	2.0777	UPLC-MS, MALDI-MS	SNGQPEPEGNYR
1805.5212	N/A	N/A	MALDI-MS	From heavy chain

1959.9318	1959.933	-0.6122	UPLC-MS, MALDI-MS	TTPPQQDVDGTYFLYSK
2305.962	N/A	N/A	MALD-MS	From heavy chain
2268.9178	2268.9147	1.3666	UPLC-MS	EEQFNSTYR + G0F (without fucose and GlcNAc)
2560.0480	2560.0461	0.74217	UPLC-MS, MALDI-MS	EAQFNSTYR + G0F
2618.0516	2618.0520	-0.1527	UPLC-MS, MALDI-MS	EEQFNSTYR + G0F
2780.0998	2780.1048	-1.7984	UPLC-MS, MALDI-MS	EEQFNSTYR + G1F
3087.1898	3087.1947	-1.5872	UPLC-MS	EEQFNSTYR +G1FS1

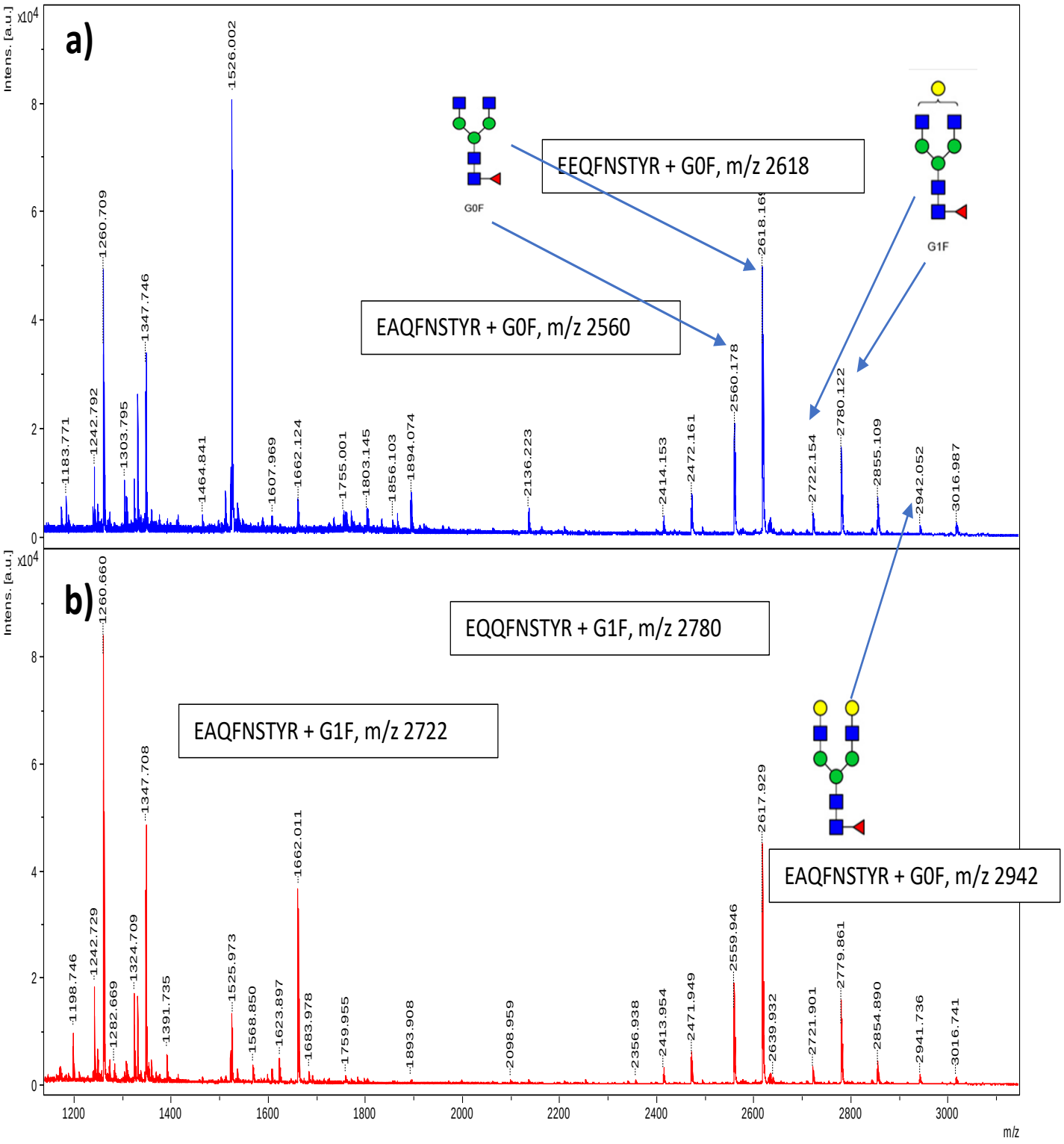

Figure 3.1.2. Reflector positive mode MALDI-TOF MS spectra of glycosylated fractions from digestion products of porcine IgG a) wildtype b) double knockout [43]. See Table 3.2 for peptide identification.

3.2 Papain fragmented analysis of antibody *N*-glycosylation for human and pig IgG by MALDI-MS

3.2.1 HPLC separation of Fab and Fc after papain digestion of human IgG

Size exclusion chromatography (SEC) has been widely used in protein analysis. Aggregates, monomers and degradation products of mAbs can be separated on SEC columns based on their molecular weights under native conditions [15]. The detector used with the HPLC system was Photo Diode Array (PDA) detects an entire spectrum simultaneously. PDA adds the third dimension (wavelength). This is convenient to determine the most suitable wavelength without repeating analyses. In this work, a small quantity of antibody remained undigested after incubation, according to HPLC-SEC results. When IgG molecules are incubated with papain in the presence of a reducing agent, one or more peptide bonds in the hinge region are split, producing three fragments of similar size: two Fab fragment and one Fc fragment [8]. When Fc fragments are of interest, papain is the enzyme of choice because it yields an intact 50,000-Dalton Fc fragment. The Fc may remain intact based on conditions and enzyme used [8].

In this experiment, using 200 μg of human IgG, the Fc along with the Fab remained intact after fragmentation. It is important to note that the use of immobilized papain with such small amounts has not been reported in the literature. The optimal pH of 7 for papain digestion was used in the experiment, with sodium phosphate buffer. The chromatogram (Figure 3.2.1) shows not so significant separation, but this by no means hampered the collection of Fab and Fc. Fractions were collected every 60 sec, which allowed to obtain Fab and Fc from the peaks' non-overlapping sections. The first major peak on Figure 3.2.1 was indicative of the Fc and the

second peak was identified as the Fab, as verified by gel electrophoresis and in-gel tryptic digestion. SDS-PAGE alone could not be used after fragmentation, as prior separation was required so that, visualization of bands on gel was possible. Later section 3.2.3, the gel results and in-gel tryptic digestion results of Fab and Fc will be discussed. Based on the order of fractions collected from the retention time of the peaks seen, the first major peak is Fc and the second is Fab. Later peaks result from buffers and were present in all SEC chromatograms.

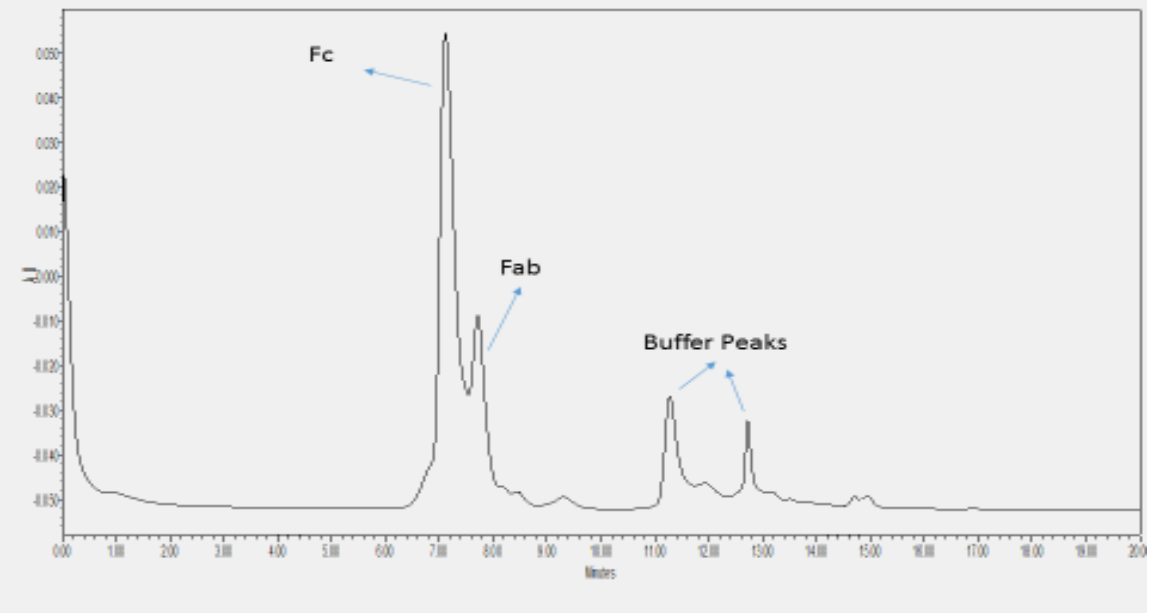

Figure 3.2.1. SEC-HPLC peaks attributed to Fab and Fc fragments after papain digestion of human IgG.

3.2.2 HPLC separated fractions analyzed by linear mode MALDI for human IgG

Figure 3.2.2 shows positive ion, linear mode MALDI-MS spectra of Fab and Fc portions of the antibody. In theory, the Fc fragment is larger in size and should elute first on the SEC column, which can be seen by MALDI in Figure 3.2.2a with the peak at *m/z* 53158 Da, present in the fraction (A3) collected before the one containing both Fab and a small amount of Fc (A4),

which was left over from fraction A3. Fraction A4 shows peaks of Fab at *m/z* 48101 Da and remaining Fc (larger) at *m/z* 53294 Da. The spectrum shows intact Fab and Fc fragments, which indicates that the papain enzyme digested IgG successfully. Peaks were small due to the low concentration of antibody used, but are still observable. These MS results after HPLC-SEC separation were used to verify that the fractions contained said fragments, confirming that the column was effective in eluting fragments based on their size, with bigger Fc fragment eluting first and Fab which is generally smaller, soon after. The closeness of peak elution is a result of the slight difference in size between fragments, which resulted in peaks not effectively being fully resolved from each other. Various variables have such as temperature of column oven, mobile phase and could obtain stationary phase better than that of SEC, due to the poor size required for such separation of Fab and Fc fragments.

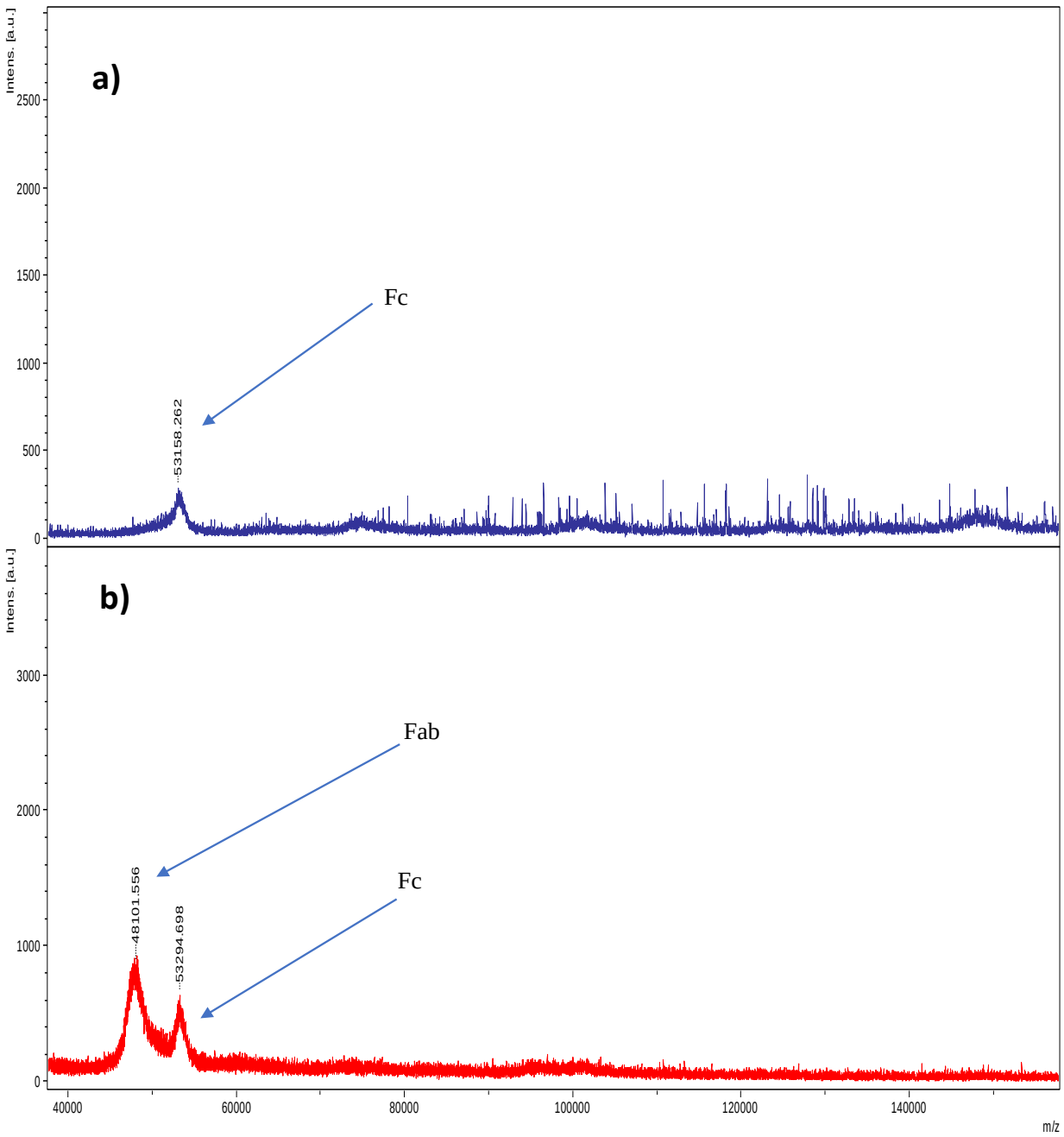

Figure 3.2.2. Linear Positive mode MALDI–TOF mass spectra of fraction a) A3 containing Fc only and b) A4 containing both Fab and small amount of Fc.

3.2.3 Gel separation of Fab and Fc of human IgG

If Fab and Fc were reduced before SDS-PAGE separation, the disulfide bonds which hold Fab and Fc chains together would be removed, and it would not be so easy to tell them apart, as the reduction would yield four ~25 kDa fragments. The non-reducing experiment allowed for individuality of the fragments to be determined when still intact, as 48 kDa and 50 kDa fragments.

After separation, the gel was stained and de-stained to reveal bands. Reduced heavy chain, reduced light chain, non-reduced Fab and Fc bands were then excised from the gel. Tryptic digestion was done on the cut bands. Cutting of bands from the gel helped to single out heavy and light chains, which were identified with their respective marker glycopeptides and peptides (Fc contains the conserved *N*-glycosylation site). The Fab and Fc bands were also distinguished and separated from each other. In lanes 5 and 9 (Figure 3.2.3), corresponding to the fractions containing both Fab and Fc, two distinguishable bands were obtained, in comparison with lanes 4 and 8, which had only the Fc band. Bands in lanes 5 and 9 have different molecular weights and as such correlate to two different fragments. The Fc is located at the top in lane 5 and 9 and the Fab band is located below it, which is indicative of a lower molecular weight. The results from the gel support the suggested contents of HPLC fractions which were viewed by MALDI MS and indicate that the Fc was eluted first by HPLC-SEC, then the Fab. The absence of low molecular weight fragments indicated that overdigestion with papain did not occur. The SDS-PAGE information also showed that fragmentation of Fab and Fc, is possible with a

1000µg Ab sample and also with an even smaller quantity of 200µg Ab sample.

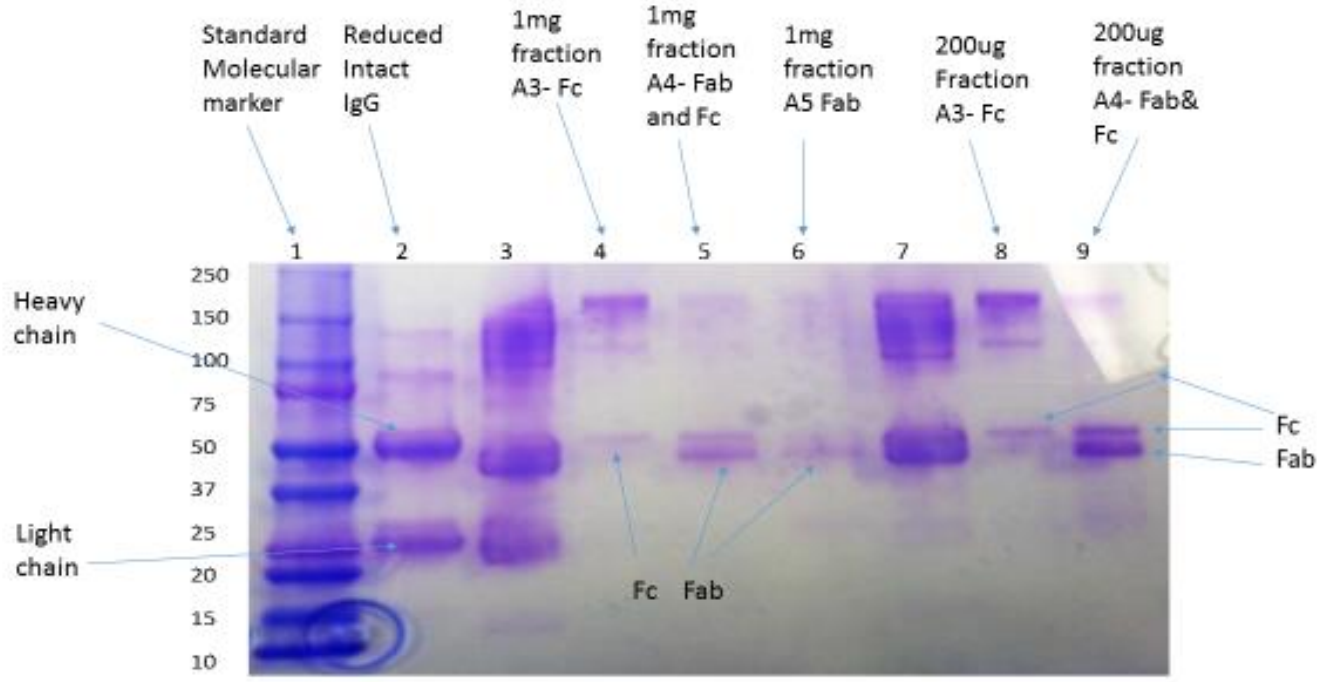

Figure 3.2.3 Fab and Fc HPLC-SEC fractions on acrylamide gel. The intact human IgG antibody was reduced, while the Fab and Fc fragments were not reduced.

3.2.4 In-gel tryptic digested Fab and Fc of human IgG

In-gel tryptic digestion was carried out on the gel bands corresponding to Fab, Fc, heavy and light chains in the oven at 37^{o}C overnight, and produced peptide and glycopeptide fractions which were collected straight from the digest in Eppendorf tubes (there was no clean up), and analysed by MS. Figure 3.2.4a shows a positive ion, reflector mode MALDI-ToF-MS spectrum of the tryptic digest of the Fc fragment of IgG. In Figure 3.2.4a, only one peptide sequence is observed (EEQYNSTYR) with the three main glycoforms G0F, G1F, and G2F (see Fig. 1.1.2) (*m/z* 2634, 2796, and 2958 respectively). Glycans are obviously present in the Fc region of IgG, also other peptides corresponding to the Fc domain are present. Known peptide masses were

retrieved from www.expasy.com, these can be seen in the additional information section, and the heavy chain was used to determine glycopeptides of the Fc region. Figure 3.2.4b below shows no obvious glycopeptides from the Fc region (see Table 3.1). It displays some non-glycosylated peptides related to the Fab region (see Table 3.1). This indicates total separation of Fab and Fc regions from each other, thus no interference in analysis as they are separate entities.

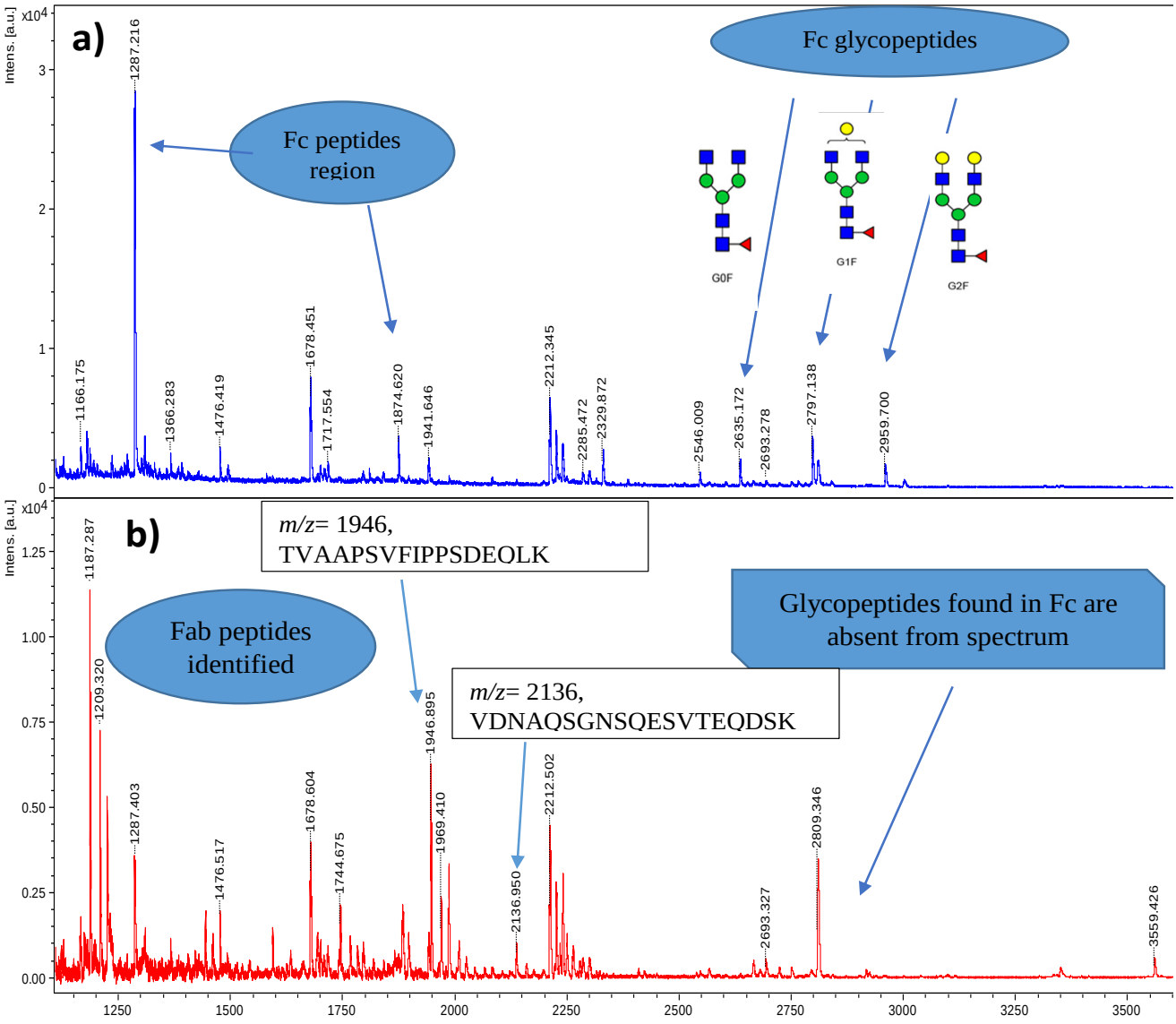

Figure 3.2.4. Reflector positive mode MALDI-ToF-MS spectra of in-gel tryptic products from the: a) Fc and b) Fab portion of commercial polyclonal human IgG. See Table 3.1 for peptide identification.

3.3 Digestion of pig IgG with papain

3.3.1 HPLC separation of Fab and Fc after papain digestion of pig IgG

Pig IgG was digested with immobilized papain for 24 h. The digestion mixture was analyzed using a HPLC SEC column. A small quantity of antibody remained undigested after incubation. This could be as a result of the pig IgG having a total eleven genomic sequences which correspond to the six subclasses of the IgG in the wild type species [44]. Eleven genomic porcine Cγ gene sequences are described that represent six putative subclasses that appear to have originated by gene duplication and exon shuffle. The genes previously described as encoding porcine IgG1 and IgG3 were shown to be the $IgG1^a$ and $IgG1^b$ allelic variants of the IGHG1 gene, IgG2a and IgG2b are allelic variants of the IGHG2 gene, while "new" IgG3 is monomorphic, has an extended hinge, is structurally unique, and appears to encode the most evolutionarily conserved porcine IgG [44].When IgG molecules are incubated with papain in the presence of a reducing agent, one or more peptide bonds in the hinge region are split, producing three fragments of similar size: two Fab fragment and one Fc fragment [8]. The Fc may remain intact depending on conditions and enzyme used [8].

In this experiment, using 200 μg of porcine IgG for digestion, the Fc did not remain intact, but the Fab remained intact as seen by MS and SDS-PAGE results. Figure 3.3.1 shows significant separation by HPLC-SEC, but as for human IgG this by no means hampered the collection of Fab and Fc. SEC collected fractions were run next on SDS gel. The first major peak on Figure 3.3.1 was indicative of the Fab and the second peak was identified as the Fc, as verified by gel electrophoresis and further tryptic digestion. This is due to the Fab fragment

being larger and eluting first due to Fc splitting, while the opposite is true for the case of the human IgG.

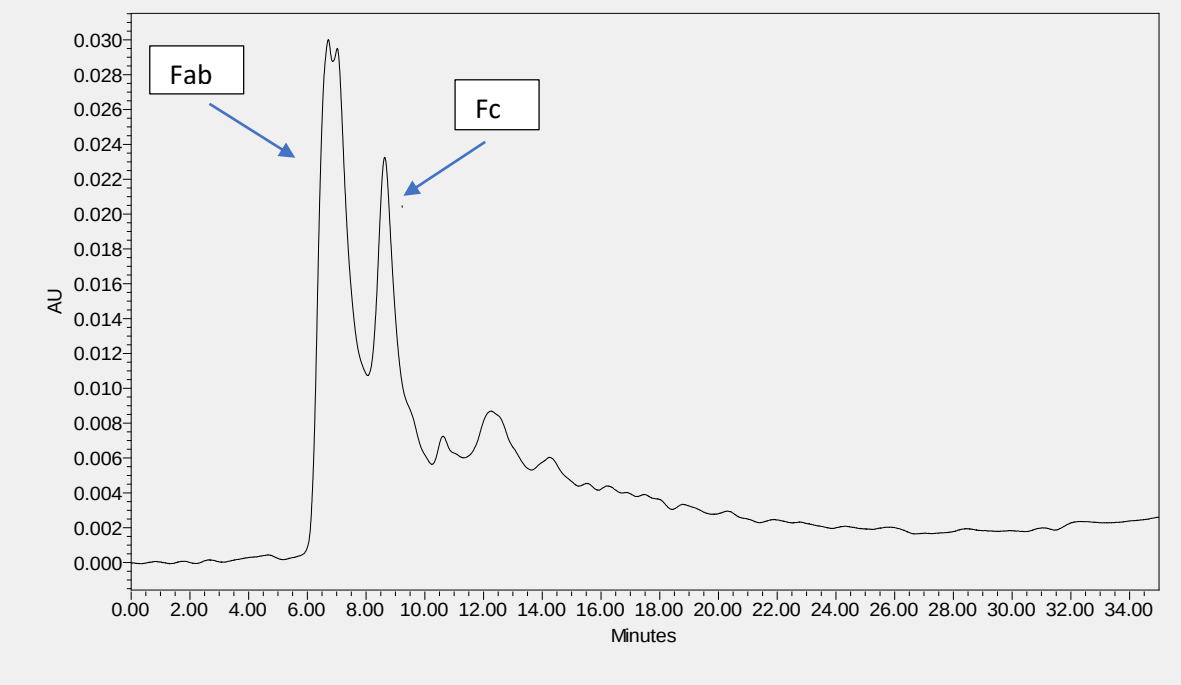

Figure 3.3.1. HPLC- SEC chromatogram for wildtype pig IgG digested with papain.

3.3.2 HPLC separated fractions analysed by MALDI for pig IgG

Figure 3.3.2 shows the positive ion, linear mode MALDI-MS spectra of Fab and Fc portions of the antibody. In theory the Fc fragment is larger in size and should elute first on the SEC column, but was split in half upon digestion, while the Fab remained intact for this experiment, which can be seen by MALDI-MS spectra in Figure 3.3.2 with the peaks at *m/z* 23639 Da and *m/z* 46486 Da present in the fraction (A3) which indicates the presence of both Fab and half the Fc. Fraction A4 shows peak of the other half Fc only, at *m/z* 24068 Da. Further

justified by SDS-PAGE. Usually in MALDI-MS linear positive mode, not having enough concentrated sample results in very low intensity peaks, as seen in figure 3.3.2. The spectrum shows intact Fab, but not intact Fc fragments, which indicates that the papain enzyme digested successfully. The complexity of porcine IgG, due to different subclasses and hinge region variation of such subclasses, caused the Fc to not remain intact under digestion conditions used previously for human IgG. These MALDI-MS results after HPLC-SEC separation were used to verify that the fractions contained said fragments, confirming that the column was effective in eluting fragments based on their size (see next section).

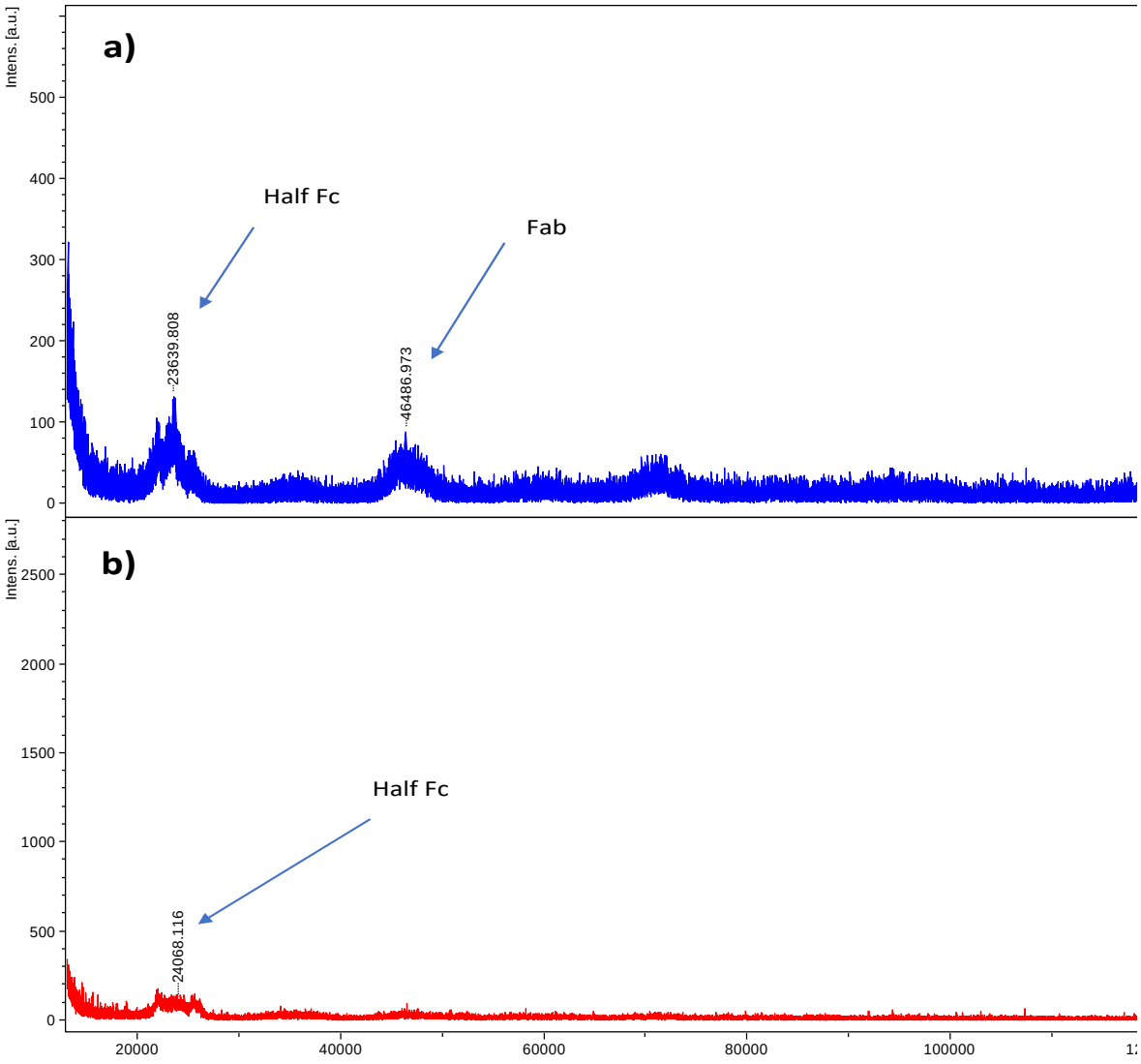

Figure 3.3.2. Linear positive MALDI-TOF mass spectra of HPLC-SEC a) fraction A3 containing ½ Fc and an intact Fab, and b) fraction A4 containing ½ Fc only.

3.3.3 Gel separation of Fab and Fc of pig IgG

A non-reducing gel experiment was performed on papain-digested pig IgG. If Fab and Fc were reduced prior to gel, the disulfide bonds which hold Fab and Fc chains together would be removed, and it would not be so easy to tell Fab and Fc apart as they would have both split into ~25 kDa fragments. The non-reducing experiment allowed for individuality of the Fab fragment

to be determined when still intact. After separation, the gel was stained and de-stained to reveal bands. The reducing experiment was performed on intact pig IgG as well as a control. Reduced heavy chain, reduced light chain, non-reduced Fab and Fc bands were then excised from the gel. Tryptic digestion was done on cut bands. Cutting of bands from the gel helped to single out heavy and light chains, which were used as a "standards" for differentiating between Fab and Fc according to their respective glycopeptides and peptides (see Table 3.2). In Figure 3.3.3a lanes 1-4 and 6- 9, corresponding to the fractions in descending order of HPLC-SEC containing both Fab and Fc, a single 50 kDa Fab band and two 25 kDa bands were obtained, as observed from Figure 3.3.2. In Figure 3.3.3a the Fab is located at the 50 kDa marker, and is positioned below the 50kda marker band, indicating that it is less than 50k Da, and in theory the Fab should have a lower molecular weight than intact Fc. The split Fc was located at the 25 kDa region. This was further revealed by the in-gel tryptic digestion procedure, which took place after the SDS gel experiment (next section).

The results from the gel support the suggested contents of HPLC fractions which were viewed by MALDI MS and indicated that the Fab was eluted first, then the Fc. Figure 3.3.3b shows a reduced gel of intact pig IgG without being digested prior to gel. This was done to show the difference between the digested form of IgG on gel and the non-digested form. Comparing Figure 3.3.3a and 3.3.3b it was proven that papain digested the pig sample, as the reduced sample shows a heavy chain above 50 kDa and a single 25 kDa light chain, while the non-reduced papain digested sample shows a Fab fragment below 50kDa and a split Fc at 25kDa. This was further proven after tryptic digestion of bands and analysis by MALDI- MS (next section).

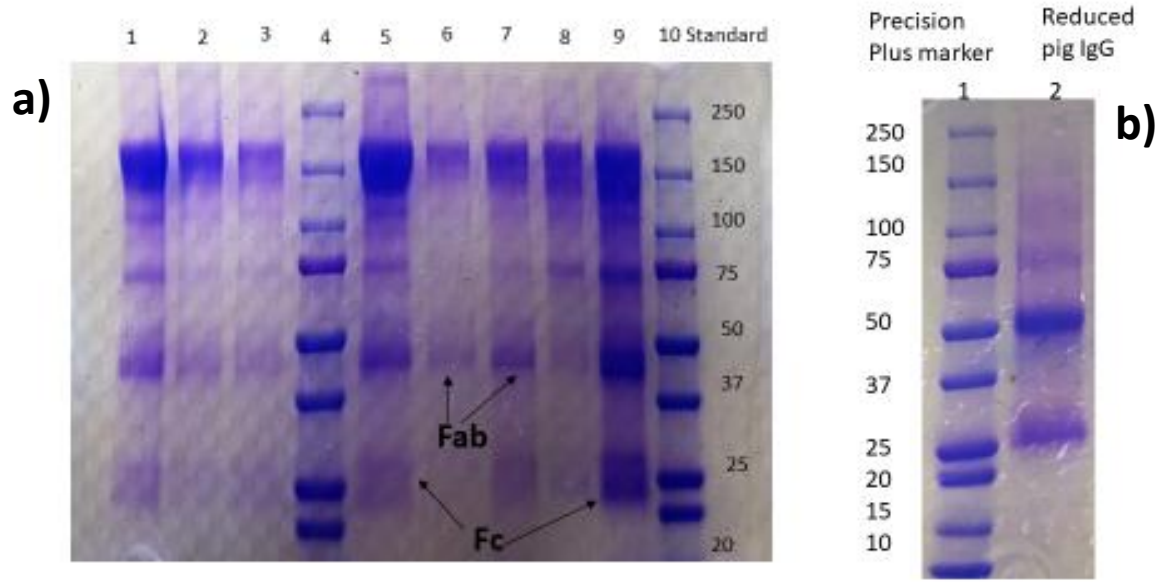

Figure 3.3.3. a) Pig IgG Fab and Fc HPLC-SEC fractions on acrylamide gel. Fab and Fc fragments are not reduced. b) Intact reduced pig IgG antibody.

3.3.4 In-gel tryptic digested Fab and Fc of pig IgG

In-gel tryptic digestion was carried out on Fab ~50kDa band, and the two Fc ~25 kDa bands this splitting of the Fc allowed for better separation as the Fab remained intact and would be at a different molecular weight on gel rather than having two bands at ~50 KDa, and produced peptide and glycopeptide fractions which were extracted by different buffers straight from the digest in Eppendorf tubes, and analyzed by MS. In Figure 3.3.4a only one peptide sequence is observed (EEQFNSTYR) with the three main glycoforms G0, G0F, and G1F (see Fig. 1.1.2) (*m/z* 2472, 2618, 2780 respectively) [45]. In Figure 3.3.4c with the other half Fc two peptide sequences are observed (EEQFNSTYR) with the glycoform G0F (see Fig. 1.1.2) (*m/z* 2618), and (EAQFNSTYR) with the main glycoform G0F (see Fig. 1.1.2) (*m/z* 2560) [45]. Glycans are obviously present in the Fc region of IgG, also other peptides corresponding to the Fc domain are

present (see Table 3.2). Figure 3.3.4b shows a positive ion, reflector mode MALDI-ToF-MS spectrum of the tryptic digest of the Fab fragment of IgG, showing absence of glycoforms found in Fc, but also peaks correlating to the light chain peptides and heavy chain Fab peptides which are absent from the Fc spectra. Known pig IgG code was retrieved from www.uniprot.com and known peptide sequences were then retrieved from www.expasy.com. Figure 3.3.4b shows no obvious glycopeptides from the Fab region (see Table 3.2). It displays some non-glycosylated peptides related to the Fab region (see Table 3.2). This indicates total separation of Fab and Fc regions from each other, thus no interference between the two portions was observed.

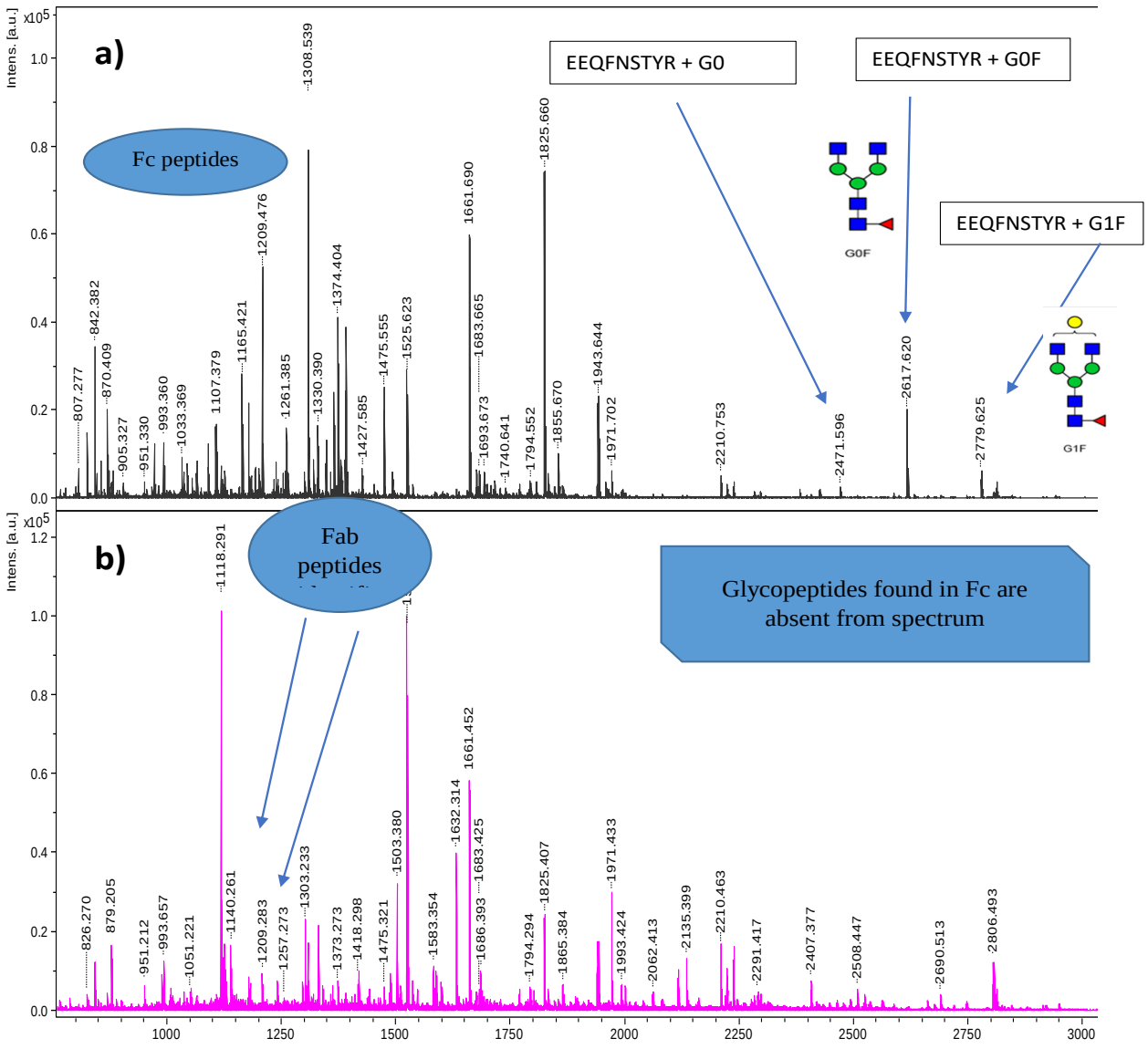

Figure 3.3.4. Reflector positive mode MALDI-MS spectra of: a) pig IgG Fc fragment, digested in-gel with trypsin and b) pig IgG intact Fab fragment, also digested in-gel with trypsin. See Table 3.2 for peak assignments.

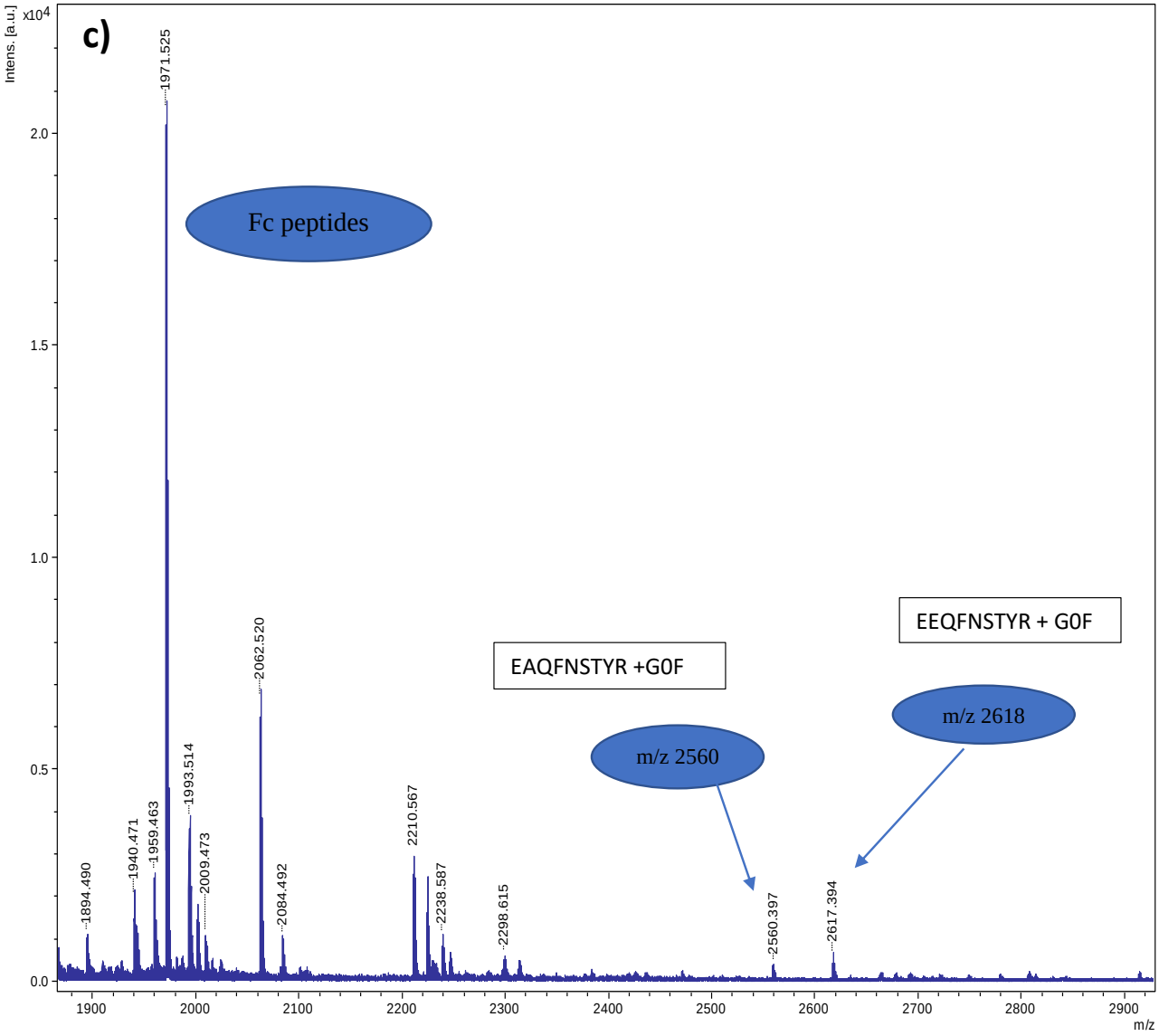

Figure 3.3.4.c Reflector positive MALDI-MS spectrum confirming the presence of 2560 and 2618 glycoforms in the lower 25 kDa gel band, indicating that the two bands at 25 kDa related to Fc. See Table 3.2 for peak assignments.

3.3.5 Analysis of pig Fc glycopeptides by LIFT™ technology PSD

A form of post source decay with LIFT™ technology was used in our experiments. This method decelerates and re-accelerates the selected precursor ions for MS/MS and enhances collisions between residual gas molecules and these precursor ions, which results in

fragmentation. MS/MS yields information about the primary structure and modifications of proteins. This was done for verification of amino acid sequence of glycopeptides found in the pig IgG Fc region. The PSD analysis conducted after tryptic digestion in this experiment produced MS/MS data on the amino acid sequences of the *m/z* 2560 and 2618 ions, which were determined to be EAQFNSTYR and EEQFNSTYR respectively. These are both from the Fc of pig IgG, and no noticeable Fab *N*-glycoforms were observed. MS/MS data are featured in Figure 3.3.5.

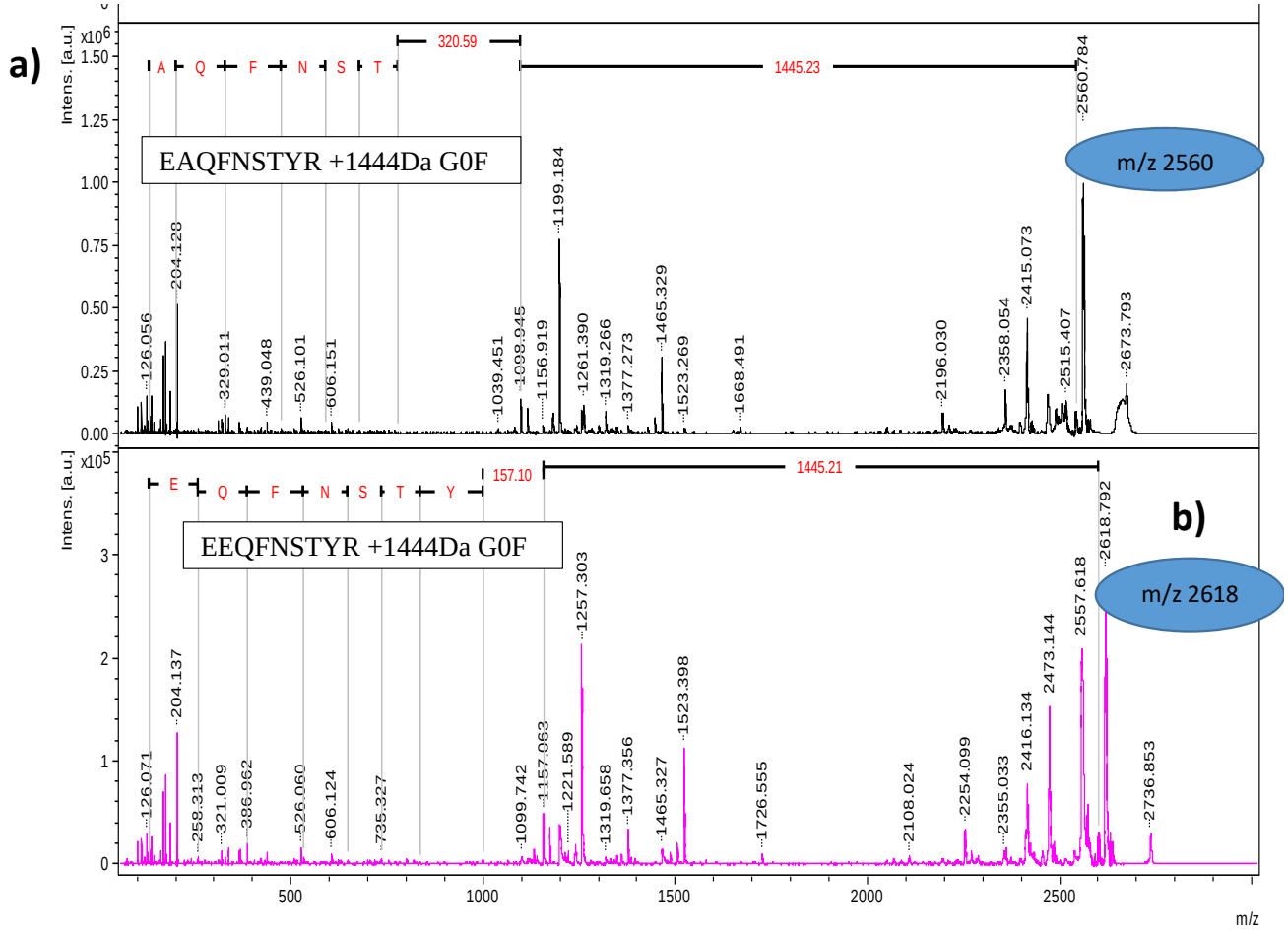

Figure 3.3.5. MALDI-MS/MS spectra of: a) the pig IgG glycoform with *m/z* 2560, with amino acid sequence EAQFNSTYR, and b) the pig IgG glycoform with *m/z* 2618, with amino acid sequence EEQFNSTYR. Both glycopeptides are G0F glycoforms.

3.3.6 PNGase F removal of *N*-glycans from pig Fc glycopeptide fraction

PNGase F is the most effective enzymatic method for removing almost all *N*-linked oligosaccharides from glycoproteins [7]. PNGase F is an amidase, which cleaves between the innermost GlcNAc and asparagine (Asn) residues of high mannose, hybrid, and complex oligosaccharides [7]. This changes the asparagine (Asn) into an aspartate (Asp) residue, after *N*-glycan mass is lost, by PNGase removal. To determine if the amino acid sequence is correct and is *N*-linked, the bond between Asn and GlcNAc must be able to be cleaved and this indicates the presence of of *N*-glycans. PNGase F analysis was done on fractions obtained by C18 peptide cleanup of tryptic digests of Fab and Fc fragments of pig IgG. In Figure 3.3.6a, the Pig IgG Fc fragment tryptic digest contains the deglycosylated peptides at *m/z* 1117 Da (EAQFNSTYR) and *m/z* 1175 Da (EEQFNSTYR), while in Figure 3.3.6b the Fab fragments did not have these or other deglycosylated peptides. The difference in peaks can be clearly seen, indicating that the Fc and Fab fragments were indeed distinct from each other, and were analyzed separately, which proves/satisfies the ultimate goal of this section (see Figure 3.3.6).

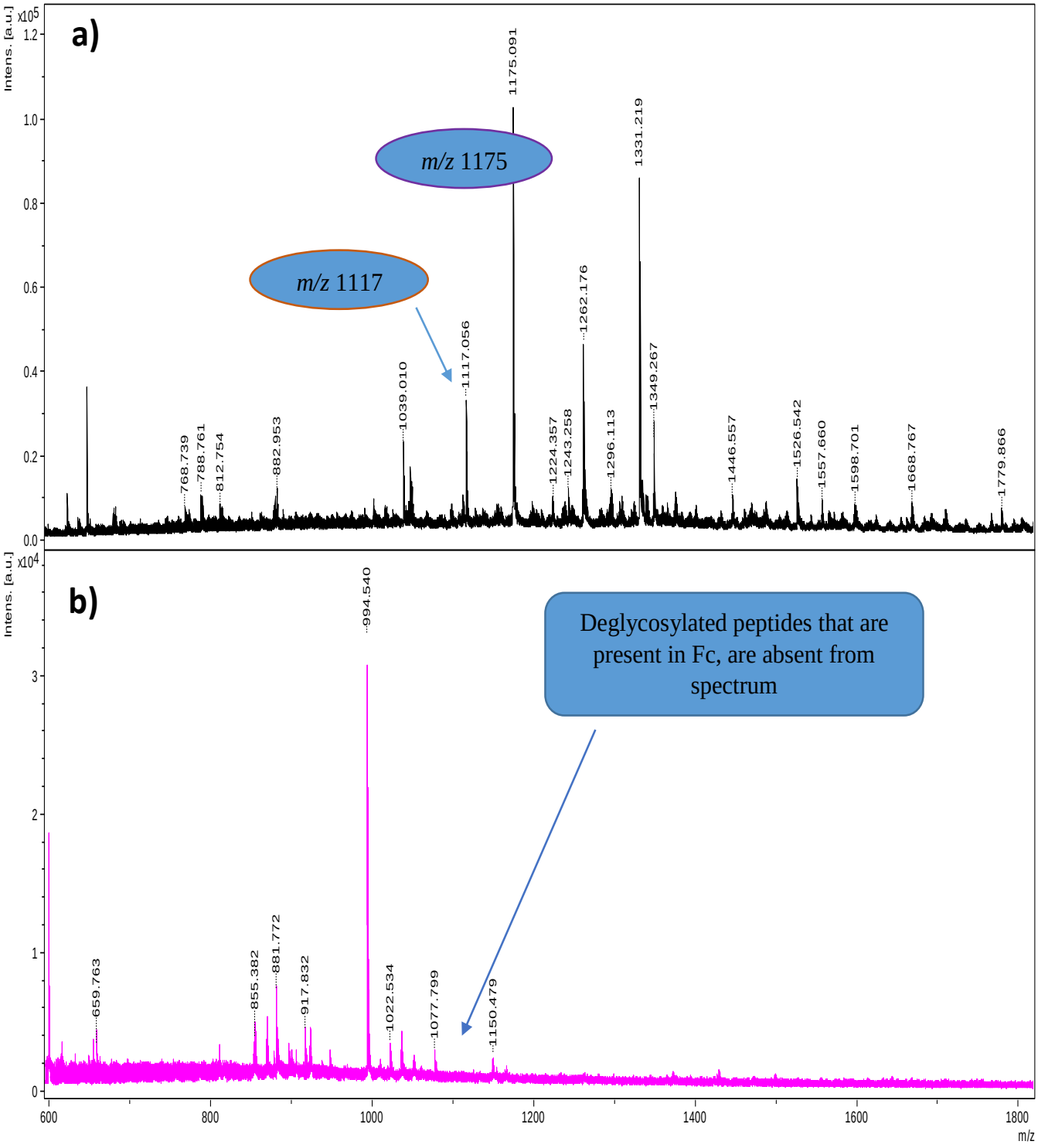

Figure 3.3.6. Shows MALDI-MS spectra of: a) reflector positive mode MALDI-MS spectrum of Fc deglycosylated peptides after removal of glycans to give peptide *m/z* 1117 (EAQFNSTYR) and 1175 (EEQFNSTYR), and b) Fab deglycosylation spectrum.

3.4 Fabricator™ fragmented analysis of antibody *N*-glycosylation for human IgG by MALDI-MS

3.4.1 HPLC separation of Fab and Fc after Fabricator™ digestion of human IgG

Pure human IgG was digested with Fabricator™ for 24 h. The digestion mixture was analyzed using a HPLC SEC column. The isocratic mobile phase used was 0.1% TFA, 0.1% formic acid in 20% ACN. A small quantity of antibody remained undigested after incubation (see Figure 3.4.1). Fabricator™ is a partially modified and His-tagged enzyme which cleaves IgG in the hinge region, leaving an intact F(ab´)$_2$ (100kda) and two residual Fc fragments (25 kDa each) [10]. The F(ab´)$_2$ fragment is divalent, which means that a single Fab fragment obtained from papain digestion similarly can be analyzed to get the same detailed information on the F(ab´)$_2$ fragment.

In this experiment, using 200 μg of human IgG, the F(ab´)$_2$ fragment remained intact, and residual Fc fragments were noticed, from SDS- PAGE gel. The HPLC chromatogram in Figure 3.4.1 shows a small Fc peak eluting after a broad F(ab´)$_2$ peak. Collection of F(ab´)$_2$ and Fc from HPLC was achieved. These fractions collected were run next on SDS gel. The identification of these peaks as such was verified by gel electrophoresis and further tryptic digestion (see Figure 3.4.1).

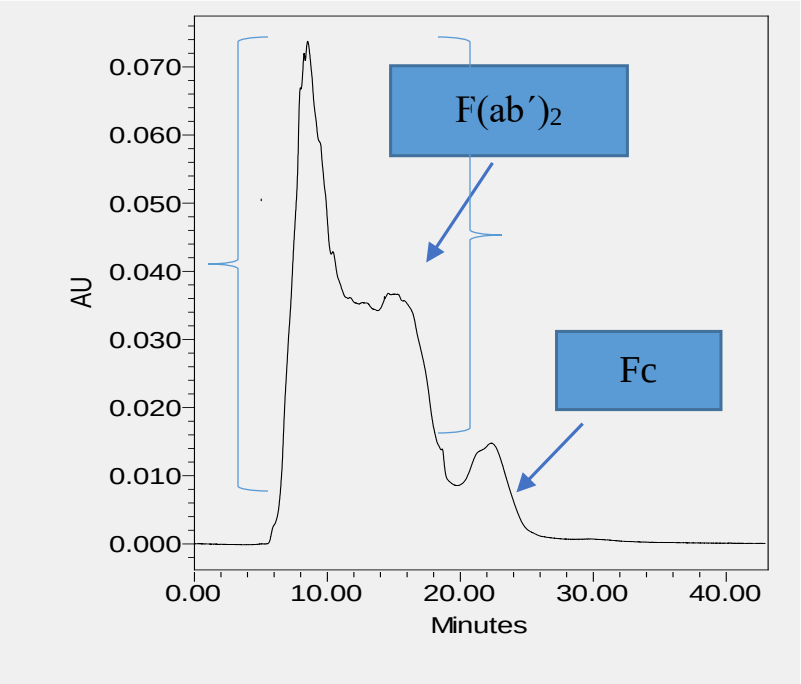

Figure 3.4.1. HPLC chromatogram of fragments after Fabricator™ digests of human IgG.

3.4.2 Gel separation of 'Fab'2 and Fc of human IgG

In Figure 3.4.2, the F(ab´)$_2$ and Fc fragments were not reduced prior to gel separation. The non-reducing experiment allowed for individuality of the fragments to be determined when still intact. After separation, the gel was stained and de-stained to reveal bands. Non-reduced F(ab´)$_2$ and Fc bands were then excised from the gel. Tryptic digestion was done on cut bands. The F(ab´)$_2$ and Fc bands were also distinguished and separated from each other. Lanes 2-5 corresponds to the HPLC-SEC fractions in descending order, containing both F(ab´)$_2$ and a split Fc, two distinguishable bands were obtained, at 100 kDa and 25 kDa respectively on gel. This was further investigated by the in-gel tryptic digestion procedure, which took place after the SDS gel experiment (next section). The results from the gel support the suggested contents of HPLC

fractions which were viewed by MALDI-MS and indicated that the F(ab´)$_2$ was eluted first, then the Fc.

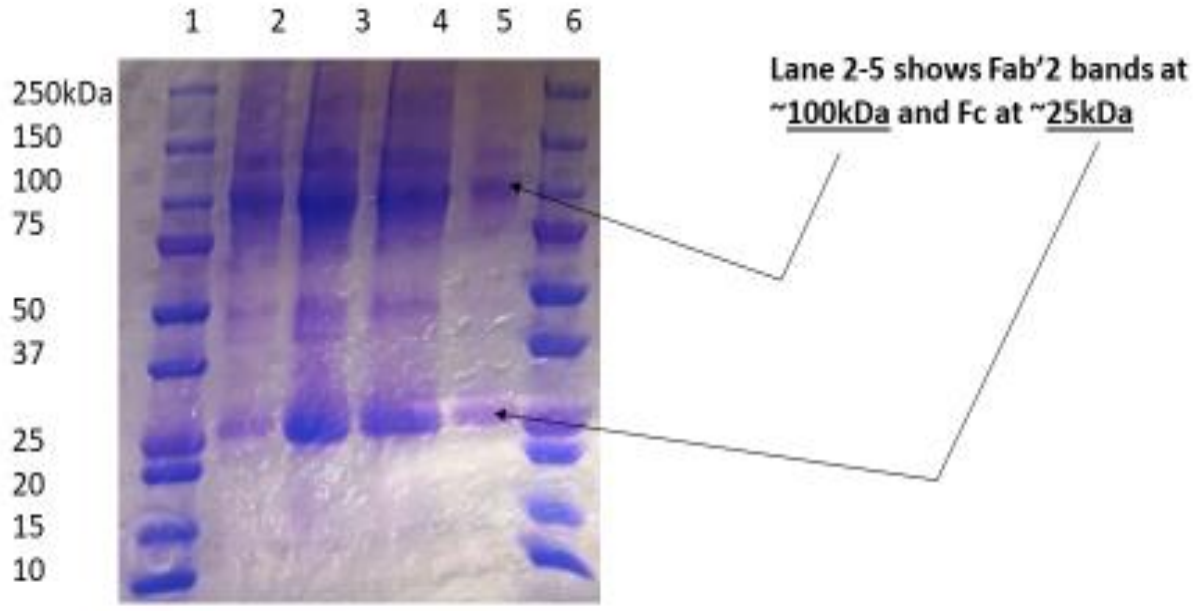

Figure 3.4.2. Human IgG F(ab´)$_2$ (100 kDa) and Fc (25 kDa) HPLC-SEC fractions on SDS acrylamide gel.

3.4.3 In-gel tryptic digested F(ab´)$_2$ and Fc of human IgG

In-gel tryptic digestion was carried out on F(ab´)$_2$, and Fc bands, and produced peptides and glycopeptides which were extracted by different buffers into these fractions, straight from the digest in Eppendorf tubes (there was no clean up), and MS analysed. Figure 3.4.3a shows a positive ion, reflector mode MALDI-ToF-MS spectrum of the tryptic digest of the Fc fragment of human IgG. In Figure 3.4.3a, two glycoform peptide sequences were observed, EEQYNSTYR with the three main glycoforms G0F, G1F, and G2F (see Fig. 1.1.2) (*m/z* 2634, 2796, and 2958 respectively), and (EAQYNSTYR) with the three main glycoforms G0F, G1F,

and G2F (see Fig. 1.1.2) (*m/z* 2602, 2764, and 2926 respectively). The figure illustrates a 162 Da hexose sugar differences between related glycoforms. Glycans are obviously present in the Fc region of IgG, also other peptides corresponding to the Fc domain are present as seen in Figure 3.4.3a (see Table 3.1). Figure 3.4.3b shows no obvious glycopeptides from the Fc region. It displays some non-glycosylated peptides related to the Fab region (see Table 3.1). This indicates total separation of Fab and Fc regions from each other, thus no interference in analysis as they are separate entities. Performance advantage of Fabricator over papain, is that Fabricator digestion yields two intact Fab fragment collectively called $F(ab')_2$, and thus allows for analysis of both homologous Fab fragments; whilst papain digestion yields only one intact Fab fragment. If no cysteine is present during papain digestion, $F(ab')_2$ fragments can be generated. These fragments are often inconsistent, and reproducibility can be a problem [8]. If cysteine is not removed completely over digestion as well of the disulfide bonds can be a problem. Thus Fabricator^{TM} is preferred when digesting human IgG samples.

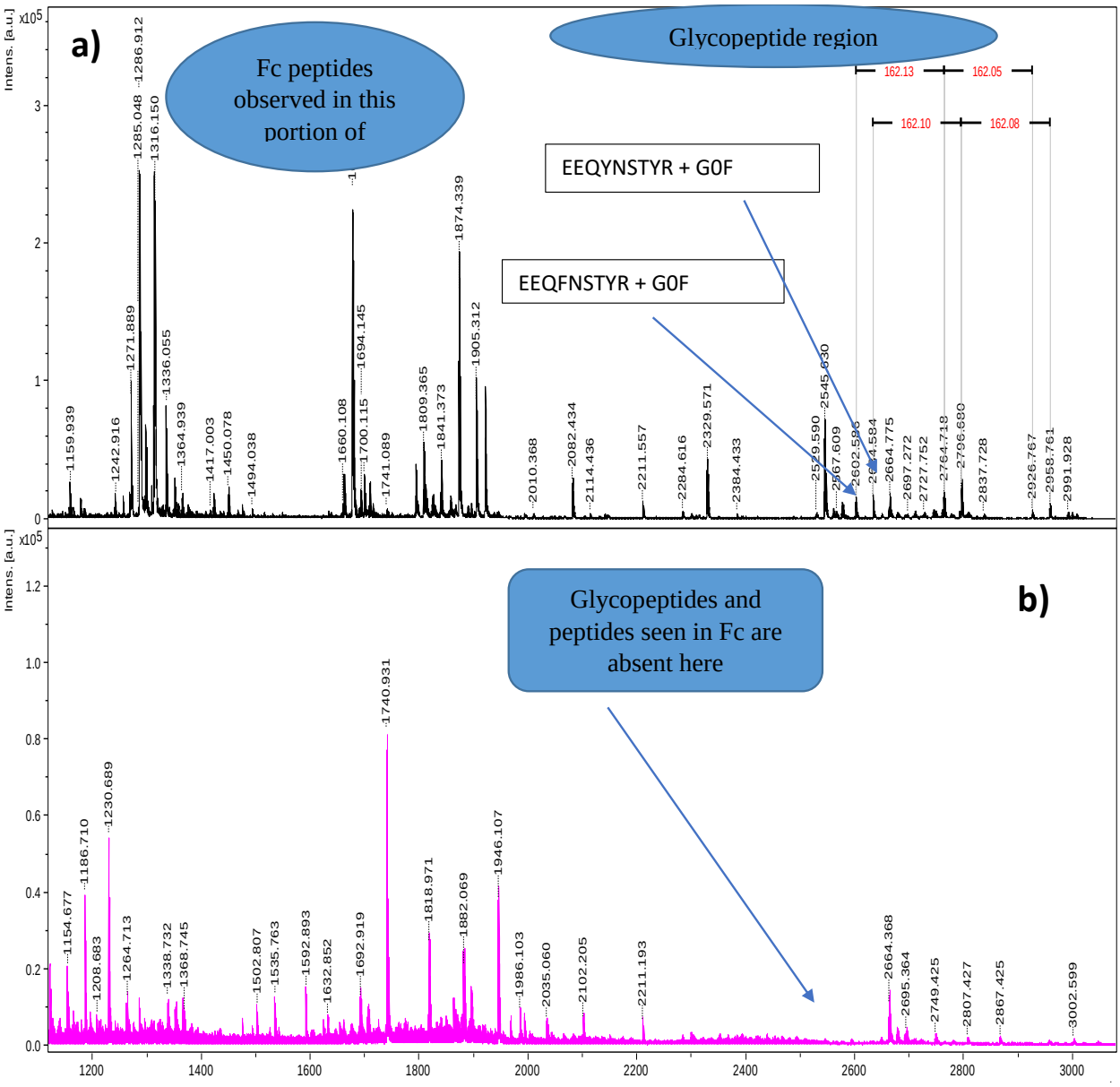

Figure 3.4.3 Reflector positive MALDI-MS of: a) human Fc peptides and glycopeptides and b) F(ab')2 peptides.

3.5 Fabulous™ fragmentation analysis of pig IgG by MALDI-MS

3.5.1 HPLC separation of Fab and Fc after Fabulous™ digestion of pig IgG

Fabulous™ (SpeB) is a recombinantly produced cysteine protease that under reduced conditions digests in the hinge region of antibodies from many species and subclasses, including human, mouse, rat and goat, yielding Fab and Fc fragments [9]. When IgG molecules are incubated with Fabulous™ in the presence of a reducing agent, one or more peptide bonds in the hinge region are split, producing two Fab and one Fc. As reducing agent is present during digestion reaction, it is likely that interchain disulfide bonds will be reduced. Pig IgG was digested with Fabulous enzyme for 24 h. The digestion mixture was analyzed using a HPLC SEC column. The isocratic mobile phase used was 0.1% TFA, 0.1% formic acid in 20% ACN. It is important to note that there was not a sufficient quantity of enzyme to test it on human IgG prior to pig IgG.

This enzyme has been used on several species before except pig, which was examined in this project and this experiment showed it digested the pig IgG in a similar action as the papain enzyme. Both enzymes are specific for the hinge region only, and reducing conditions usually affect interchain disulfide bonds of the Fc thus splitting it into two halves. In this experiment with Fabulous™, using 200 μg of pig IgG, the Fc was split into two halves and the Fab remained intact. The HPLC chromatogram in Figure 3.5.1 showed separation, of Fab and Fc. The first major peak in Figure 3.5.1 was indicative of the Fab fragment and the second peak eluting after identified as the Fc, as verified later by gel electrophoresis and further tryptic digestion. A small amount of the intact IgG remained as seen in Figure 3.5.1, which indicates that digestion was

successful and was not incomplete. The gel results and in-gel tryptic digestion results of Fab and Fc will be discussed in next section.

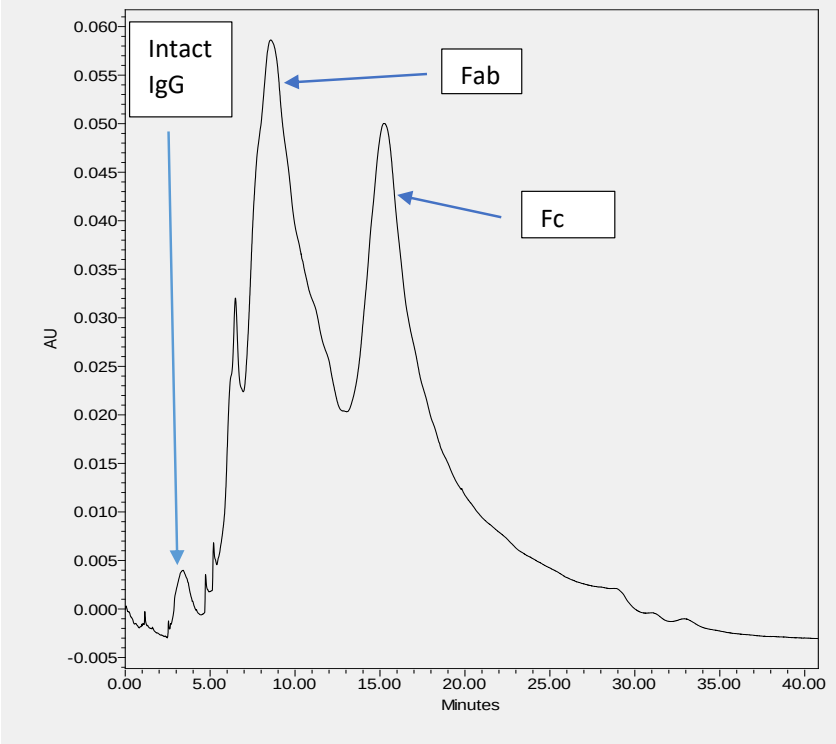

Figure 3.5.1. HPLC-SEC chromatogram for pig IgG after Fabulous™ digestion.

3.5.2 Gel separation of Fab and Fc of pig IgG

The HPLC chromatogram results, correlated with the non-reducing SDS-PAGE gel electrophoresis results in Figure 3.5.2, which shows Fab and Fc separated as an intact less than 50 kDa Fab band and two split Fc bands at ~25 kDa, which indicates that the Fab was heavier and eluted first, due to the split of the Fc fragment. The non-reduced Fab and Fc bands were then excised from the gel. Tryptic digestion was done on the intact Fab and two split Fc cut bands. In lanes 1-3, corresponding to the HPLC-SEC fractions in descending order containing both Fab

and Fc, two distinguishable bands were obtained, the Fab below 50 kDa and Fc which is indicated by a band above and below 25 kDa. This was further revealed by the in-gel tryptic digestion procedure, which took place after the SDS gel experiment (next section).

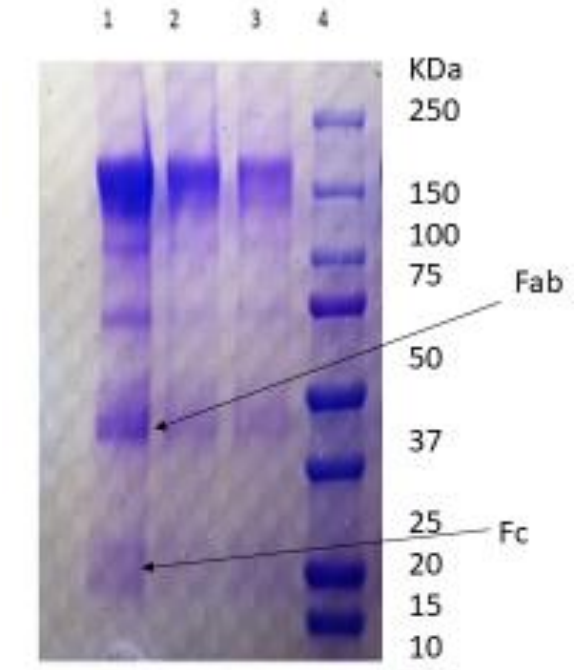

Figure 3.5.2. Pig IgG Fab and Fc HPLC-SEC fractions on acrylamide gel. The Fab remained intact and Fc fragments were split in halves.

3.5.3 MALDI-MS after in-gel tryptic digestion of Fab and Fc of pig IgG

In-gel tryptic digestion was carried out on Fab and Fc and produced peptide and glycopeptide fractions which were collected straight from the digest in Eppendorf tubes by different extraction buffers, and MS analyzed. In Figure 3.5.3a two peptide sequences were observed 1) EEQYNSTYR with the main glycoforms, G0F, G1F, and G2F (*m/z* 2618, 2780, and 2942 respectively), and 2) EAQYNSTYR with the main glycoforms G0, G0F, G2F (*m/z* 2414, 2560, and 2884 respectively). This latter figure indicates the presence of low intensity peaks of

the glycoform of *m/z* 2560. Figure 3.5.3b shows a positive ion, reflector mode MALDI-ToF-MS spectrum of the tryptic digest of the Fab fragment of pig IgG. Figure 3.5.3b shows no obvious glycopeptides from the Fab region (see Table 3.2) and Fc glycopeptides of Figure 3.5.3a are not present in spectrum. It displays some non-glycosylated peptides related to the Fab region (see Table 3.2).

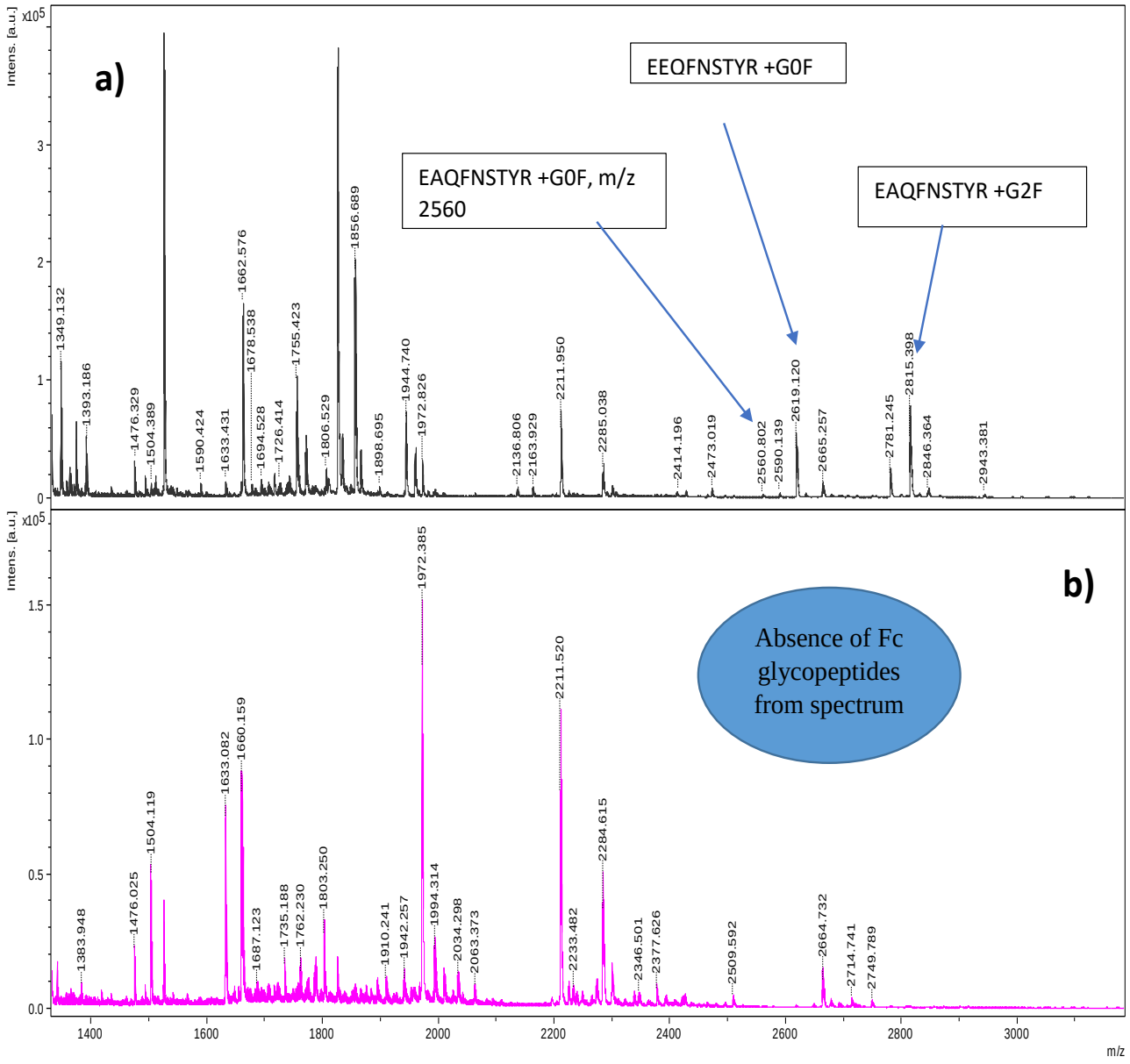

Figure 3.5.3. Reflector positive mode MALDI-MS spectra of tryptic digests products of: a) pig IgG Fc (two bands at 25 kDa) and b) pig IgG intact Fab.

3.6 Analysis of Fabricator™ digested Human and Fabulous™ digested pig IgG by UPLC-MS

3.6.1. Fabricator fragmented analysis of human IgG by reverse phase UPLC-MS: MS results after in-gel tryptic digestion of Fab and Fc of human IgG

In-gel tryptic digestion was carried out on F(ab′)$_2$, and Fc fragments of human IgG. After buffer extraction of IgG fragments, peptide and glycopeptide fractions were collected and MS analysed by UPLC ESI-MS. UPLC separated the peptides in Fab, Fc and F(ab′)$_2$ tryptic digested samples. A list of the human Fc and Fab peptides detected by UPLC-MS can be observed in Table 3.1.

In Figure 3.6.1, LC-MS total ion chromatograms are shown for the digestion products of human F(ab′)$_2$ and Fc and showed differences in retention time for peaks observed in the glycopeptide or peptide regions. At early stages of the runs, the glycopeptide and peptides eluted and the peak retention times were different between human F(ab′)$_2$ and Fc.

Figure 3.6.2a shows a full range positive mode ESI spectrum of the human Fc tryptic peptides and glycopeptides eluting at retention time 8.3 min. Figure 3.6.2b and c show closeup views of this spectrum with focus on b) *m/z* 1302 doubly charged ions and c) m/z 868 triply charged ions of a glycopeptide of mass 2601 Da. This corresponds to the G0F form of EEQFNSTFR from human IgG2.

Figure 3.6.3a shows a portion of the ESI spectrum obtained at retention time 4.4 min, focusing on human IgG1 Fc glycopeptide of mass 2633 with charge state of +3 (*m/z* 878). This

corresponds to the G0F glycoform of peptide EEQYNSTYR. Figure 3.6.3b focuses on the G2FS glycoform of EEQFNSTFR (mass 3054) with charge state +3 (*m/z* 1019). The form of sialic acid observed here in human IgG is *N*-acetyl neuraminic acid (NeuAc). The disialylated form G2FS2 of the same peptide (mass 3248, charge +3, *m/z* 1083) is observed in Figure 3.6.3c. This disialylated form of the peptide was not observed by reflector positive mode MALDI-MS. These two last figures are only examples of human IgG tryptic digest products observed by UPLC-MS. The observation that sialylated species elute at 4.4 min earlier than neutral glycoforms is logical, as they are more polar than the neutral species. Fab peptides were identified in total ion chromatogram, under the peak areas shown in figure 3.6.1. In ESI-MS spectra done after UPLC run the m/z ratio of each of the peptides was observed under the peak areas identified. Fab peptides mass and retention time can be seen in table S5. All peptides and glycopeptides observed in these experiments are listed in Table 3.1.

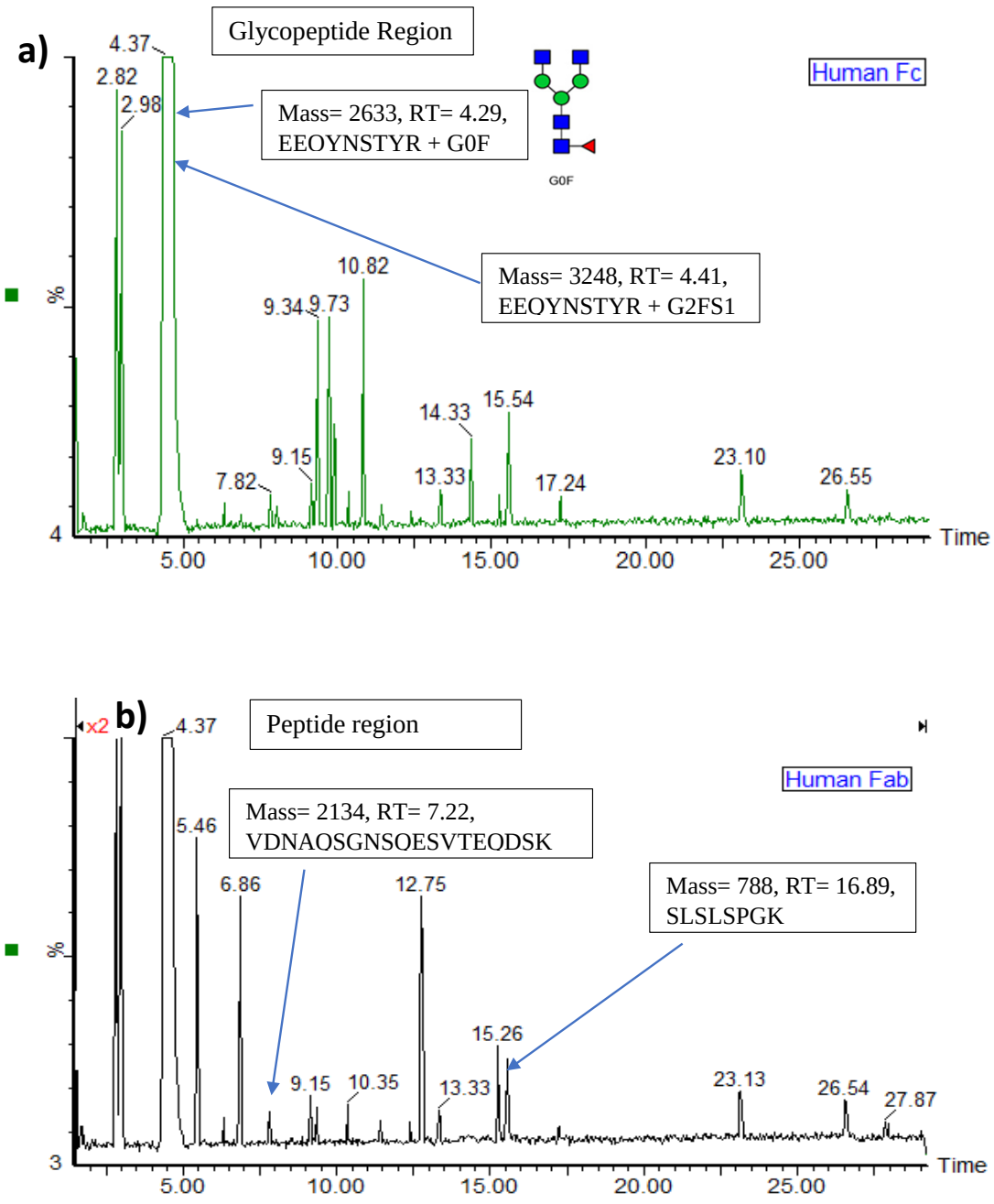

Figure 3.6.1 UPLC total ion chromatogram of: a) human Fc tryptic digestion products, and b) human Fab tryptic digestion products.

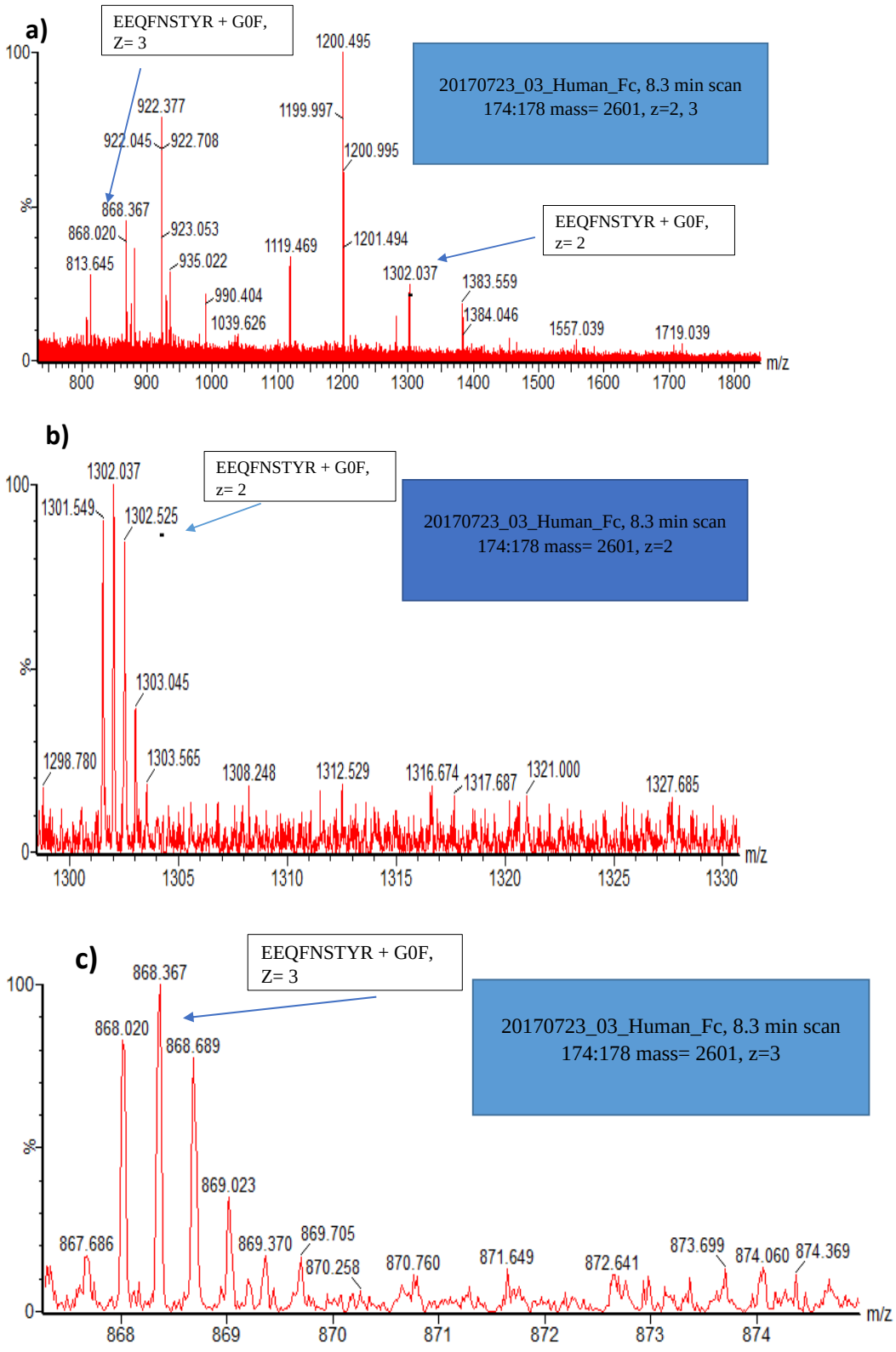

Figure 3.6.2 a) Full range positive ESI spectrum corresponding to the peak at 8.3 min observed in Figure 3.6.1a for the tryptic digest of human IgG Fc; b) exploded view of the *m/z* 1302 doubly

charged ions and **c)** of the *m/z* 868 triply charged ions. These characterize a glycopeptide of mass 2601, the G0F glycoform of EEQFNSTFR (human IgG2).

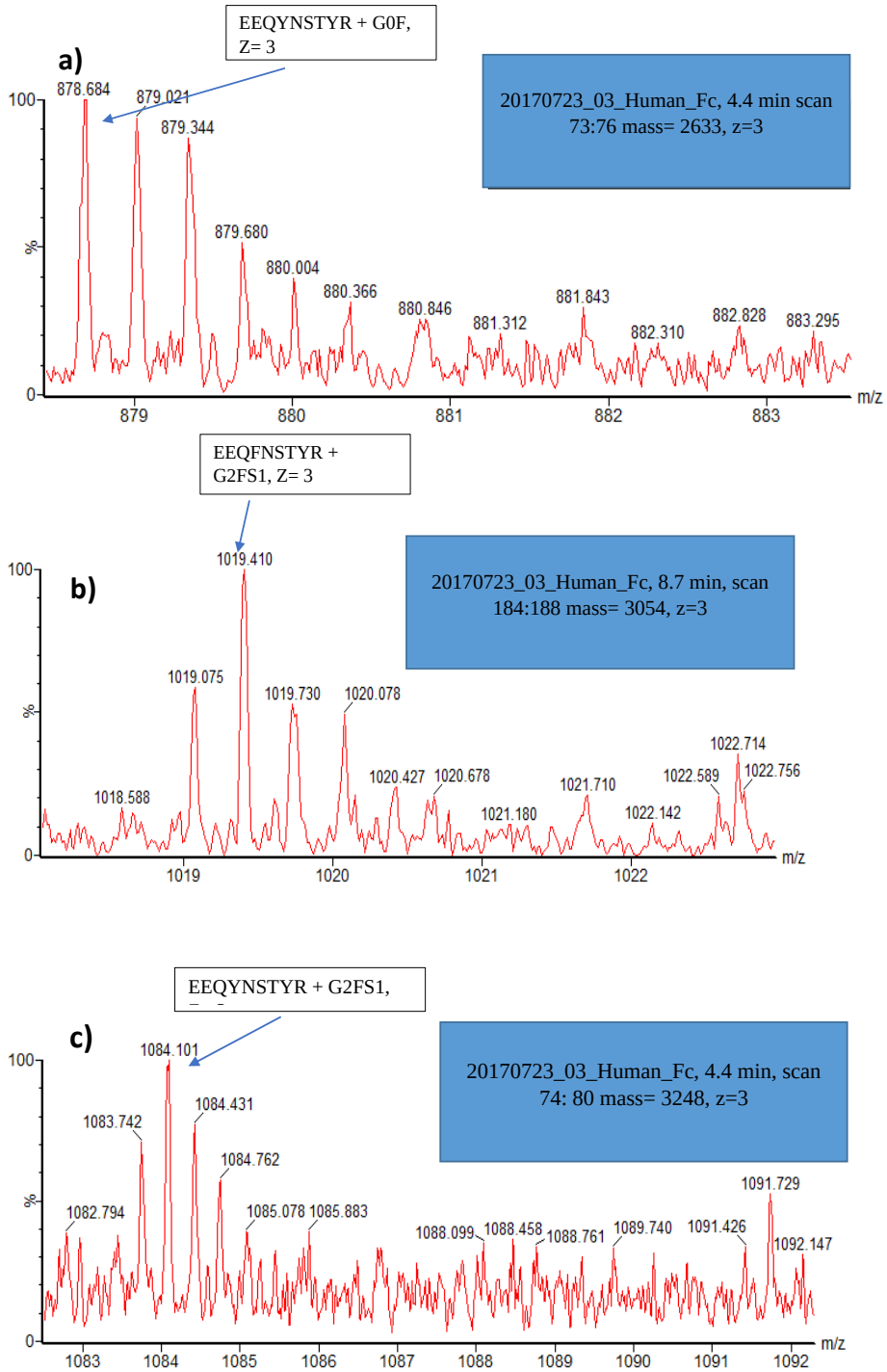

Figure 3.6.3 In a), *m/z* 878 ions correspond to the triply charged ions of a glycopeptide of mass 2633, i.e. the G0F glycoform of human IgG1 EEQYNSTYR. In b), the monosialylated G2FS

glycoform of EEQFNSTFR (human IgG2, mass 3054) produces triply charged ions at *m/z* 1019. Finally in c), the same IgG2 peptide's glycoform G2FS2 (mass 3248) has triply charged ions at *m/z* 1083. All three spectral portions were obtained at retention time 4.4 min.

The Progenesis™ software used for UPLC-MS data interpretation comprises a feature called "distribution charts" which allows to graphically represent the peaks of analytes of known masses through a whole chromatographic experiment. The axes in these charts have arbitrary units, as the charts compress the ion mass chromatograms to fit in a small window. This is qualitative, but useful for the comparison of LC-MS data obtained from two different samples, i.e. in our case the tryptic digestion products of human IgG Fc and Fab fragments. For instance, Figure 3.6.4 shows distribution charts indicating the absence or presence of a) neutral glycoforms of peptide EEQFNSTFR and b) acidic forms of EEQFNSTFR throughout all the Fab and Fc species from human and pig IgGs. According to the pictograms, these glycopeptides exist only in human. The neutral forms in a) are G0F, G1F and G2F (second panel from the left). The acidic glycoforms in b), G2FS and G2FS2, where S = NeuAc, are featured also in the second left panel. These results are not surprising, given that the sequence EEQFNSTFR does not exist in any of the 6 subclasses of pig IgG but does exist in human IgG (IgG2).

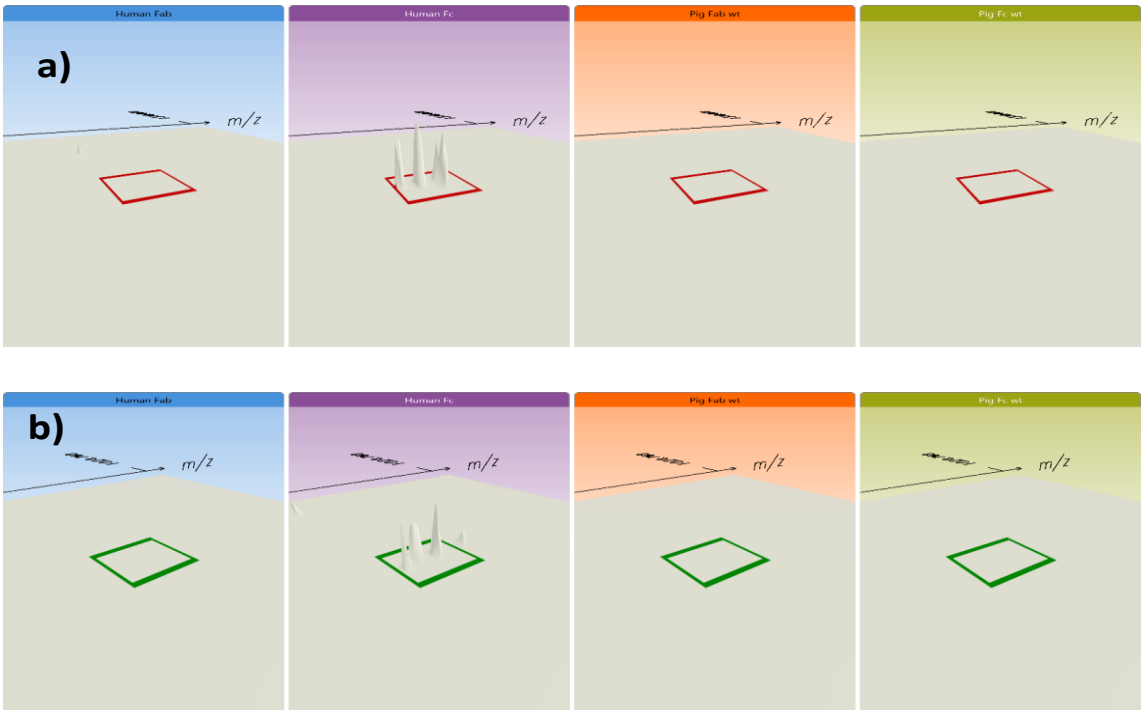

Figure 3.6.4 UPLC-MS distribution charts indicating the absence or presence of glycoforms of EEQFNSTFR: a) neutral glycoforms G0F, G1F and G2F; and b) acidic glycoforms G2FS and G2FS2 (where S = NeuAc) in the Fab and Fc species from human and pig IgGs.

3.6.5. Fabulous fragmented analysis of pig IgG by UPLC-MS: MS results after in-gel tryptic digested Fab and Fc of pig IgG

In-gel tryptic digestion was carried out on Fab, Fc, fragments of pig IgG. After buffer extraction of IgG fragments, peptide and glycopeptide fractions were collected and MS analyzed by UPLC ESI-MS. In Figure 3.6.5, the LC-MS total ion chromatograms of the pig Fab and Fc showed differences in the retention times of peaks observed in glycopeptide or peptide regions. At early stages of the experiments, the glycopeptides and peptides were eluted and the peaks' retention times were different for pig IgG Fab and Fc glycopeptides. Pig Fab and Fc peptides observed by this method are listed in Table 3.2.

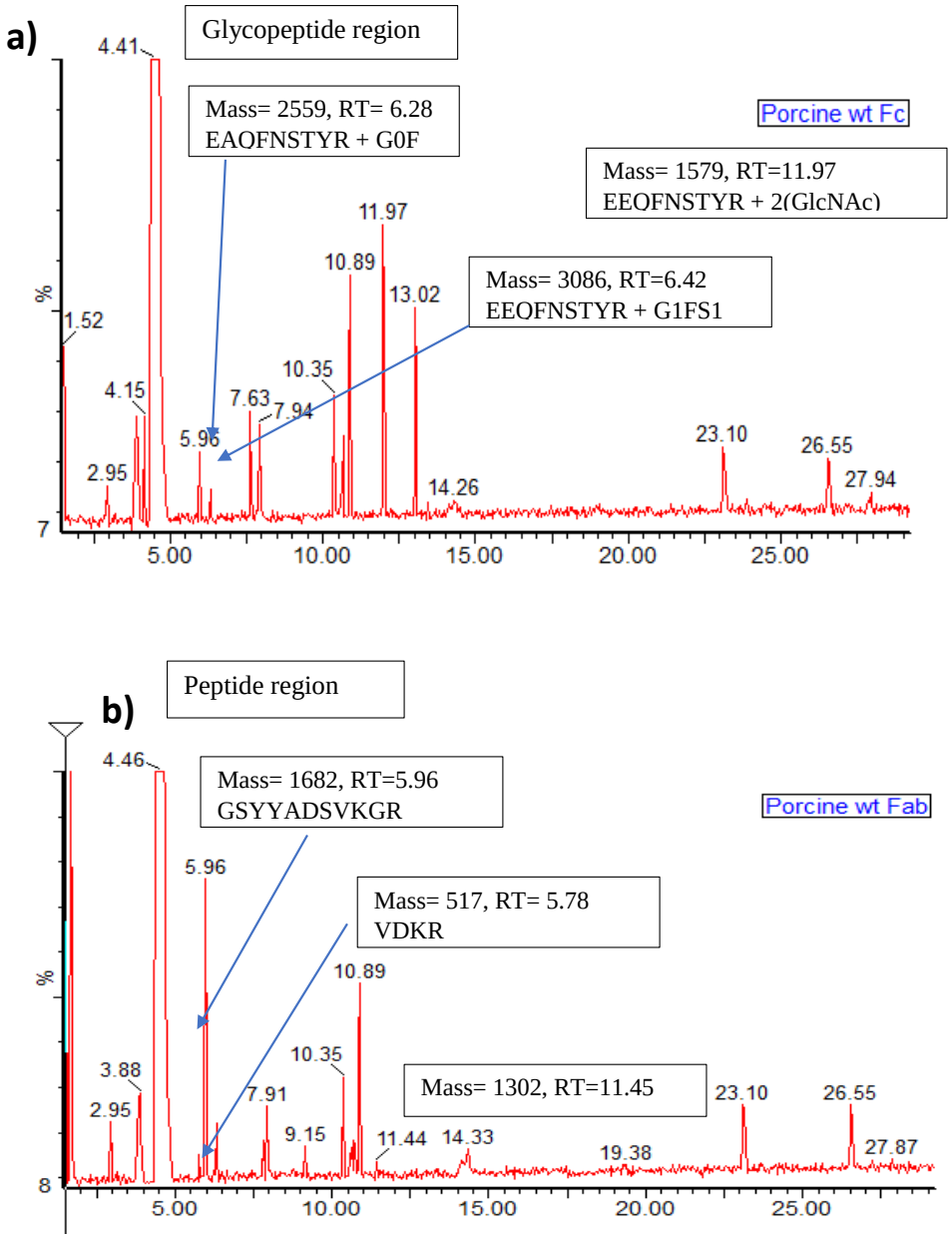

Figure 3.6.5 Total ion chromatograms obtained by UPLC-MS measurements on the separation of a) pig IgG Fab tryptic digests and b) pig IgG Fc tryptic digests.

As examples of results, Figure 3.6.6a shows the positive ESI mass spectrum of the G0F glycoform of pig Fc's EEQFNSTYR of mass 2617 Da, with charge state of +2 (*m/z* 1310) and retention time of 6.17 minutes. Figure 3.6.6b shows the MS spectrum of the monosialylated

G2FS glycoform of the same peptide (mass 3086 Da), where S = *N*-glycolylneuraminic acid (NeuGc). The charge state observed is +3 (*m/z* 1030). This monosialylated glycoform of the peptide was not observed by reflector positive mode MALDI-MS in this study.

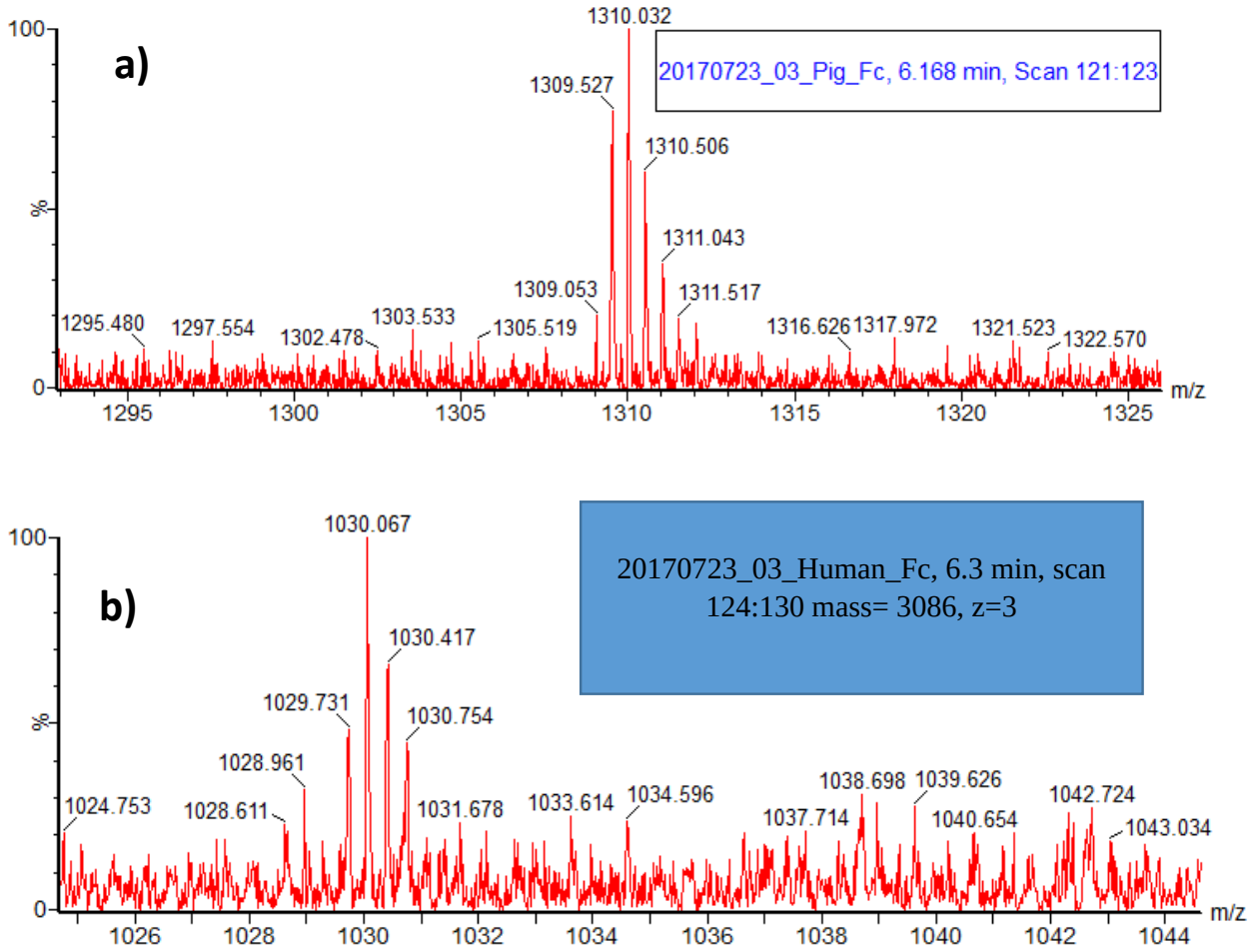

Figure 3.6.6 a) Positive mode ESI-MS spectrum of the pig Fc G0F glycoform of EEQFNSTYR, mass 2617. Charge state +2, *m/z* 1310 and retention time of 6.17 minutes; b) ESI-MS spectrum of the monosialylated G2FS form of pig Fc peptide EEQFNSTYR, mass 3086, where S = NeuGc. The charge state observed is +3, *m/z* 1030.

Again, distribution charts were used to highlight the presence or absence of neutral and sialylated glycoforms of EEQFNSTYR, a peptide prevalent in pig IgG but of very low concentration in human IgG [45]. In Figure 3.6.7 a-b, the last pictogram on the right represents the elution of neutral (a) and sialylated (b) glycoforms of this peptide, only detected in the pig IgG Fc fragment.

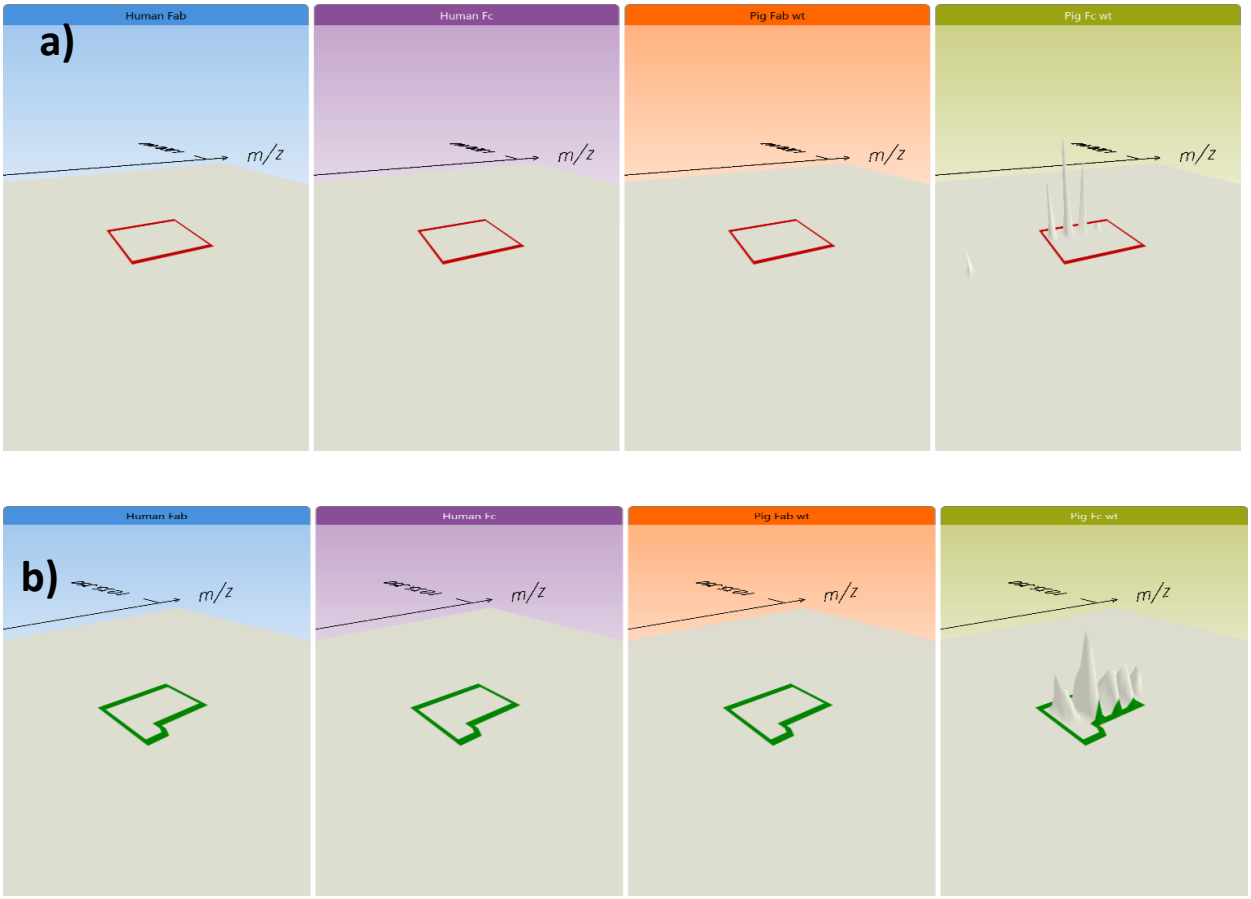

Figure 3.6.7 Distribution charts indicating absence or presence of a) neutral glycoforms G0F, G1F, G2F of EEQFNSTYR and b) sialylated G2FS (S = NeuGc) glycoforms of EEQFNSTYR. Samples represented are form left to right: human Fab, human Fc, pig Fab, Pig Fc. All the Fab peptides detected and other Fc glycopeptides of pig IgG are listed Table 3.2.

In general MALDI typically produces far fewer multiply charged ions than UPLC-ESI-MS. By producing mostly singly charged protonated ions, interpretation of data is simpler than with ESI-MS, where masses need to be calculated from the *m/z* values of multiply charged ions. In this study UPLC-ESI-MS in positive mode gave more detailed data on sialylated glycopeptides than MALDI-TOF-MS for both human and pig Abs. No MS/MS data were acquired, so sequences could not be determined, therefore peptide peak masses were retrieved, and compared to known peptide masses and sequences [45].

Analyses of IgG fragments by both MALDI-ToF MS and UPLC-ESI-MS reveal separation of Fab and Fc, in that the UPLC chromatograms show no similar retention times of peptides and glycopeptides in the Fab *vs.* Fc samples. Also clear distinctions were further seen at the MS level, in that the ESI-MS spectra for Fab and Fc, revealed differences in the tryptic peptides and the glycopeptides obtained from these fragments. Indeed, peptide masses obtained were entirely different from Fc spectra to Fab spectra of both human and pig IgG. There was a high level of correlation between MALDI-MS and UPLC-MS data.

The UPLC-MS mass spectra obtained for human and porcine Fab tryptic peptides are not shown in this thesis, as there was no evidence of *N*-glycosylation in these species. Instead, a peptide mass list was created (see table 3.2) for known and unknown sequences based on database findings and peptide masses obtained from both ESI-MS and MALDI-MS data acquisitions. For example, human Fab peptides with mass 2134 and 788 with retention times 7.22 and 6.89 min, respectively, were observed. The pig Fab peptides observed were at mass 517, and 1302 with retention times 5.78, and 11.45 respectively. A few more of these peptides and sequences are listed in Table 3.2.

Chapter 4

Conclusions and Future work

4.1 Comparison of data obtained from whole antibody digestion and fragmented antibody digests of human and pig IgG

The general analysis seen from whole tryptic digestion of the pig IgG show peptides and glycopeptides of whole IgG, and no distinguishing information on where glycosylation is occurring in the IgG or on the origin of peptides within the antibody. Fragmentation of IgG into its Fc, Fab, or $F(ab')_2$ portions can be achieved using papain, Fabulous™ and Fabricator™ enzymes. Fragments can be separated by use of HPLC-SEC and SDS-PAGE and analyzed by MALDI-TOF-MS or UPLC ESI-MS. The second workflow is preferred to the more conventional method of whole antibody digestion with trypsin, as it is more effective and accurate in determining exact glycosylation sites, and whether the peptides/glycopeptides are from the Fab or Fc region. Antibody fragmentation results in elimination of interferences between the Fab and Fc regions, allowing for more detailed characterization.

The second workflow procedure, with all above mentioned enzymes used in the thesis work along with HPLC-SEC, SDS-PAGE, and MALDI-TOF-MS or UPLC ESI-MS is an efficient method for structural characterization of human and pig IgGs. This results in better targeted glycosylation and amino acid sequence analysis, without interference of Fab and Fc fragments with each other. In the context of this work, no glycopeptides from the Fab fragments of either human or pig IgG were observed, due to the intrinsic properties of the antibodies

studied. However there have been cases where ~ 40% of the Fab portion of human IgG were glycosylated [46-47] and in such cases this work would usefully apply to the analysis of these antibodies.

4.2 Discussion on MS information obtained from MALDI and UPLC for human and pig antibody fragments

MALDI-MS and UPLC-ESI-MS techniques are relatively soft (low fragmentation) ways of obtaining ions of large molecules in the gas phase, though MALDI typically produces far fewer multiply charged ions. The comparison of analyses by MALDI-ToF-MS and UPLC-ESI-MS shows consistency, in that differences between peptides and glycopeptides from the Fc and Fab fragments of both species were observed.

As for sample fractionation prior to MS, for MALDI-MS either crude tryptic digests from in-gel were analyzed at once, or separated by manual C_{18} cartridges. Therefore, most peptides and glycopeptides were analyzed at once in one sample spot on the MALDI target. As MALDI is a competitive ionization process, some analyte signals may be compromised by the presence of better ionizing species. This is the case for sialylated glycopeptides in this work, which due to their inherent negative charge do not produce much signal in the positive mode relative to neutral glycopeptides. In comparison, the UPLC-ESI-MS interface allows for much better separation of tryptic digestion products prior to MS, and then better selectivity for compound detection. As ESI is also a competitive ionization mechanism, it is preferable to have as few coeluting compounds as possible. As a result of the better selectivity, sialylated species not observed by

MALDI were detected by UPLC-MS. Overall UPLC-ESI-MS in positive mode gives more detailed data on sialylated glycopeptides, and few other glycofoms, that are usually absent from MALDI-MS spectra in positive mode. UPLC ESI-MS provided exceptionally clean product (fragment) ion chromatograms, which allowed for increased accuracy in peptide and protein analysis. Lastly, for practical reasons, it is much less time consuming to work with a direct on-line interface such as UPLC-MS than with off-line manual C_{18} cartridge-MALDI-MS.

4.3 General Conclusions

Immunoglobulin Gs are to be thoroughly analyzed after production, so as to determine the functionality of a particular antibody, and any alteration or PTM present after production. This type of analysis can be carried out by use of MALDI-TOF-MS and UPLC ESI-MS after intensive sample preparation techniques. Whole antibody digestion by trypsin is effective in producing peptides and glycopeptides, but yields no information as to where the glycopeptides, glycans and peptides are derived from exactly (from the heavy or light chain, Fab or Fc region of the antibody). Fragmentation of IgG into its Fc, Fab, or $F(ab')_2$ portions can be achieved using papain, Fabulous™ and Fabricator™ enzymes and the fragments can be separated by use of HPLC-SEC and further separated by SDS-PAGE and analyzed by MALDI-TOF-MS or UPLC ESI-MS after in-gel tryptic digestion. This latter workflow is preferred to the more conventional method of whole antibody digestion with trypsin, as it is more effective and accurate in determining exact glycosylation sites, and showing whether the glycopeptides are from the heavy or light chain or Fab or Fc regions. The analysis of antibody fragments results in elimination of interferences between the Fab and Fc regions, allowing for more detailed characterization.

4.4 Future work

Future experimentation involves further UPLC-ESI-MS or MALDI MS/MS analyses of peptides from these fragments other than the *N*-glycopeptides determined so far. Characterization of *O*-glycopetides from pig and human IgG Fab region is also to be investigated. First approach: describes the protein glycosylation-targeting enrichment technologies mainly employing solid-phase extraction methods such as hydrizide-capturing, lectin-specific capturing, and affinity separation techniques based on porous graphitized carbon, hydrophilic interaction chromatography, or immobilized boronic acid. Another approach: describes MS-based quantitative analysis strategies coupled with enzyme deglycosylation of O- glycosylated protein and O-glycosylation-targeting enrichment technologies, by using a label-free MS, stable isotope-labeling, or targeted multiple reaction monitoring (MRM) MS [48].

The enrichment techniques based on solid-phase extraction are for subsequent sample manipulations like enzyme digestion and therefore feasible to construct high-throughput MS-based analytical systems. The O-glycoprotein samples separated would subjected to enzymatic digestion to afford low molecular-weight peptide mixture or glycan mixture, and can further be purified and separated by on-line or off-line technique prior to mass analysis [48]. Mass spectrometry of the prepared peptides can be realized by label-free quantification methods or by quantitative methods with stable isotope-labeling or stable isotope-coded internal standard spiking to the prepared sample. Since great number of studies have been reported involving sample separation technologies and quantitative mass analysis technologies in the proteomic field.

References

1. Komatsu E, Buist M, Roy R, Gomes de Oliveira AG, Bodnar E, Salama A, Soulillou JP, Perreault H. Characterization of immunoglobulins through analysis of N-glycopeptides by MALDI- TOF MS. (2016)

2. http://pdl.com/technology-products/what-are-antibodies/

3. http://www.sigmaaldrich.com/catalog/product/sigma

4. .Stadlmann J, Pabst M, Kolarich D, Kunert R, Altmann F. Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. Proteomics. (2008) (14): 2858–2871.

5. www.thermofisher.com/ca/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/protein-glycosylation.html

6. Rawlings ND, Barrett AJ. Families of serine peptidases. Meth. Enzymol. Methods in Enzymology. (1994) 244: 19–61

7. Plummer, T.H. Jr. and Tarentino, A.L. Glycobiology. (1991), 257-263.

8. www.thermofisher.com/us/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/antibody-fragmentation.html

9. https://www.genovis.com/products/igg-proteases/fabulous/

10. Saito, K. Decreased Fc gamma receptor III (CD16) expression on peripheral blood mononuclear cells in patients with Sjogren's syndrome. J. Rheumatol. (1998) 689-696.

11. De Reys, S. Fc-independent crosslinking of a novel platelet membrane protein by a monoclonal antibody causes platelet activation. Blood. (1994) 547-555.

12. von Pawel-Rammingen, U. Johansson, B.P., and Björck, L. IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. EMBO Journal 21. (2002)1607-1615.

13. http://www.chemguide.co.uk/analysis/chromatography/hplc.html

14. http://www.waters.com/waters/en_CA/HPLC---High-Performance-Liquid-Chromatography-Explained/

15. Chen H, McLaughlin, K. Monoclonal Antibody Fragment Separation and Characterization Using Size Exclusion Chromatography Coupled with Mass Spectrometry.

Sepax Technologies, Inc. 5 Innovation Way Newark, DE 19711 USA

16. http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6040.pdf

17. Ninfa A, Ballou D, Benore M. Fundamental laboratory Approaches for Biochemistry and Biotechnology. United States: Wiley. (2010) 164–179. ISBN 978-0-470-08766-4

18. Weber K, Osborn M . The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. J Biol Chem. (1969) 244, 16, 4406–4412. PMID 5806584

19. Quandt N, Stindl A, Keller U. Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis for Mr Estimations of High-Molecular-Weight Polypeptides. Analytical Biochemistry. (1993)214, 2, 490–494. doi:10.1006/abio.1993.1527

20. Karas M, Hillenkamp F. Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10 000 Daltons. Anal. Chem. (1988) 60, 259–280.

21. Hillenkamp F, Karas M, Beavis RC, Chait BT. Matrixassisted Laser Desorption Ionization Mass Spectrometry of Biopolymers. Anal. Chem. (1991) 63, A1193–A1202.

22. Hillenkamp F, Karas M, Holtkamp D, Klusener P. Energy Deposition in Ultraviolet Laser Desorption Mass Spectrometry of Biomolecules. Int. J. Mass Spectrom. Ion Phys.(1986) 69, 265–276.

23. Beavis RC, Chait BT. Standing, Physics. (1990), 66, 269.

24. https://masspec.scripps.edu/publications/public_pdf/64_art.pdf

25. Yanjun J, Luzal V. The Waters Acquity Ultra-Performance Liquid Chromatograph and the Micromass Quattro Premier Triple Quadrupole Mass Spectrometer (2012).

26. IUPAC. Compendium of Chemical Terminology, 2nd ed. (the Gold Book) (1997).

27. Louris JN, Wright LG, Cooks RG, Schoen AE. New scan modes accessed with a hybrid mass spectrometer. Analytical Chemistry. (1985) 57 (14): 2918–2924. doi:10.1021/ac00291a039

28. deHoffman, E, Stroobant, V. Mass Spectrometry: Principles and Applications. Toronto: Wiley. (2003) p. 133. ISBN 0-471-48566-7.

29. Angel TE, Aryal UK, Hengel SM, Baker ES, Kelly RT, Robinson EW, Smith RD. Mass spectrometry-based proteomics: existing capabilities and future directions. Chemical Society Reviews. (2012) 41 (10): 3912–28. doi:10.1039/c2cs15331a. PMID 22498958

30. Hardouin J. Protein sequence information by matrix-assisted laser desorption/ionization in-source decay mass spectrometry. Mass Spectrometry Reviews. (2007)26(5): 672 PMID 17492750.

31. Shadforth I, Crowther D, Bessant C. Protein and peptide identification algorithms using MS for use in high-throughput, automated pipelines. Proteomics. (2005) 5 (16): 4082–95. doi:10.1002/pmic.200402091. PMID 16196103.

32. Mørtz E, O'Connor PB, Roepstorff P, Kelleher NL, Wood TD, McLafferty FW, Mann M. Sequence tag identification of intact proteins by matching tanden mass spectral data against sequence data bases. Proceedings of the National Academy of Sciences of the United States of America. (1996) 93 (16): 8264. PMID 8710858.

33. Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomedical Mass Spectrometry. (1984).11 (11): 601. doi:10.1002/bms.1200111109. PMID 6525415

34. Johnson RS, Martin SA, Biemann K. Collision-induced fragmentation of (M + H)+ ions of peptides. Side chain specific sequence ions. International Journal of Mass Spectrometry and Ion Processes. (1988) 86: 137–154. doi:10.1016/0168-1176(88)80060-0.

35. Falick AM, Hines WM, Medzihradszky KF, Baldwin MA, Gibson BW. Low-mass ions produced from peptides by high-energy collision-induced dissociation in tandem mass spectrometry. Journal of the American Society for Mass Spectrometry. (1993).PMID 24227532.

36. Downard KM, Biemann K. Methionine specific sequence ions formed by the dissociation of protonated peptides at high collision energies. Journal of Mass Spectrometry. (1995). doi:10.1002/jms.1190300106

37. Zaia J. Mass spectrometry of oligosaccharides. Mass Spectrometry Reviews. (2004). 23(3): 161–227. doi:10.1002/mas.10073. PMID 14966796.

38. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconj. J. 5 (1988). (4): 397–409. doi:10.1007/BF01049915.

39. Spina E, Cozzolino R, Ryan E, Garozzo D. Sequencing of oligosaccharides by collision-induced dissociation matrix-assisted laser desorption/ionization mass spectrometry. Journal of Mass Spectrometry. (2000). 35 (8): 1042–8. doi:10.1002/1096-9888(200008)35:8<1042::AID-JMS33>3.0.CO;2-Y. PMID 10973004

40. Morgan JW, Hettick JM, Russell DH. Peptide sequencing by MALDI 193-nm photodissociation TOF MS. Methods in Enzymology. (2005). 402: 186–209. doi:10.1016/S0076-6879(05)02006-9. PMID 16401510

41. Wells JM, McLuckey SA. Collision-induced dissociation (CID) of peptides and proteins. Methods in Enzymology. (2005). 402: 148–85. doi:10.1016/S0076-6879(05)02005-7. PMID 16401509.

42. Sleno L, Volmer DA. Ion activation methods for tandem mass spectrometry. Journal of Mass Spectrometry. (2004) 39 (10): 1091–112. doi:10.1002/jms.703. PMID 15481084

43. Buist M, Komatsu E, Lopez PG, Girard L, Bodnar E, Salama A, Sachs DH, Galli C, Perota A, Conchon S, Judor J-P, ConcordetJP, Lazzari G, Soulillou JP, Perreault H. J. Anal. Bioanal. Tech. (2016). doi:10.4172/2155-98(7(4))

44. Butler JE, Wertz N, Deschacht N. Immunogenetics (2009) 61: 209. https://doi.org/10.1007/s00251-008-0336-

45. Lopez PG , Girard L , Buist M, de Oliveira AG, Bodnar E, Salama A, Soulillou JP, Perreault H. Glycoconj. J. (2016). 33 (1):79-91. doi:10.1007/s10719-015-9637-z.

46. Dwek RA, Lellouch AC, Wormald MR. Glycobiology: the function of sugar in the IgG molecule. Journal of anatomy. (1995). 187 (Pt 2):279-292

47. Wormald MR, Rudd PM, Harvey DJ, Chang SC, Scragg IG, Dwek RA. Variations in oligosaccharide-protein interactions in immunoglobulin G determine the site-specific glycosylation profiles and modulate the dynamic motion of the Fc oligosaccharides. Biochemistry. (1997). 36 (6):1370-1380. doi:10.1021/bi9621472

48. Yeong HA, Jin YK, Jong SY. Mass Spectrometry Reviews Published by Wiley Periodicals, Inc. 2015 Apr; 34(2): 148–165.

49. Mattes MJ. Binding parameters of antibodies reacting with multivalent antigens: functional affinity or pseudo-affinity. J Immunol Methods (1997); 202:97–101

50. Kovari LC, Momany C, Rossmann MG. The use of antibody fragments for crystallization and structure determinations. (1995); 3:1291–1293

51. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3121151/

52. http://www.bio-rad.com/webroot/web/pdf/ps/literature/Bulletin_5913.pdf

53. Holliger P, Hudson PJ. Engineered antibody fragments and the rise of single domains. Nat Biotechnol (2005); 23:1126–1136

54. Ljungberg UK, Jansson B, Niss U, Nilsson R, Sandberg BE, Nilsson B. The interaction between different domains of Staphylococcal Protein A and human polyclonal IgG, IgA, IgM and F(ab′)2: separation of affinity from specificity. Mol Immunol (1993);30:1279–1285

55. Zhao Y, Gutshall L, Jiang H, et al. Two routes for production and purification of Fab fragments in biopharmaceutical discovery research: papain digestion of mAb and transient expression in mammalian cells. Protein Expr Purif (2009); 67:182–189

56. Hillson JL, Karr NS, Oppliger IR, Mannik M, Sasso EH. The structural basis of germline-encoded VH3 immunoglobulin binding to Staphylococcal Protein A. J Exp Med (1993); 178:331–336

57. Starovasnik MA, O'Connell MP, Fairbrother WJ, Kelley RF. Antibody variable region binding by Staphylococcal Protein A: thermodynamic analysis and location of the Fv binding site on E-domain. Protein Sci (1999); 8:1423–1431

58. Nelson AL, Dhimolea E, Reichert JM. Development trends for human monoclonal antibody therapeutics. Nat Rev Drug Discov (2010); 9:767–774

59. Cook GP, Tomlinson IM. The human immunoglobulin VH repertoire. Immunol Today (1995); 16:237–242

60. Davidkova G, Pettersson S, Holmberg D, Lundkvist I. Selective usage of VH genes in adult human B lymphocyte repertoires. Scand J Immunol (1997); 45:62–73

61. Sheets MD, Amersdorfer P, Finnern R, et al. Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens. Proc Natl Acad Sci USA (1998); 95:6157–6162

62. Jespers L, Schon O, Famm K, Winter G. Aggregation-resistant domain antibodies selected on phage by heat denaturation. Nat Biotechnol (2004); 22:1161–1165

63. Nilsson B, Moks T, Jansson B, et al. A synthetic IgG-binding domain based on Staphylococcal Protein A. Protein Eng (1987); 1:107–113

64. Linhult M, Gulich S, Graslund T, et al. Improving the tolerance of a Protein A analogue to repeated alkaline exposures using a bypass mutagenesis approach. Proteins (2004); 55:407–416

65. Ghose S, Allen M, Hubbard B, Brooks C, Cramer SM. Antibody variable region interactions with Protein A: implications for the development of generic purification processes. Biotechnol Bioeng (2005); 92:665–673

66. Goding, J. Monoclonal Antibodies: Principles and Practice. Academic Press Inc., London, U.K. (1983)

67. Mariani, M., et al. Mol. Immunol. (1991) 28, 69-77.

68. Sykaluk, L. Pierce Chemical Company, unpublished results (1992).

69. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2905857/

70. Lehringer A, Cox M, Nelson D. Lehringer Principles of Biochemistry, fifth edition. W.H. Freeman and company (2008).

Supplemental Information

Table S1 list of the amino acids used in determination of sequences based on their mass in IgG peptides

Amino acid	1-letter code	3-letter code	Chemical formula(-H_2O)	Monoisotopic mass (-H_2O)	Average mass (-H_2O)
Alanine	A	Ala	C_3H_5ON	71.03711	71.0788
Arginine	R	Arg	$C_6H_{12}ON_4$	156.10111	156.1875
Asparagine	N	Asn	$C_4H_6O_2N_2$	114.04293	114.1038
Aspartic Acid	D	Asp	$C_4H_5O_3N$	115.02694	115.0886
Cysteine	C	Cys	C_3H_5ONS	103.00919	103.1388
Glutamic Acid	E	Glu	$C_5H_7O_3N$	129.04259	129.1155
Glutamine	Q	Gln	$C_5H_8O_2N_2$	128.05858	128.1307
Glycine	G	Gly	C_2H_3ON	57.02146	57.0519
Histidine	H	His	$C_6H_7ON_3$	137.05891	137.1411
Isoleucine	I	Ile	$C_6H_{11}ON$	113.08406	113.1594
Leucine	L	Leu	$C_6H_{11}ON$	113.08406	113.1594
Lysine	K	Lys	$C_6H_{12}ON_2$	128.09496	128.1741
Methionine	M	Met	C_5H_9ONS	131.04049	131.1926
Phenylalanine	F	Phe	C_9H_9ON	147.06841	147.1766
Proline	P	Pro	C_5H_7ON	97.05276	97.1167
Serine	S	Ser	$C_3H_5O_2N$	87.03203	87.0782
Threonine	T	Thr	$C_4H_7O_2N$	101.04768	101.1051
Tryptophan	W	Trp	$C_{11}H_{10}ON_2$	186.07931	186.2132
Tyrosine	Y	Tyr	$C_9H_9O_2N$	163.06333	163.176
Valine	V	Val	C_5H_9ON	99.06841	99.1326

Table S2 Pig Fc *N*-glycopeptides obtained from UPLC-ESI-MS experiments

#	m/z	Retention time (min)	Mass	Charge
1319	1029.735	6.417817	**3086.182**	3
1331	1002.413	5.7548	**3004.216**	3
1329	981.3958	6.0348	**2941.166**	3
1328	940.0182	6.1491	**2817.033**	3
1327	934.6971	6.16625	**2801.07**	3
1269	927.3714	6.16625	**2779.092**	3
1326	898.6601	6.217817	**2692.958**	3
1325	891.3255	6.183383	**2670.955**	3
1268	886.0019	6.200667	**2654.984**	3
1264	880.6759	6.200667	**2639.006**	3
1340	1309.526	6.183383	**2617.038**	2
1263	873.3526	6.200667	**2617.036**	3
1339	1289.011	6.1491	**2576.008**	2
1267	854.0031	6.286383	**2558.988**	3
1338	1226.963	6.217817	**2451.911**	2
1337	1218.977	6.217817	**2435.939**	2
490	1207.987	6.200667	**2413.96**	2

Table S-3 Pig Fab peptides obtained from UPLC-ESI-MS experiments

#	m/z	Retention time (min)	Mass	Charge
773	698.637	11.72088	2092.889	3
1161	632.3182	12.13233	1893.933	3
1106	947.9706	12.13233	1893.927	2
1144	601.6572	9.6008	1801.95	3
255	425.7101	5.983383	1698.811	4
452	567.2769	5.983383	1698.809	3
1157	561.9503	5.983383	1682.829	3
988	421.7144	5.983383	1682.829	4
43	554.6285	5.983383	1660.864	3
727	831.4379	5.983383	1660.861	2
232	548.6243	5.983383	1642.851	3
862	542.6209	5.983383	1624.841	3
510	806.3457	8.7722	1610.677	2
662	795.9188	5.983383	1589.823	2

Table S-4 Human Fc *N*-glycopeptides obtained from UPLC-ESI-MS experiments

#	m/z	Retention time (min)	Mass	Charge
1324	1083.756	4.4118	3248.246	3
1311	1073.09	8.55505	3216.249	3
1310	1019.075	8.6522	3054.202	3
1309	989.7279	8.2236	2966.162	3
1306	988.7072	8.109317	2963.1	3
1272	986.7224	4.246083	2957.145	3
1305	983.3924	8.109317	2947.155	3
1317	976.0564	8.109317	2925.147	3
1304	935.7141	8.32075	2804.12	3
1303	934.6913	8.257883	2801.052	3
1270	932.7022	4.263217	2795.085	3
1266	922.0425	8.24075	2763.106	3
1316	1382.55	8.2236	2763.086	2
1300	880.6731	8.303617	2638.998	3
1298	878.6853	4.2975	2633.034	3
1296	877.6616	4.246083	2629.963	3
1299	875.3513	8.303617	2623.032	3
1294	872.3438	4.263217	2614.01	3
1315	1301.541	8.32075	2601.068	2
1260	868.0245	8.32075	2601.052	3
1262	866.995	8.137883	2597.963	3
1259	1297.007	4.246083	2591.999	2

Table S-5 Human Fab peptides obtained from UPLC-ESI-MS experiments

m/z	Retention time (min)	Mass	Charge
712.6612	7.229283	2134.962	3
640.8802	4.377517	1279.746	2
633.8856	4.377517	1265.757	2
412.9481	4.657533	1235.823	3
408.8692	14.4867	1223.586	3
605.8788	4.377517	1209.743	2
597.881	4.446083	1193.747	2
593.8258	14.46955	1185.637	2
589.8877	4.377517	1177.761	2
588.884	4.377517	1175.753	2
548.9314	4.560383	1095.848	2
528.9487	4.526083	1055.883	2
494.8993	4.428933	987.7841	2
493.9061	4.4118	985.7977	2
475.9228	4.377517	949.831	2
473.8829	4.39465	945.7512	2
472.2567	18.53275	942.4987	2
467.9272	4.62895	933.8398	2
465.8912	4.4118	929.7678	2
462.4568	4.377517	922.899	2
454.9625	4.428933	907.9105	2
450.2429	17.98417	898.4712	2
449.9329	4.428933	897.8513	2
445.9785	4.62895	889.9425	2
443.9158	4.594667	885.817	2

Table S-6 Pig Fc matrices of *N*-glycopeptides obtained from UPLC-ESI-MS experiments. This matrix was created to show correlation between related glycoforms by a loss or gain of monosaccharide or disaccharide in mass. The masses 146, 162, 203, 307 and 365Da corresponds to: (Fucose), (Hexose (Mannose, or Galactose)), (GlcNAc), (NeuGc) and (Hexose + GlcNAc) respectively.

				1319	1331	1329	1328	1327	1269	1326	1325	1268	1264	1340	1263	1339
				1029.735	1002.413	981.3958	940.0182	934.6971	927.3714	898.6601	891.3255	886.0019	880.6759	1309.526	873.3526	1289.011
				6.4	5.8	6.0	6.1	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.1
#	m/z	Retention time (min)	Mass	3086.182	3004.216	2941.166	2817.033	2801.07	2779.092	2692.958	2670.955	2654.984	2639.006	2617.038	2617.036	2576.008
1319	1029.735	6.4	3086.182													
1331	1002.413	5.8	3004.216	82												
1329	981.3958	6.0	2941.166	145	63											
1328	940.0182	6.1	2817.033	269	187	124										
1327	934.6971	6.2	2801.07	285	203	140	16									
1269	927.3714	6.2	2779.092	307	225	162	38	22								
1326	898.6601	6.2	2692.958	393	311	248	124	108	86							
1325	891.3255	6.2	2670.955	415	333	270	146	130	108	22						
1268	886.0019	6.2	2654.984	431	349	286	162	146	124	38	16					
1264	880.6759	6.2	2639.006	447	365	302	178	162	140	54	32	16				
1340	1309.526	6.2	2617.038	469	387	324	200	184	162	76	54	38	22			
1263	873.3526	6.2	2617.036	469	387	324	200	184	162	76	54	38	22	0		
1339	1289.011	6.1	2576.008	510	428	365	241	225	203	117	95	79	63	41	41	
1267	854.0031	6.3	2558.988	527	445	382	258	242	220	134	112	96	80	58	58	17
1338	1226.963	6.2	2451.911	634	552	489	365	349	327	241	219	203	187	165	165	124
1337	1218.977	6.2	2435.939	650	568	505	381	365	343	257	235	219	203	181	181	140
490	1207.987	6.2	2413.96	672	590	527	403	387	365	279	257	241	225	203	203	162

Table S-7 Human Fc matrices of *N*-glycopeptides obtained from UPLC-ESI-MS experiments. The masses 146, 162, 203, 291 and 365Da corresponds to: (Fucose), (Hexose (Mannose, or Galactose)), (GlcNAc), (NeuAc) and (Hexose + GlcNAc) respectively.

				1324	1311	1310	1309	1306	1272	1305	1317	1304	1303	1270	1266	1316
				1083.756	1073.09	1019.075	989.7279	988.7072	986.7224	983.3924	976.0564	935.7141	934.6913	932.7022	922.0425	1382.55
				4.4118	8.55505	8.6522	8.2236	8.109317	4.246083	8.109317	8.109317	8.32075	8.257883	4.263217	8.24075	8.2236
#	m/z	Retention time (min)	**Mass**	3248.246	3216.249	3054.202	2966.162	2963.1	2957.145	2947.155	2925.147	2804.12	2801.052	2795.085	2763.106	2763.086
1324	1083.756	4.4118	3248.246													
1311	1073.09	8.55505	3216.249	32												
1310	1019.075	8.6522	3054.202	194	162											
1309	989.7279	8.2236	2966.162	282	250	88										
1306	988.7072	8.109317	2963.1	285	253	91	3									
1272	986.7224	4.246083	2957.145	291	259	97	9	6								
1305	983.3924	8.109317	2947.155	301	269	107	19	16	10							
1317	976.0564	8.109317	2925.147	323	291	129	41	38	32	22						
1304	935.7141	8.32075	2804.12	444	412	250	162	159	153	143	121					
1303	934.6913	8.257883	2801.052	447	415	253	165	162	156	146	124	3				
1270	932.7022	4.263217	2795.085	453	421	259	171	168	162	152	130	9	6			
1266	922.0425	8.24075	2763.106	485	453	291	203	200	194	184	162	41	38	32		
1316	1382.55	8.2236	2763.086	485	453	291	203	200	194	184	162	41	38	32	0	
1300	880.6731	8.303617	2638.998	609	577	415	327	324	318	308	286	165	162	156	124	124
1298	878.6853	4.2975	2633.034	615	583	421	333	330	324	314	292	171	168	162	130	130
1296	877.6616	4.246083	2629.963	618	586	424	336	333	327	317	295	174	171	165	133	133
1299	875.3513	8.303617	2623.032	625	593	431	343	340	334	324	302	181	178	172	140	140
1294	872.3438	4.263217	2614.01	634	602	440	352	349	343	333	311	190	187	181	149	149
1315	1301.541	8.32075	2601.068	647	615	453	365	362	356	346	324	203	200	194	162	162

NB* All information in proceeding data tables were obtained from the UniprotKB site, for the determination and verification of the amino acid sequences of the peptide masses of Fab and Fc of both human and pig IgG observed in tables 3.1 and 3.2, which were derived from thesis experiments. All amino acid sequences in tables below are as results of tryptic digests, with number of missed cleavage (#MC) observed, in theoretical experiments done on UniProtKB.

Table S-8

IGHG1_HUMAN (P01857) from UniProtKB/Swiss-Prot sequence:

```
         10         20         30         40         50         60
aSTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS

         70         80         90        100        110        120
GLYSLSSVVT VPSSSLGTQT YICNVNHKPS NTKVDKKVEP KSCDKTHTCP PCPAPELLGG

        130        140        150        160        170        180
PSVFLFPPKP KDTLMISRTP EVTCVVVDVS HEDPEVKFNW YVDGVEVHNA KTKPREEQYN

        190        200        210        220        230        240
STYRVVSVLT VLHQDWLNGK EYKCKVSNKA LPAPIEKTIS KAKGQPREPQ VYTLPPSRDE

        250        260        270        280        290        300
LTKNQVSLTC LVKGFYPSDI AVEWESNGQP ENNYKTTPPV LDSDGSFFLY SKLTVDKSRW

        310        320        330
QQGNVFSCSV MHEALHNHYT QKSLSLSPGK
```

Chain Ig gamma-1 chain C region at positions <1 - 330

Mass	position	#Missed Cleavage (MC)	peptide sequence
7901.9316	17-93	1	STSGGTAALGCLVKDYFPEP VTVSWNSGALTSGVHTFPAV LQSSGLYSLSSVVTVPSSSL GTQTYICNVNHKPSNTK
6998.4832	31-96	1	DYFPEPVTVSWNSGALTSGV HTFPAVLQSSGLYSLSSVVT VPSSSLGTQTYICNVNHKPS NTKVDK

6656.2929	31-93	0	DYFPEPVTVSWNSGALTSGV HTFPAVLQSSGLYSLSSVVT VPSSSLGTQTYICNVNHKPS NTK
4399.0353	254-292	1	GFYPSDIAVEWESNGQPENN YKTTPPVLDSDGSFFLYSK
3740.7900	139-171	1	TPEVTCVVVDVSHEDPEVKF NWYVDGVEVHNAK
3629.7216	244-275	1	NQVSLTCLVKGFYPSDIAVE WESNGQPENNYK
3546.8309	106-138	1	THTCPPCPAPELLGGPSVFL FPPKPKDTLMISR
3513.6790	300-330	1	WQQGNVFSCSVMHEALHNHY TQKSLSLSPGK
3163.5777	102-131	1	SCDKTHTCPPCPAPELLGGP SVFLFPPKPK
2987.3788	298-322	1	SRWQQGNVFSCSVMHEALHN HYTQK
2978.5006	176-200	1	EEQYNSTYRVVSVLTVLHQD WLNGK
2898.4223	132-157	1	DTLMISRTPEVTCVVVDVSH EDPEVK
2744.2456	300-322	0	WQQGNVFSCSVMHEALHNHY TQK
2730.4145	106-131	0	THTCPPCPAPELLGGPSVFL FPPKPK
2544.1313	254-275	0	GFYPSDIAVEWESNGQPENN YK
2432.2853	5-30	1	GPSVFPLAPSSKSTSGGTAA LGCLVK
2430.2438	276-297	1	TTPPVLDSDGSFFLYSKLTV DK
2228.2073	185-203	1	VVSVLTVLHQDWLNGKEYK
2160.0984	158-175	1	FNWYVDGVEVHNAKTKPR
2082.0059	139-157	0	TPEVTCVVVDVSHEDPEVK

1873.9218	276-292	0	TTPPVLDSDGSFFLYSK
1872.9701	228-243	1	EPQVYTLPPSRDELTK
1808.0064	185-200	0	VVSVLTVLHQDWLNGK
1724.9078	224-238	1	GQPREPQVYTLPPSR
1690.9044	239-253	1	DELTKNQVSLTCLVK
1677.8019	158-171	0	FNWYVDGVEVHNAK
1671.8085	172-184	1	TKPREEQYNSTYR
1573.8584	1-16	1	ASTKGPSVFPLAPSSK
1286.6739	228-238	0	EPQVYTLPPSR
1267.7620	210-221	1	ALPAPIEKTISK
1266.7416	206-217	1	VSNKALPAPIEK
1264.6565	17-30	0	STSGGTAALGCLVK
1189.5120	176-184	0	EEQYNSTYR
1186.6466	5-16	0	GPSVFPLAPSSK
1104.6081	244-253	0	NQVSLTCLVK
905.4397	98-105	1	VEPKSCDK
838.5032	210-217	0	ALPAPIEK
835.4342	132-138	0	DTLMISR
818.4730	293-299	1	LTVDKSR

788.4512	323-330	0	SLSLSPGK
678.3603	204-209	1	CKVSNK
670.3229	201-205	1	EYKCK
656.3838	222-227	1	AKGQPR
647.4086	218-223	1	TISKAK
605.3141	239-243	0	DELTK
600.3715	97-101	1	KVEPK
575.3399	293-297	0	LTVDK
501.3143	172-175	0	TKPR
1085.6425	218-227	2	TISKAKGQPR

Table S-9 Human IgG Kappa 1 light chain variable region, peptide sequence recovered from UniProtKB

mass	position	#MC	peptide sequence
5199.4581	56-103	1	TGVPSRFSGGGSATNFTVTI SSLQPEDFATYYCQQYHHLP FTFGPGTK
5091.3933	62-107	1	FSGGGSATNFTVTISSLQPE DFATYYCQQYHHLPFTFGPG TKVDFK
4602.1346	62-103	0	FSGGGSATNFTVTISSLQPE DFATYYCQQYHHLPFTFGPG TK
4239.0233	1-38	1	DIQMTQSPSSLSASVGDRVT FICQASQDIANHLNWYQK
3086.5516	19-45	1	VTFICQASQDIANHLNWYQK KPGEAPK
2379.1550	19-38	0	VTFICQASQDIANHLNWYQK
1910.0421	39-55	1	KPGEAPKFLIYDGSFLK
1878.8862	1-18	0	DIQMTQSPSSLSASVGDR
1799.9690	46-61	1	FLIYDGSFLKTGVPSR
1202.6455	46-55	0	FLIYDGSFLK
947.5309	108-116	1	RTVAAPSVF
791.4297	109-116	0	TVAAPSVF
726.4144	39-45	0	KPGEAPK
664.3777	104-108	1	VDFKR
616.3413	56-61	0	TGVPSR
508.2765	104-107	0	VDFK

Table S-10 Human Ig kappa chain C region peptide sequence recovered from UniProtKB

Mass	Position	#MC	peptide sequence
3667.8828	1-34	1	TVAAPSVFIFPPSDEQLKSG TASVVCLLNNFYPR
3619.7093	42-75	1	VDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSK
2677.2700	38-61	1	VQWKVDNALQSGNSQESVTE QDSK
2323.1499	83-103	1	VYACEVTHQGLSSPVTKSFN R
2135.9687	42-61	0	VDNALQSGNSQESVTEQDSK
2109.0234	62-80	1	DSTYSLSSTLTLSKADYEK
2084.0593	81-99	1	HKVYACEVTHQGLSSPVTK
2069.0484	19-37	1	SGTASVVCLLNNFYPREAK
1946.0269	1-18	0	TVAAPSVFIFPPSDEQLK
1818.9054	83-99	0	VYACEVTHQGLSSPVTK
1740.8737	19-34	0	SGTASVVCLLNNFYPR
1502.7584	62-75	0	DSTYSLSSTLTLSK
890.4366	76-82	1	ADYEKHK
888.4937	35-41	1	EAKVQWK
812.3355	100-106	1	SFNRGEC
625.2828	76-80	0	ADYEK
560.3191	38-41	0	VQWK
523.2623	100-103	0	SFNR

Table S-11 IGHG3_HUMAN (P01860) sequence recovered from UniProtKB/Swiss-Prot. Immunoglobulin heavy constant gamma 3 (Ig gamma-3 chain C region)

```
        10         20         30         40         50         60
ASTKGPSVFP LAPCSRSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS

        70         80         90        100        110        120
GLYSLSSVVT VPSSSLGTQT YTCNVNHKPS NTKVDKRVEL KTPLGDTTHT CPRCPEPKSC

       130        140        150        160        170        180
DTPPPCPRCP EPKSCDTPPP CPRCPEPKSC DTPPPCPRCP APELLGGPSV FLFPPKPKDT

       190        200        210        220        230        240
LMISRTPEVT CVVVDVSHED PEVQFKWYVD GVEVHNAKTK PREEQYNSTF RVVSVLTVLH

       250        260        270        280        290        300
QDWLNGKEYK CKVSNKALPA PIEKTISKTK GQPREPQVYT LPPSREEMTK NQVSLTCLVK

       310        320        330        340        350        360
GFYPSDIAVE WESSGQPENN YNTTPPMLDS DGSFFLYSKL TVDKSRWQQG NIFSCSVMHE

       370
ALHNRFTQKS LSLSPGK
```

mass	position	#MC	peptide sequence
7889.8952	17-93	1	STSGGTAALGCLVKDYFPEP VTVSWNSGALTSGVHTFPAV LQSSGLYSLSSVVTVPSSSL GTQTYTCNVNHKPSNTK
6986.4468	31-96	1	DYFPEPVTVSWNSGALTSGV HTFPAVLQSSGLYSLSSVVT VPSSSLGTQTYTCNVNHKPS NTKVDK
6644.2565	31-93	0	DYFPEPVTVSWNSGALTSGV HTFPAVLQSSGLYSLSSVVT VPSSSLGTQTYTCNVNHKPS NTK
5475.5347	291-339	1	NQVSLTCLVKGFYPSDIAVE WESSGQPENNYNTTPPMLDS DGSFFLYSK
4946.2665	301-344	1	GFYPSDIAVEWESSGQPENN YNTTPPMLDSDGSFFLYSKL TVDK
4389.9444	301-339	0	GFYPSDIAVEWESSGQPENN YNTTPPMLDSDGSFFLYSK
3754.8057	186-218	1	TPEVTCVVVDVSHEDPEVQF KWYVDGVEVHNAK
3173.5493	179-206	1	DTLMISRTPEVTCVVVDVSH EDPEVQFK

3147.5827	149-178	1	SCDTPPPCPRCPAPELLGGP SVFLFPPKPK
2962.5057	223-247	1	EEQYNSTFRVVSVLTVLHQD WLNGK
2910.5619	159-185	1	CPAPELLGGPSVFLFPPKPK DTLMISR
2761.3086	347-369	1	WQQGNIFSCSVMHEALHNRF TQK
2500.1721	345-365	1	SRWQQGNIFSCSVMHEALHN R
2476.2686	5-30	1	GPSVFPLAPCSRSTSGGTAA LGCLVK
2357.1329	186-206	0	TPEVTCVVVDVSHEDPEVQF K
2257.0389	347-365	0	WQQGNIFSCSVMHEALHNR
2228.2073	232-250	1	VVSVLTVLHQDWLNGKEYK
2094.1456	159-178	0	CPAPELLGGPSVFLFPPKPK
1904.9422	275-290	1	EPQVYTLPPSREEMTK
1898.9871	207-222	1	WYVDGVEVHNAKTKPR
1852.8680	102-118	1	TPLGDTTHTCPRCPEPK
1808.0064	232-247	0	VVSVLTVLHQDWLNGK
1767.9058	98-113	1	VELKTPLGDTTHTCPR
1724.9078	271-285	1	GQPREPQVYTLPPSR
1722.8764	286-300	1	EEMTKNQVSLTCLVK
1655.8136	219-231	1	TKPREEQYNSTFR
1626.7073	114-128	1	CPEPKSCDTPPPCPR
1626.7073	119-133	1	SCDTPPPCPRCPEPK

1626.7073	129-143	1	CPEPKSCDTPPPCPR
1626.7073	134-148	1	SCDTPPPCPRCPEPK
1626.7073	144-158	1	CPEPKSCDTPPPCPR
1617.8417	1-16	1	ASTKGPSVFPLAPCSR
1416.6906	207-218	0	WYVDGVEVHNAK
1298.6157	102-113	0	TPLGDTTHTCPR
1292.7208	366-377	1	FTQKSLSLSPGK
1286.6739	275-285	0	EPQVYTLPPSR
1267.7620	257-268	1	ALPAPIEKTISK
1266.7416	253-264	1	VSNKALPAPIEK
1264.6565	17-30	0	STSGGTAALGCLVK
1230.6299	5-16	0	GPSVFPLAPCSR
1173.5171	**223-231**	**0**	**EEQYNSTFR**
1104.6081	291-300	0	NQVSLTCLVK
1072.4550	119-128	0	SCDTPPPCPR
1072.4550	134-143	0	SCDTPPPCPR
1072.4550	149-158	0	SCDTPPPCPR
838.5032	257-264	0	ALPAPIEK
835.4342	179-185	0	DTLMISR
818.4730	340-346	1	LTVDKSR
788.4512	370-377	0	SLSLSPGK

686.3944	269-274	1	TKGQPR
678.3603	251-256	1	CKVSNK
677.4192	265-270	1	TISKTK
670.3229	248-252	1	EYKCK
644.4090	97-101	1	RVELK
637.2861	286-290	0	EEMTK
575.3399	340-344	0	LTVDK
573.2701	114-118	0	CPEPK
573.2701	129-133	0	CPEPK
573.2701	144-148	0	CPEPK
523.2875	366-369	0	FTQK
517.3092	94-97	1	VDKR
501.3143	219-222	0	TKPR

Table S-12 IGHG2_HUMAN (P01859) sequence recovered from UniProtKB/Swiss-Prot. Chain Immunoglobulin heavy constant gamma 2

-
```
         10         20         30         40         50         60
aSTKGPSVFP LAPCSRSTSE STAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS

         70         80         90        100        110        120
GLYSLSSVVT VPSSNFGTQT YTCNVDHKPS NTKVDKTVER KCCVECPPCP APPVAGPSVF

        130        140        150        160        170        180
LFPPKPKDTL MISRTPEVTC VVVDVSHEDP EVQFNWYVDG VEVHNAKTKP REEQFNSTFR

        190        200        210        220        230        240
VVSVLTVVHQ DWLNGKEYKC KVSNKGLPAP IEKTISKTKG QPREPQVYTL PPSREEMTKN

        250        260        270        280        290        300
QVSLTCLVKG FYPSDISVEW ESNGQPENNY KTTPPMLDSD GSFFLYSKLT VDKSRWQQGN

        310        320
VFSCSVMHEA LHNHYTQKSL SLSPGK
```

mass	position	#MC	peptide sequence
8053.9062	17-93	1	STSESTAALGCLVKDYFPEP VTVSWNSGALTSGVHTFPAV LQSSGLYSLSSVVTVPSSNF GTQTYTCNVDHKPSNTK
7048.4261	31-96	1	DYFPEPVTVSWNSGALTSGV HTFPAVLQSSGLYSLSSVVT VPSSNFGTQTYTCNVDHKPS NTKVDK
6706.2358	31-93	0	DYFPEPVTVSWNSGALTSGV HTFPAVLQSSGLYSLSSVVT VPSSNFGTQTYTCNVDHKPS NTK
4557.1700	128-167	1	DTLMISRTPEVTCVVVDVSH EDPEVQFNWYVDGVEVHNAK
4447.0023	250-288	1	GFYPSDISVEWESNGQPENN YKTTPPMLDSDGSFFLYSK
4223.0501	135-171	1	TPEVTCVVVDVSHEDPEVQF NWYVDGVEVHNAKTKPR
3740.7536	135-167	0	TPEVTCVVVDVSHEDPEVQF NWYVDGVEVHNAK
3645.7166	240-271	1	NQVSLTCLVKGFYPSDISVE WESNGQPENNYK
3513.6790	296-326	1	WQQGNVFSCSVMHEALHNHY TQKSLSLSPGK

3496.7321	102-134	1	CCVECPPCPAPPVAGPSVFL FPPKPKDTLMISR
2987.3788	294-318	1	SRWQQGNVFSCSVMHEALHN HYTQK
2932.4951	172-196	1	EEQFNSTFRVVSVLTVVHQD WLNGK
2808.4107	101-127	1	KCCVECPPCPAPPVAGPSVF LFPPKPK
2744.2456	296-318	0	WQQGNVFSCSVMHEALHNHY TQK
2680.3158	102-127	0	CCVECPPCPAPPVAGPSVFL FPPKPK
2578.3003	5-30	1	GPSVFPLAPCSRSTSESTAA LGCLVK
2560.1262	250-271	0	GFYPSDISVEWESNGQPENN YK
2462.2159	272-293	1	TTPPMLDSDGSFFLYSKLTV DK
2214.1917	181-199	1	VVSVLTVVHQDWLNGKEYK
1905.8939	272-288	0	TTPPMLDSDGSFFLYSK
1904.9422	224-239	1	EPQVYTLPPSREEMTK
1793.9908	181-196	0	VVSVLTVVHQDWLNGK
1724.9078	220-234	1	GQPREPQVYTLPPSR
1722.8764	235-249	1	EEMTKNQVSLTCLVK
1639.8187	168-180	1	TKPREEQFNSTFR
1617.8417	1-16	1	ASTKGPSVFPLAPCSR
1366.6882	17-30	0	STSESTAALGCLVK
1286.6739	224-234	0	EPQVYTLPPSR
1253.7463	206-217	1	GLPAPIEKTISK
1252.7259	202-213	1	VSNKGLPAPIEK

1230.6299	5-16	0	GPSVFPLAPCSR
1157.5222	**172-180**	**0**	**EEQFNSTFR**
1104.6081	240-249	0	NQVSLTCLVK
846.4679	94-100	1	VDKTVER
835.4342	128-134	0	DTLMISR
824.4876	206-213	0	GLPAPIEK
818.4730	289-295	1	LTVDKSR
788.4512	319-326	0	SLSLSPGK
686.3944	218-223	1	TKGQPR
678.3603	200-205	1	CKVSNK
677.4192	214-219	1	TISKTK
670.3229	197-201	1	EYKCK
637.2861	235-239	0	EEMTK
632.3726	97-101	1	TVERK
575.3399	289-293	0	LTVDK
504.2776	97-100	0	TVER
501.3143	168-171	0	TKPR

Table S-13 Sus Scrofa (Pig) IgG- Uniprotkb code identifier-(L8B0S7_PIG), IgG2 heavy chain. Sequence recovered:

-
```
        10         20         30         40         50         60
             E EKLVESGGGL VQPGGSLRLS CVGSGFTFNS TYINWVRQAP

        70         80         90        100        110        120
GKGLEWLAGI ARGSYSGSSY YADSVKGRFT ISRDNSRNTD YLQMNSLRTE DTARYYCARG

       130        140        150        160        170        180
VSYCYSFGAY CYDQYYYALD LWGPGVEVVV SSAPKTAPSV YPLAPCSRDT SGPNVALGCL

       190        200        210        220        230        240
ASSYFPEPVT VTWNSGALSS GVHTFPSVLQ PSGLYSLSSM VTVPASSLSS KSYTCNVNHP

       250        260        270        280        290        300
ATTTKVDKRV GTKTKPPCPI CPACESPGPS VFIFPPKPKD TLMISRTPQV TCVVVDVSQE

       310        320        330        340        350        360
NPEVQFSWYV DGVEVHTAQT RPKEEQFNST YRVVSVLPIQ HQDWLNGKEF KCKVNNKDLP

       370        380        390        400        410        420
APITRIISKA KGQTREPQVY TLPPHAEELS RSKVSITCLV IGFYPPDIDV EWQRNGQPEP

       430        440        450        460        470        480
EGNYRTTPPQ QDVDGTYFLY SKFSVDKASW QGGGIFQCAV MHEALHNHYT QKSISKTPGK
```

mass	position	#MC	peptide sequence
7889.8299	169-245	1	DTSGPNVALGCLASSYFPEP VTVTWNSGALSSGVHTFPSV LQPSGLYSLSSMVTVPASSL SSKSYTCNVNHPATTTK
7714.8069	156-231	1	TAPSVYPLAPCSRDTSGPNV ALGCLASSYFPEPVTVTWNS GALSSGVHTFPSVLQPSGLY SLSSMVTVPASSLSSK
6372.1366	169-231	0	DTSGPNVALGCLASSYFPEP VTVTWNSGALSSGVHTFPSV LQPSGLYSLSSMVTVPASSL SSK
5362.4886	120-168	1	GVSYCYSFGAYCYDQYYYAL DLWGPGVEVVVSSAPKTAPS VYPLAPCSR
5326.5385	287-332	1	TPQVTCVVVDVSQENPEVQF SWYVDGVEVHTAQTRPKEEQ FNSTYR
4988.4556	280-323	1	DTLMISRTPQVTCVVVDVSQ ENPEVQFSWYVDGVEVHTAQ TRPK
4676.0923	115-155	1	YYCARGVSYCYSFGAYCYDQ YYYALDLWGPGVEVVVSSAP K

4172.0393	287-323	0	TPQVTCVVVDVSQENPEVQF SWYVDGVEVHTAQTRPK
4019.8182	120-155	0	GVSYCYSFGAYCYDQYYYAL DLWGPGVEVVVSSAPK
3691.7849	394-425	1	VSITCLVIGFYPPDIDVEWQ RNGQPEPEGNYR
3657.8482	23-57	1	LVESGGGLVQPGGSLRLSCV GSGFTFNSTYINWVR
3553.8077	254-286	1	TKPPCPICPACESPGPSVFI FPPKPKDTLMISR
3389.5942	443-472	1	FSVDKASWQGGGIFQCAVMH EALHNHYTQK
3228.5465	448-476	1	ASWQGGGIFQCAVMHEALHN HYTQKSISK
3201.4759	415-442	1	NGQPEPEGNYRTTPPQQDVD GTYFLYSK
3122.6239	250-279	1	VGTKTKPPCPICPACESPGP SVFIFPPKPK
2987.5009	324-348	1	EEQFNSTYRVVSVLPIQHQD WLNGK
2813.3035	448-472	0	ASWQGGGIFQCAVMHEALHN HYTQK
2737.3914	254-279	0	TKPPCPICPACESPGPSVFI FPPKPK
2665.3694	392-414	1	SKVSITCLVIGFYPPDIDVE WQR
2632.2976	39-62	1	LSCVGSGFTFNSTYINWVRQ APGK
2537.2306	63-86	1	GLEWLAGIARGSYSGSSYYA DSVK
2536.2242	426-447	1	TTPPQQDVDGTYFLYSKFSV DK
2450.2424	394-414	0	VSITCLVIGFYPPDIDVEWQ R
2308.1680	372-391	1	GQTREPQVYTLPPHAEELSR
2237.2077	333-351	1	VVSVLPIQHQDWLNGKEFK
2151.0328	39-57	0	LSCVGSGFTFNSTYINWVR
2081.0661	376-393	1	EPQVYTLPPHAEELSRSK
2027.9451	98-114	1	NTDYLQMNSLRTEDTAR

1959.9334	426-442	0	TTPPQQDVDGTYFLYSK
1912.0134	20-38	1	EEKLVESGGGLVQPGGSLR
1878.9014	232-248	1	SYTCNVNHPATTTKVDK
1865.9392	376-391	0	EPQVYTLPPHAEELSR
1833.0017	333-348	0	VVSVLPIQHQDWLNGK
1826.8450	94-108	1	DNSRNTDYLQMNSLR
1683.7609	73-88	1	GSYSGSSYYADSVKGR
1566.8750	58-72	1	QAPGKGLEWLAGIAR
1536.7111	232-245	0	SYTCNVNHPATTTK
1525.8332	23-38	0	LVESGGGLVQPGGSLR
1470.6383	73-86	0	GSYSGSSYYADSVK
1361.6882	156-168	0	TAPSVYPLAPCSR
1354.6420	98-108	0	NTDYLQMNSLR
1348.5950	109-119	1	TEDTARYYCAR
1337.7535	354-365	1	VNNKDLPAPITR
1323.7994	358-369	1	DLPAPITRIISK
1260.5603	415-425	0	NGQPEPEGNYR
1173.5171	**324-332**	**0**	**EEQFNSTYR**
1095.5541	89-97	1	FTISRDNSR
1085.6102	63-72	0	GLEWLAGIAR
882.5043	358-365	0	DLPAPITR
836.4737	87-93	1	GRFTISR
835.4342	280-286	0	DTLMISR
817.4778	473-480	1	SISKTPGK

705.3712	352-357	1	CKVNNK
692.3209	109-114	0	TEDTAR
675.2919	115-119	0	YYCAR
660.3787	370-375	1	AKGQTR
659.4450	366-371	1	IISKAK
654.3279	349-353	1	EFKCK
623.3511	89-93	0	FTISR
595.3086	443-447	0	FSVDK
560.3514	249-253	1	RVGTK
517.3092	246-249	1	VDKR
500.2827	58-62	0	QAPGK

Table S-14 L8B180_PIG (L8B180) from UniProtKB/TrEMBL. IgG1 heavy chain sequence:

```
        10         20         30         40         50         60
                    E VKLVESGGGL VQPGGSLRLS CVGSGFTFNT YNMIWVRQVP

        70         80         90        100        110        120
GKGLEGLAYI GYTGVITDYA DSVKGRFTIS RDNSENTAYL QMNSLRTEDT ARYYCARNYF

       130        140        150        160        170        180
AGDLYAMDLW GPGVEVVVSS APKTAPSVYP LAPCGRDTSG PNVALGCLAS SYFPEPVTVT

       190        200        210        220        230        240
WNSGALTSGV HTFPSVLQPS GLYSLSSTVT VPARSSSRKC FTCNVNHPAT TTKVDLCVGR

       250        260        270        280        290        300
PCPICPACEG NGPSVFIFPP KPKDTLMISR TPEVTCVVVD VSQENPEVQF SWYVDGVEVH

       310        320        330        340        350        360
TAQTRPKEAQ FNSTYRVVSV LPIQHQDWLK GKEFECKVNN KDLPAPITRI ISKAKGPSRE

       370        380        390        400        410        420
PQVYTLSPSA EELSRSKVSI TCLVTGFYPP DIDVEWKSNG QPEPEGNYRT TPPQQDVDGT

       430        440        450        460
YFLYSKFSVD KARLQSGDTF QCAVMHEALH NHYTQKSISK TQGN
```

mass	position	#MC	peptide sequence
7235.6132	144-214	1	TAPSVYPLAPCGRDTSGPNV ALGCLASSYFPEPVTVTWNS GALTSGVHTFPSVLQPSGLY SLSSTVTVPAR
6340.1506	157-218	1	DTSGPNVALGCLASSYFPEP VTVTWNSGALTSGVHTFPSV LQPSGLYSLSSTVTVPARSS SR
5922.9534	157-214	0	DTSGPNVALGCLASSYFPEP VTVTWNSGALTSGVHTFPSV LQPSGLYSLSSTVTVPAR
5269.5170	271-316	1	TPEVTCVVVDVSQENPEVQF SWYVDGVEVHTAQTRPKEAQ FNSTYR
4989.4396	264-307	1	DTLMISRTPEVTCVVVDVSQ ENPEVQFSWYVDGVEVHTAQ TRPK
4658.2307	220-263	1	CFTCNVNHPATTTKVDLCVG RPCPICPACEGNGPSVFIFP PKPK
4173.0233	271-307	0	TPEVTCVVVDVSQENPEVQF SWYVDGVEVHTAQTRPK

4098.0139	118-156	1	NYFAGDLYAMDLWGPGVEVV VSSAPKTAPSVYPLAPCGR
3956.9715	234-270	1	VDLCVGRPCPICPACEGNGP SVFIFPPKPKDTLMISR
3701.8566	23-57	1	LVESGGGLVQPGGSLRLSCV GSGFTFNTYNMIWVR
3610.7158	378-409	1	VSITCLVTGFYPPDIDVEWK SNGQPEPEGNYR
3441.6282	113-143	1	YYCARNYFAGDLYAMDLWGP GVEVVVSSAPK
3288.5079	398-426	1	SNGQPEPEGNYRTTPPQQDV DGTYFLYSK
3140.5551	234-263	0	VDLCVGRPCPICPACEGNGP SVFIFPPKPK
3073.4618	434-460	1	LQSGDTFQCAVMHEALHNHY TQKSISK
2885.3570	432-456	1	ARLQSGDTFQCAVMHEALHN HYTQK
2814.4559	58-84	1	QVPGKGLEGLAYIGYTGVIT DYADSVK
2785.3541	118-143	0	NYFAGDLYAMDLWGPGVEVV VSSAPK
2758.4311	308-330	1	EAQFNSTYRVVSVLPIQHQD WLK
2704.3374	39-62	1	LSCVGSGFTFNTYNMIWVRQ VPGK
2658.2187	434-456	0	LQSGDTFQCAVMHEALHNHY TQK
2536.2242	410-431	1	TTPPQQDVDGTYFLYSKFSV DK
2518.2823	63-86	1	GLEGLAYIGYTGVITDYADS VKGR
2497.2683	376-397	1	SKVSITCLVTGFYPPDIDVE WK
2429.0997	92-112	1	DNSENTAYLQMNSLRTEDTA R
2360.1299	87-106	1	FTISRDNSENTAYLQMNSLR
2305.1598	63-84	0	GLEGLAYIGYTGVITDYADS VK
2282.1413	378-397	0	VSITCLVTGFYPPDIDVEWK
2203.0989	356-375	1	GPSREPQVYTLSPSAEELSR
2195.0412	39-57	0	LSCVGSGFTFNTYNMIWVR

2021.0185	360-377	1	EPQVYTLSPSAEELSRSK
1959.9334	410-426	0	TTPPQQDVDGTYFLYSK
1882.0392	20-38	1	EVKLVESGGGLVQPGGSLR
1847.0537	317-332	1	VVSVLPIQHQDWLKGK
1805.8915	360-375	0	EPQVYTLSPSAEELSR
1755.7966	92-106	0	DNSENTAYLQMNSLR
1664.7883	219-233	1	KCFTCNVNHPATTTK
1661.9373	317-330	0	VVSVLPIQHQDWLK
1536.6933	220-233	0	CFTCNVNHPATTTK
1525.8332	23-38	0	LVESGGGLVQPGGSLR
1348.5950	107-117	1	TEDTARYYCAR
1347.5924	398-409	0	SNGQPEPEGNYR
1337.7535	338-349	1	VNNKDLPAPITR
1331.6776	144-156	0	TAPSVYPLAPCGR
1323.7994	342-353	1	DLPAPITRIISK
1115.5116	**308-316**	**0**	**EAQFNSTYR**
1110.5248	333-341	1	EFECKVNNK
882.5043	342-349	0	DLPAPITR
840.3920	331-337	1	GKEFECK
836.4737	85-91	1	GRFTISR
835.4342	264-270	0	DTLMISR

834.4316	457-464	1	SISKTQGN
822.4468	427-433	1	FSVDKAR
692.3209	107-112	0	TEDTAR
675.2919	113-117	0	YYCAR
659.4450	350-355	1	IISKAK
655.2756	333-337	0	EFECK
623.3511	87-91	0	FTISR
615.3573	354-359	1	AKGPSR
595.3086	427-431	0	FSVDK
564.3100	215-219	1	SSSRK
528.3140	58-62	0	QVPGK