

Thesis for the Master's degree in
chemistry

Nebiyu Abshiru

**Oxidative footprinting and
mass spectrometry based
structural characterization of
mAb 14F7-NeuGc GM3
ganglioside complex.**

60 study points

DEPARTMENT OF CHEMISTRY
Faculty of mathematics and natural sciences

UNIVERSITY OF OSLO 05/2010



Acknowledgements

I would like to express my deepest gratitude to my supervisor, Prof. Ute Krenzel, for accepting me into this interesting project. Her guidance and supervision throughout the duration of my thesis enabled me to work enthusiastically.

I am also heartily thankful to my co-supervisors Prof. Robert Woods and Dr. Wolfgang Egge-Jacobsen, who allowed me to work in their lab throughout the project. I am thankful to all the group members in Rob's lab at the CCRC, University of Georgia, for all the help they provided me. They made my stay in Georgia pleasant experience.

I also want to express my sincere thanks to lab members in Oslo, Gabriele, Andre, Vincent, Al and Anders (IMBV) for proving themselves helpful and enjoyable fellows.

A big thank you goes to my friends and families for the encouragement and support during all of these years.

Last but not least I would like to acknowledge the Norwegian quota scheme for accepting my application and for the financial support

Abstract

The characterization of antibody-antigen interactions provides crucial information on the structural basis of the specificity of binding of antibodies to their target antigens. In this project, an attempt was made to characterize the binding interactions between a monoclonal antibody (mAb) designated as 14F7 and NeuGc-GM3 ganglioside antigen. The 14F7 strongly recognizes the ganglioside antigen on human melanoma and breast tumors. The crystal structure of the 14F7 Fab fragment has been solved by X-ray crystallography and binding to the ganglioside predicted by docking and MD simulations. However, despite significant efforts, so far no experimental structural analysis of the 14F7-NeuGc-GM3 complex is available. Such information would be extremely valuable for developing new and improved antibodies for cancer immunotherapy.

Therefore, in this thesis, an alternative approach was taken to get experimental data. The binding interactions within the complex were examined using a newly developed MS-based oxidative footprinting technique. Briefly, the antibody sample, with and without the ligand, was exposed to an excimer laser source in order to oxidize solvent accessible amino acid side chains. The oxidized samples were then digested with trypsin. Finally, the tryptic fragments were analyzed by MALDI-TOF MS and ESI-LTQ-FT MS.

Sequencing of the MS/MS spectra and database searching identified two residues, Trp325 of the VL-CDR3 and Met112 of the VH-CDR3, that were oxidized in both the free and ligand-bound 14F7. The percentages of oxidation of the peptides containing the residues were quantified based on the areas under the selected ion chromatograms of the oxidized and non-oxidized forms of the peptides. The data suggest that the extent of oxidation of both peptides is significantly lower in the ligand-bound 14F7. Moreover, an analysis of the spectral intensities and side chain solvent accessibilities confirmed that residues Met112 and Trp325 are protected from radical oxidation upon ligand binding to the antibody. A new docking model of the complex is presented that is consistent with all experimental data. The combined result, in contrast to the previous model, suggests that the ligand binds in close proximity to the light chain, in addition to binding to the heavy chain CDR3 of the antibody. This information can now be used for further development of anti-tumor antibody with improved potency and affinity for target tumor antigens.

Abbreviations

AB	Applied biosystems
CCRC	Complex carbohydrate research center
CDR	Complementarity determining region
CID	Collision induced dissociation
CIM	Center for molecular immunology (Havana, Cuba)
Da	Dalton
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECD	Electron capture dissociation
EDTA	Ethylenediamine-tetraacetate
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
ETD	Electron transfer dissociation
ExPASy	Expert protein analysis system
FTICR	Fourier transform ion cyclotron resonance
GM3	The ganglioside: Neu5Ac(Gc) α 3Gal β 4Glc β 1Cer
IgG	Immunoglobulin G
LC	Liquid chromatography
LTQ	Linear trap quadrupole
mAb	Monoclonal antibody
MALDI	Matrix assisted laser desorption/ionization
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
<i>m/z</i>	Mass to charge ratio
NMR	Nuclear magnetic resonance
Neu5Ac	N-acetylneuraminic acid
Neu5Gc	N-glycolylneuraminic acid
N-14F7	Non-oxidized 14F7 mAb alone
O-14F7	Oxidized 14F7 mAb alone

OL-14F7	Oxidized ligand-bound 14F7 mAb
PMF	Peptide mass fingerprinting
pNPP	para-nitrophenyl phosphate
PTM	Post-translational modification
QIT	Quadrupole ion trap
SASA	Solvent accessible surface area
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIC	Selected ion chromatogram
TIC	Total ion chromatogram
TOF	Time-of-flight
TBS	Tris buffered saline
TBST	Tris buffered saline – Tween 20
UV	Ultraviolet
3-D	3-dimensional
VH	Variable heavy chain
VL	Variable light chain
r.m.s.d.	Root mean square deviation/distance

Table of Contents

Acknowledgements.....	2
Abstract.....	3
Abbreviations	4
1. INTRODUCTION.....	8
1.1 Cancer.....	8
1.2 Cancer immunotherapy based on monoclonal antibodies.....	9
1.2.1 Monoclonal antibodies	9
1.3 Gangliosides as target tumor antigens.....	12
1.4 Antibody-antigen interactions	14
1.5 Aim of the study.....	15
2. METHODS: A BRIEF BACKGROUND.....	18
2.1 Mass spectrometry (MS).....	19
2.1.1 Basic components of a mass spectrometer	21
2.1.2 Tandem mass spectrometry (MS/MS).....	26
2.1.3 Liquid chromatography/mass spectrometry (LC/MS).....	27
2.2 Peptide identification and characterization using MS.....	29
2.2.1 Peptide mass fingerprinting.....	29
2.2.2 Tandem MS of peptides and database searching.....	32
2.3 Characterization of protein-ligand interactions using oxidative footprinting and MS.....	35
2.3.1 Hydroxyl radicals	37
2.3.2 Methods to generate hydroxyl radicals for protein footprinting.....	37
2.3.3 Amino acid oxidation by hydroxyl radicals	39
3. EXPERIMENTAL.....	48
3.1 Materials and Reagents	48
3.2 Equipments.....	49
3.3 Methods.....	50
3.3.1 Sample preparation.....	50
3.3.2 Laser photolysis.....	51
3.3.3 Proteolytic digestion.....	51
3.3.4 Peptide mass fingerprinting.....	52
3.3.5 LC-MS/MS analysis	53
3.3.6 Database searching and spectral assignment	55

3.3.7 Docking simulations of the 14F7 - NeuGc-GM3 complex	56
4. RESULTS AND DISCUSSION.....	59
4.1 Sample preparation.....	59
4.1.1 SDS-PAGE.....	59
4.1.2 ELISA.....	60
4.2 MALDI-MS analysis of N-14F7 and O-14F7 mAb digests.....	61
4.2.1 Peptide mass matching using the ExPASy database search	61
4.2.2 Percentage sequence coverage of 14F7 digest	64
4.2.3 MALDI-TOF spectra - why are some peptides less abundant?.....	65
4.3 ESI-LC/MS/MS analysis of O-14F7 and OL-14F7 digests	70
4.3.1 LC-MS/MS total ion chromatograms	70
4.3.2 Peptide and modification sites identified using ByOnic	71
4.3.3 MS/MS spectra – sequence and oxidation site confirmations	75
4.3.4 Percentage oxidation and level of protection of modified peptides	80
4.4 A new structural model of 14F7-NeuGc-GM3 complex	84
4.5 Correlation of the footprinting result with SASA of residue side chains.....	88
5. CONCLUSION AND FURTHER WORKS	91
APPENDIX.....	93
Appendix 1	93
Appendix 2	94
Appendix 3	95
Appendix 4	97
Appendix 5	104
Appendix 6.....	105
REFERENCES	106

1. INTRODUCTION

1.1 Cancer

Cancer is one of the major public health problems in our world. In the United States only, a total of 1.5 million cases and 0.5 million deaths from cancer were projected to occur in 2009 [1]. Of the 57% of the estimated new cancer cases in women in the year, breast cancer alone was expected to account for 27%. The disease is characterized by uncontrolled multiplication of cells to form lumps of tissue called tumors [2]. Under normal conditions, rapid cell divisions occur in our body when there is a need to replace damaged cells or during wound healing. For instance, the lymphoid tissue generates enormous amounts of white blood cells when there is a need to fight infections [3]. However, cancerous cells tend to grow and divide rapidly without a needful signal by other cells. The abnormal growth or division can occur in two ways [4]. In one way, the cells multiply aberrantly without spreading to adjacent tissues. This forms a solid tumor called benign tumor. The second and more dangerous one is known as malignant tumor. Unlike the benign tumor that normally remains in one spot, a malignant tumor spreads throughout the body to conquer as many tissues as possible.

Some of the known causes of cancer are genetic abnormalities during DNA replication (or cell division), carcinogens and infectious diseases. Gene mutations that may occur during faulty DNA replication can cause abnormal cell divisions leading to cancer. Mutations that cause cancer can be inherited from our parents or sometimes they are due to environmental factors such as carcinogenic substances like asbestos, tobacco and toxic chemical compounds.

Tumor cells usually display special molecules called antigens on their outer surface [5-6]. Our immune system may recognize the cells via these signal molecules. Generally, the tumor antigens are absent in normal cells (or they exist in a relatively lower concentration). However, in individuals bearing a tumor, the concentration of the antigens aberrantly increases in concentration [7]. The immune system fails to respond effectively despite this unexpected rise in the concentration of the antigen on the surface of the tumor. Moreover, it fails to discriminate between the tumor and the normal cells.

Thus, boosting the potency of our immunity by generating substances outside our body, which specifically target the tumor-associated surface antigens, is currently a promising method of treatment. Other treatment methods such as chemotherapy and surgery are also practiced. In chemotherapy, pharmaceutically produced drugs are used to prevent tumor cells from damaging adjacent cells or from growing rapidly by dividing.

1.2 Cancer immunotherapy based on monoclonal antibodies

Cancer immunotherapy is a treatment method by which the potency of our immune system is enhanced against tumor surface antigens [8]. Generally, there are two classes of cancer immunotherapy: Active immunotherapy and passive immunotherapy. Active immunotherapy aims at the induction of patients' immune responses by immunization with tumor associated antigens or anti-idiotypic antibodies. In passive immunotherapy, the biological functions of monoclonal antibodies against tumor antigens are utilized to enhance the immune response. In general, antibody-based cancer immunotherapy is a form of passive immunotherapy whereby antibodies generated outside the body are used to target tumor associated surface antigens [9].

1.2.1 Monoclonal antibodies

Antibodies are proteins made by certain white blood cells (B cells) in response to antigens [10]. They tag the antigen for destruction by the immune system. Structurally, they are 'Y' shaped molecules composed of four polypeptides – two identical heavy chains each containing one variable (VH) and three constant (CH1, CH2 and CH3) domains and two light chains each containing one variable (VL) and one constant (CL) domain (see Figure 1.1) [11]. Proteolytic cleavage of antibodies by papain releases three fragments- one Fc (Fragment, crystallizable) and two Fab (Fragment, antigen binding) fragments. Within the variable domains there are hypervariable regions called complementarity-determining regions (CDR). These regions determine the antibody's specificity and affinity for binding an antigen [12].

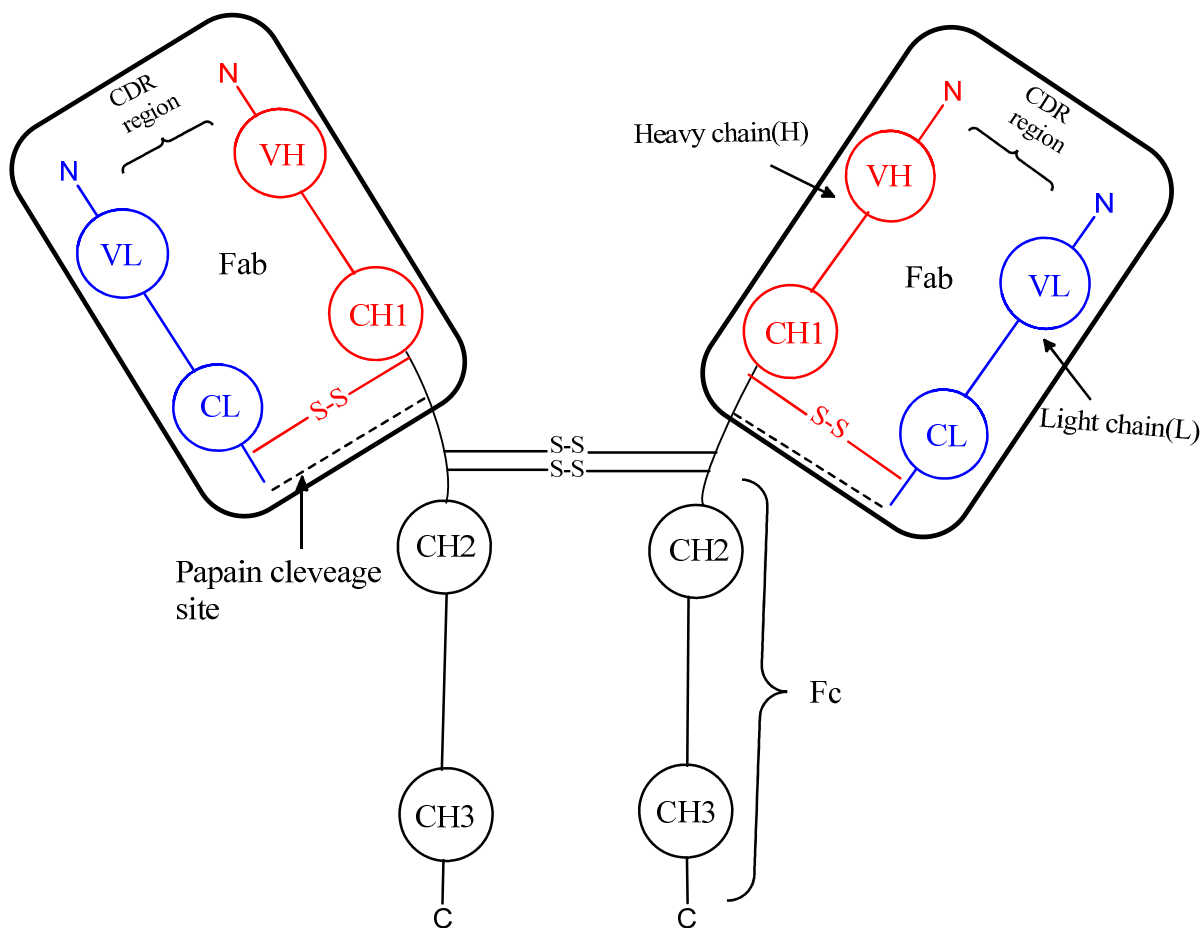


Figure 1.1 Antibody structure (adapted from Ref. [13]). Picture made with ChemBioDraw Ultra 12 [14].

Antibodies generated from the same cell line by fusion of an antibody forming cell (such as mouse spleen cell) and a cancerous cell (such as multiple myeloma cells) are known as monoclonal antibodies (mAb) [15]. A given spleen cell isolated from mice produces a unique antibody and grows only for few days in culture. However, fusion with a myeloma cell that can divide indefinitely produces a hybridoma (see Figure 1.2). The hybridoma cell is capable of generating multiple copies of a unique monoclonal antibody specific for the antigen of interest. The mAbs are generated from a single (mono) clone (the hybridoma cell) hence the name monoclonal antibody. Until now, several mAbs that can uniquely detect tumor associated ganglioside antigens have been generated [16-18].

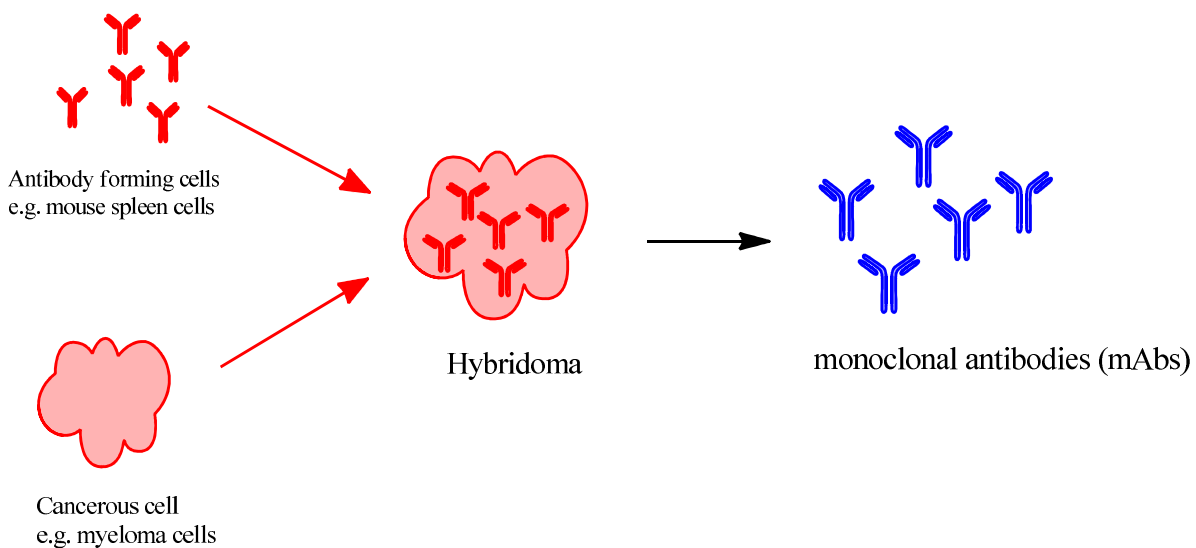


Figure 1.2: A schematic diagram showing the generation of monoclonal antibodies from a hybridoma cell line. (Picture made with ChemBioDraw Ultra 12)

1.3 Gangliosides as target tumor antigens

Some biological macromolecules - proteins, carbohydrates and lipids can covalently combine to form hybrid macromolecules. The combination of a carbohydrate and a protein or lipid forms a glycoconjugate. There are three types of glycoconjugates: Proteoglycans (composed of heteropolysaccharide and polypeptide), glycoproteins (composed of protein and carbohydrate) and glycolipids (hybrid of lipid and carbohydrate). Gangliosides are a class of membrane-associated sialic acid-containing glycolipids [19]. They contain a special type of lipid component called a sphingolipid and a sialic acid-containing carbohydrate component. Gangliosides play an important role in cell growth and immune recognition and serve as surface membrane receptors [20].

The sialic acid moiety in membrane-associated gangliosides acts as the specific site for molecular recognition on cell surfaces. In the past, various types of sialic acids have been detected in different parts of our cells including in cancerous cells. The two most abundant sialic acids found in animals are N-acetyl (NeuAc) and N-glycolyl (NeuGc) neuraminic acids [21]. The former one is found in human tissues, whereas Neu5Gc, which is structurally very similar to the N-acetylated form (see Figure 1.3), is generally absent in normal human tissues [22]. NeuGc-GM3 and NeuAc-GM3 gangliosides are examples of glycosphingolipids containing these sialic acids. Irie *et al.* reported that the minimum amount of an immunostained N-glycolyl GM3 ganglioside detected in melanoma tissue is 5 nanogram [22]. On the other hand at most 5 picogram of the ganglioside is detected in a milligram of normal human tissue. Moreover, an increased amount of the ganglioside was observed in human breast cancer [23]. These findings led to the idea that the N-glycolyl variant of the GM3 ganglioside can be used as a target antigen in antibody-based breast and melanoma cancer immunotherapy.

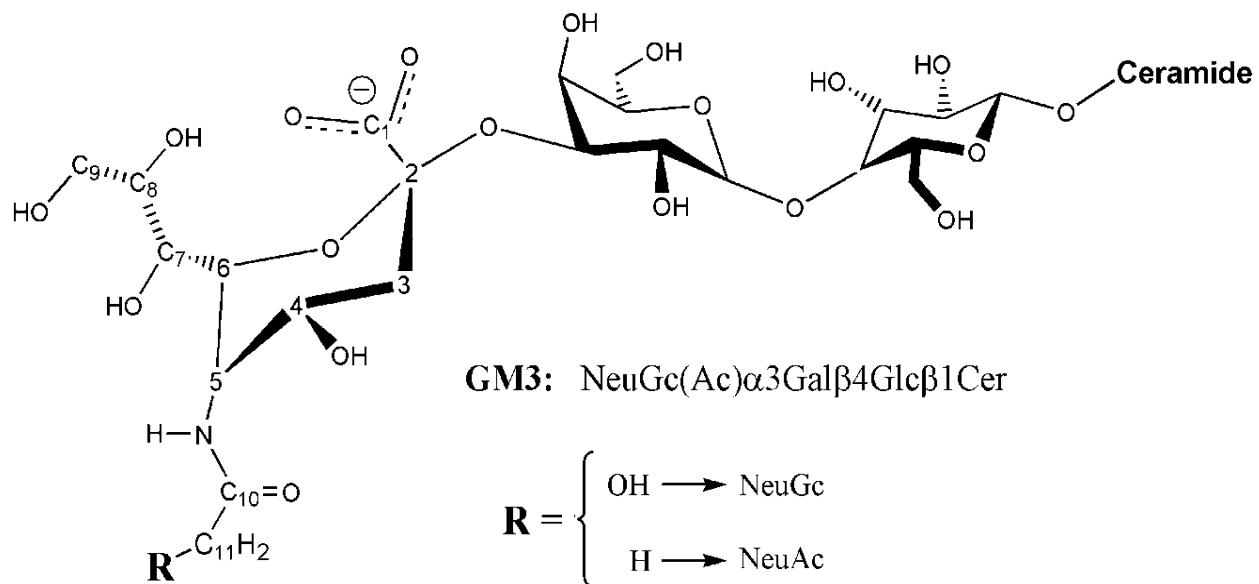


Figure 1.3: The structure of the N-glycolyl (NeuGc) and the N-acetyl (NeuAc) GM3 gangliosides (Repeated from Ref. [24]). The gangliosides differ only by a single oxygen atom in the R-group of the N-glycolyl form.

1.4 Antibody-antigen interactions

Proteins accomplish their task by interacting with other molecules - ligands. A ligand can be another protein, a carbohydrate, a nucleic acid or it can be a simple molecule such as oxygen. The interaction between an antibody and an antigen is a typical example of a protein-ligand interaction. The antigen-specificity of antibodies is the basis for the immune system's ability to recognize several disease-causing antigens. The variations in the amino acid sequences in the variable (VL and VH) domains of the Fab region provide the basis for the antigen-specificity of the antibodies [25]. Each of these domains consists of three CDRs (VL CDR1, CDR2, CDR3 and VH CDR1, CDR2, CDR3) at their tip, ready to bind a target antigen. Potential hydrogen bonding interactions between the residues within the CDRs and the antigen hold the antibody-antigen complex stable. In addition, complementarity of the surface of the molecules and several van der Waals interactions contribute for the stability of the complex. A review of antibody structures by Davies *et al.* [11] shows that aromatic residues such as phenylalanine, tyrosine and tryptophan indeed are often exposed to solvent and contribute large areas to the antigen-binding surface of the CDRs.

The three-dimensional structures of a large number of antibody-antigen complexes have been determined [26-27]. These complex structures provide us with a general structural picture of the antigen-binding sites of antibodies. In general, the CDR loops in the VH and VL domains are clustered to form the binding sites. In addition, the amino acid side chains of the CDRs form continuous binding surface with protrusions and depressions. In most of the Fab-antigen complex structures, the binding interactions involve all six CDRs, whereby each CDR contributes at least one residue to the interaction with the antigen [28-29]. Furthermore, ligand binding to an antibody often does not cause a large conformational change to the structure of the complementarity determining sites [24].

In this project, in particular, an attempt is made to characterize the binding interactions between a unique anti-tumor antibody, 14F7, and the N-glycolyl GM3 ganglioside. This will be discussed in detail in the following section.

1.5 Aim of the study

The primary focus of this study is on a monoclonal anti-tumor antibody called 14F7. This antibody was obtained from hybridoma cells prepared by combining the spleen cell and myeloma cell lines of Balb/c mice, immunized with a vaccine containing N-glycolyl GM3 ganglioside [30]. Reactivity tests of the 14F7 mAb against several gangliosides showed that the antibody did not react with NeuGc containing gangliosides (for instance NeuGc-GM1, NeuGc-GM2), except NeuGc-GM3. In addition, 14F7 is able to discriminate between the N-glycolyl and N-acetyl forms of GM3 and specifically binds the N-glycolylated GM3 ganglioside. The fact that the structure of the N-glycolyl GM3 differs from the N-acetyl GM3 only by a single oxygen atom (as shown previously in Figure 1.3) makes the 14F7 specificity very unique. *In vitro* tissue specificity tests showed that the antibody recognizes NeuGc-GM3 ganglioside antigens expressed on breast and melanoma tumor sections. *In vivo*, the anti-tumor effect of the antibody was demonstrated in mice myeloma cells [31]. The antibody effectively inhibits myeloma tumor growth. Furthermore, clinical trials with the 14F7 mAb in human cancer patients are in progress, with promising results (E. Moreno, CIM, personal communication).

The crystal structure of the 14F7 Fab has been solved at 2.5 Å resolution by the Krengel group [24]. The conformations of the six CDR regions of the Fab fragments are shown in Figure 1.4A. Because the antibody has so far resisted crystallization with its ligand, docking and molecular dynamics simulations were used to predict the interactions between the 14F7 Fab and the terminal disaccharide NeuGc(α 2-3)Gal β of the N-glycolyl GM3 ganglioside. In the model, the saccharide is exclusively bound to the heavy chain CDR regions of the 14F7 Fab, through a large number of hydrogen bonding interactions (see Figure 1.4B). Six out of eleven potential hydrogen bonding interactions between the disaccharide and the Fab fragment are predicted to be with amino acids in the VH-CDR3 region [24]. A phage display and light-chain shuffling analysis indicated that the specificity of 14F7 resides entirely in its VH region, as the VL domain could be extensively mutated without compromises in antibody binding properties [32]. In contrast, two site-directed mutations of VH Asp52 to valine or isoleucine were shown to completely abolish binding to the ligand, pointing to a critical role of this residue for ligand binding.

The target of the present study was therefore to provide new detailed experimental data on the binding interactions within the 14F7-NeuGc-GM3 complex. The 14F7 mAb was analyzed by oxidative footprinting and mass spectrometry in the free and bound state. Characterization of any possible interactions that may exist between the CDRs of the 14F7 mAb and the NeuGc-GM3 ganglioside was the focus of this project. The levels of oxidation of CDR peptides containing modified residues were examined in the free and ligand-bound 14F7. The regions participating in the interaction with the ligand, as compared to the free 14F7, were expected to have a reduced level of oxidation.

The study, in addition to providing new experimental evidence on the binding interactions in the complex, contributes to the development of the footprinting and mass spectrometry techniques utilized in this project.

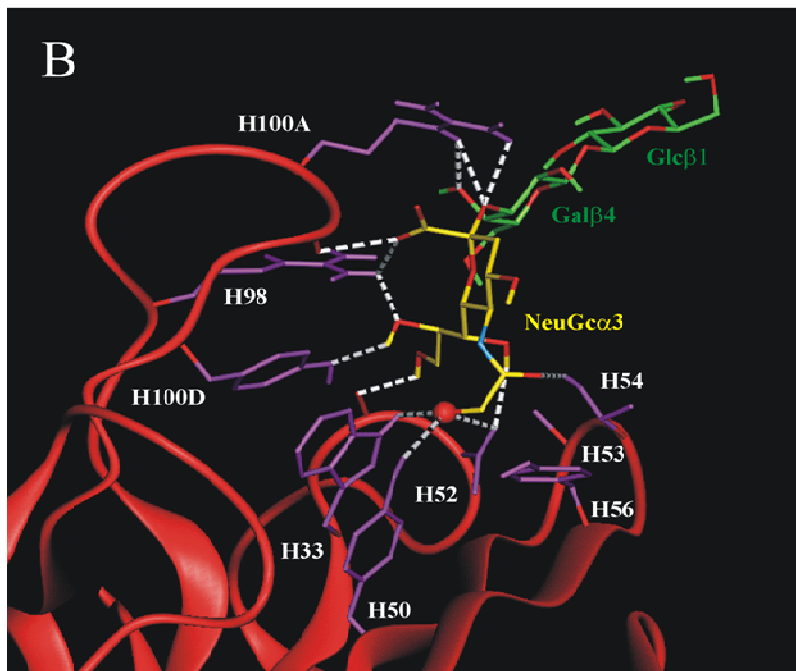


Figure 1.4: Fab 14F7. (A) Crystal structure, (B) 14F7-NeuGc-GM3 complex model [24].

2. METHODS: A BRIEF BACKGROUND

The structure of a protein-ligand complex is studied for various purposes. First, it provides us with a three-dimensional view of the binding interactions in the complex. Second, based on the structure of the complex it is possible to predict the molecular recognition mechanism, catalysis and regulatory functions of the proteins [33]. Several methods are being used for determining the structure and dynamics of protein-ligand complexes. X-ray crystallography and NMR are by far the leading methods. X-ray crystallography is providing a high resolution and detailed structural solution for vast number of proteins (and peptides) that have pharmaceutical and biological importance [34]. NMR has been a method of choice for in-solution structural and dynamic studies of protein-ligand complexes, although sample analysis is mostly limited to lower molecular weight biomolecules [35]. In this project, a newly developed MS-based oxidative footprinting technique is used to study the interaction between the anti-tumor antibody 14F7 and the N-glycolyl GM3 ganglioside.

2.1 Mass spectrometry (MS)

A mass spectrometer is an analytical tool that measures the mass-to-charge ratio (m/z) of ionized (charged) analytes, based on the response of the ions to applied electric and/or magnetic fields. This analytical tool is mainly used to determine accurate molecular weight of gas-phase ions of organic and biomolecules [36]. The result from a mass spectrometric analysis of a sample is obtained in the form of mass spectra. Each mass spectrum is a plot indicating the m/z value of a molecular ion (or its fragments) versus the relative abundance. The three basic components of the mass spectrometry are the ion source, analyzer and detector as illustrated in Figure 2.1. Each component is discussed briefly in the following subsection.

The fundamental requirement for a mass spectrometer to determine the m/z values of ions is that the analyte should be in the gas phase. As a result, in the past, only volatile molecules were amenable to analysis by MS. In the 1980s, the advent of the major soft ionization techniques such as electrospray ionization (ESI) [37] and matrix assisted laser desorption/ionization (MALDI) [38] enabled the analysis of nonvolatile and thermally labile molecules such as proteins and nucleic acids. Since then mass spectrometry has become a method of choice in many analytical experiments involving biomolecules. It has greatly revolutionized biological science in protein identification and quantitation and localization of post-translational modifications. The delicate selectivity, sensitivity and accuracy of MS together with the ability of 2-D based SDS-PAGE to separate complex protein mixtures make it a forefront analytical tool in the field of proteomics. Mass spectrometry core facilities are now established in many research institutes in most parts of the world.

The other area where MS has achieved a significant development is the structural characterization of macromolecular complexes. Together with other methods such as oxidative footprinting and hydrogen-deuterium exchange, mass spectrometry has enabled the study of the structure and dynamics of protein-ligand complexes. This has a special significance for drug design in pharmaceutical industries. Mass spectrometry is being extensively used in the discovery and assay of biomarkers in biopharmaceuticals, in medicine and forensic sciences [39-40].

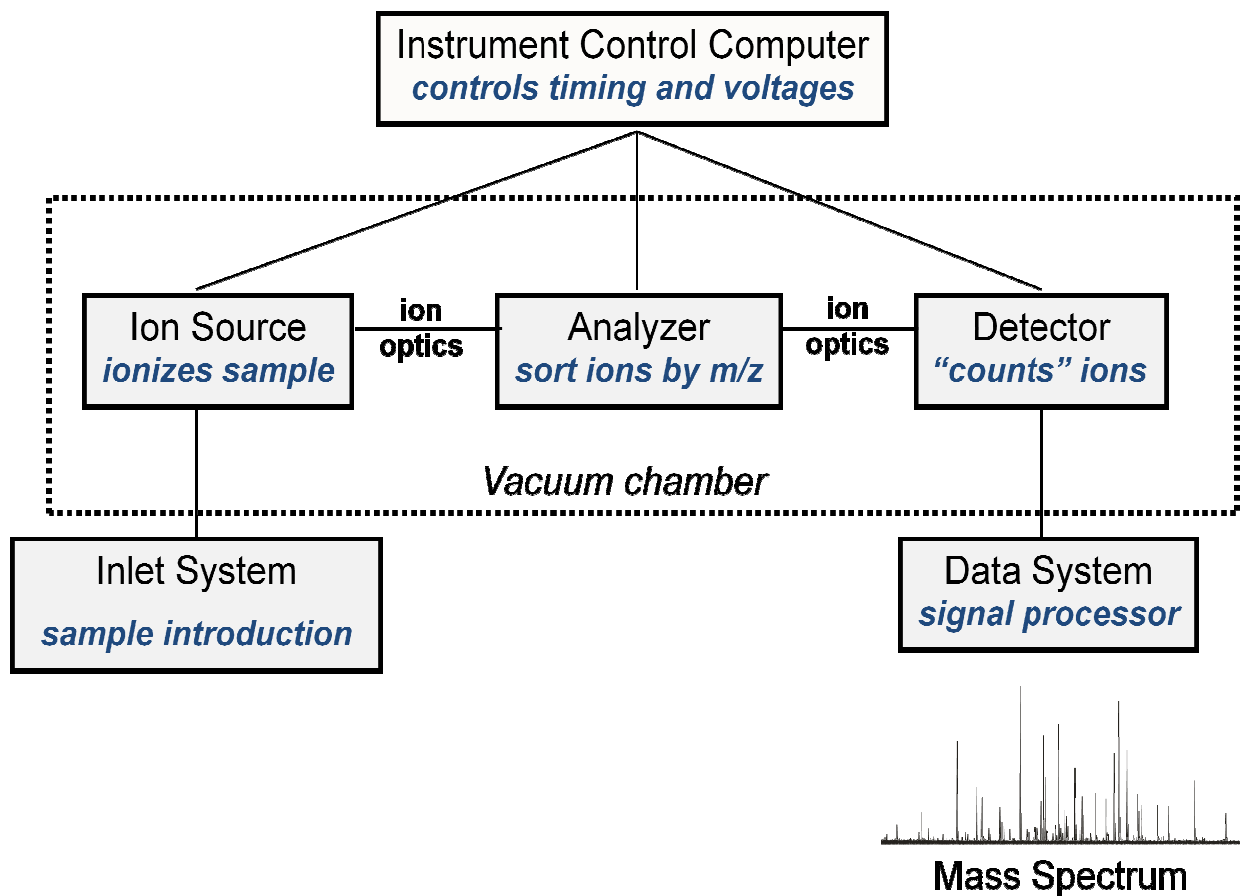


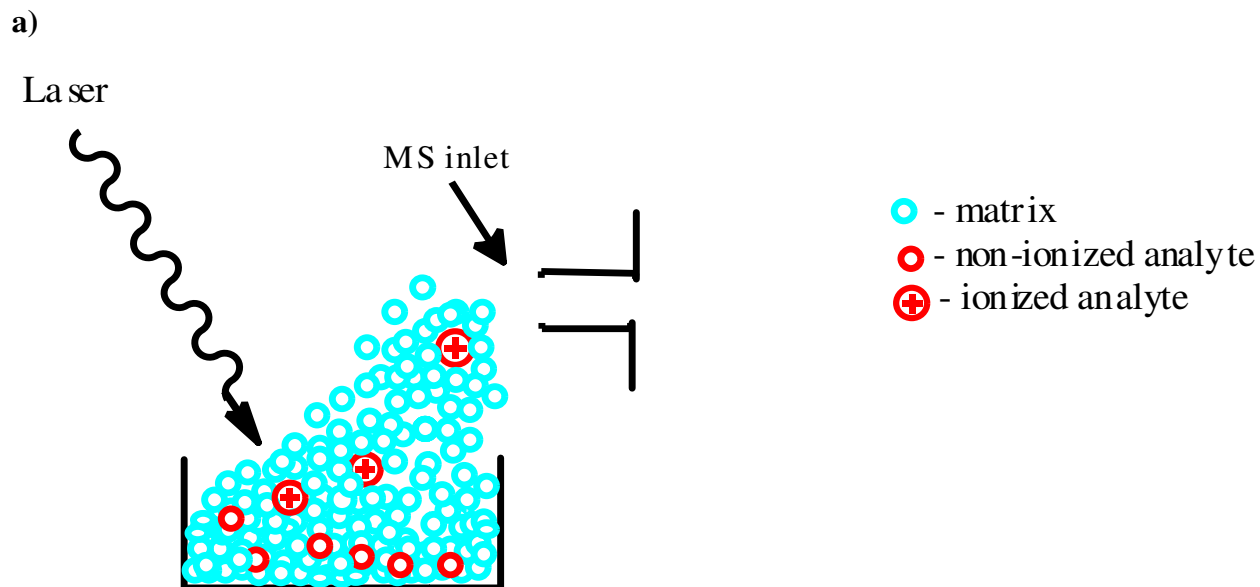
Figure 2.1: Basic components of a mass spectrometer (picture repeated from the MS lecture by Prof. I. Jonathan Amster, University of Georgia, U.S.A.)

2.1.1 Basic components of a mass spectrometer

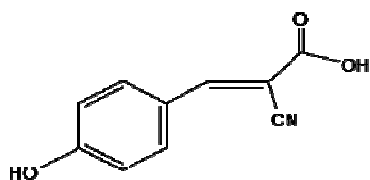
2.1.1.1 Ion source

Ion source is the component of a mass spectrometer where ionization of analyte molecules takes place. There are two types of ionization methods: Soft ionization and hard ionization. The hard ionization method, in addition to forming ions of the molecules, results in breaking of bonds. It is suitable for ionization of thermally stable compounds. The soft ionization method allows formation of ions of analyte molecules without causing fragmentation. The two most commonly used soft ionization techniques for biomolecules are MALDI and ESI.

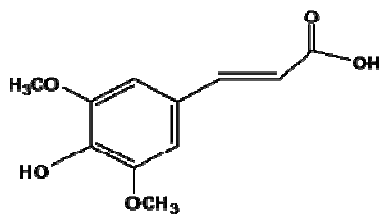
Matrix assisted laser desorption/ionization (MALDI) is a soft ionization technique which enables the analysis of high-mass ions of biological macromolecules. The original research report for the development of MALDI was published in 1988 [41], for which K. Tanaka was awarded the Nobel Prize in chemistry in 2002 [42]. Previously, plasma desorption (PD-MS) [43] and fast atom bombardment (FAB) [44] techniques were used to effectively ionize analytes with relatively lower molecular mass. These days, analytes having a molecular mass greater than 300 kDa can be ionized using MALDI [45]. Laser desorption of biomolecules such as nucleic acids, peptides and proteins most commonly generate singly protonated ions (see Figure 2.2a). In practice, sample molecules should be co-crystallized with a volatile matrix to minimize fragmentation during laser exposure. The quality of the MS spectra largely depends on the choice of matrix and the composition of the matrix/analyte mixture. Commonly used matrices include α -cyano-4-hydroxy cinnamic acid (α -CHCA) [46], 3,5-dimethoxy-4-hydroxy cinnamic acid (sinapinic acid) [47] and 2,5-dihydroxybenzoic acid (DHB) [48]. The α -CHCA is suitable for peptides with masses less than 10 kDa, whereas sinapinic acid is suitable for peptides/proteins with masses greater than 10 kDa.



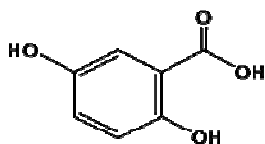
b)



α -cyano-4-hydroxycinnamic acid
(α -cyano)
Peptide matrix



4-hydroxy-3,5-methoxy cinnamic acid
(sinapinic acid)
Protein matrix



2,5-dihydroxybenzoic acid acid
(DHB)
Peptide matrix

Figure 2.2: MALDI MS (a) sample ionization process, (b) matrices.

(Pictures made with ChemBioDraw Ultra 12).

Electrospray ionization (ESI) is a soft ionization technique that results in little to no fragmentation of analyte molecules. The ionization of large molecular weight biomolecules was a major challenge until the 1980s. Then, for the first time J. B. Fern described the generation of ions using an electrospray method [37]. He won the Nobel Prize in 2002 for this outstanding achievement in the field of chemistry.

ESI allows liquid droplets of analytes to be sprayed via a capillary tube into a mass spectrometer. In typical ESI, the analyte is dissolved in a mixture of polar (such as formic acid) and volatile (such as acetonitrile) organic solvents. Then, it is sprayed into an electric field generated between the tip of the spraying capillary tube and a counter electrode. The electric field helps to form highly charged liquid droplets and keeps them from freezing. The charged droplets are then evaporated into the atmosphere in the presence of a drying gas such as nitrogen. The gas passes across the tip of the spray needle, thereby evaporating the solvent molecules leaving a solvent-free charged analyte with a diminished size. The decrease in droplet size in turn results in increased repulsive force between neighboring like charges in the droplet. When the repulsive force between the charges reaches the Rayleigh limit (the maximum repulsive force that can exist between like charges before they overcome cohesive forces), they disintegrate into smaller droplets due to the effect known as ‘coulombic explosion’. This process generates ions with different mass to charge ratio, which directly flow into a mass spectrometer (see Figure 2.3a). Moreover, the use of the ESI method produces ions with multiple charges, thereby enabling the determination of the molecular weight of analytes that have a larger mass than the upper mass limit of the mass spectrometer. So far, analysis of a protein molecule up to at least 130 kDa using ESI method was reported [49]. The different charge states in protein MS are formed due to multiple protonation of basic sites (e.g. the NH_2 group) in the amino acid sequences.

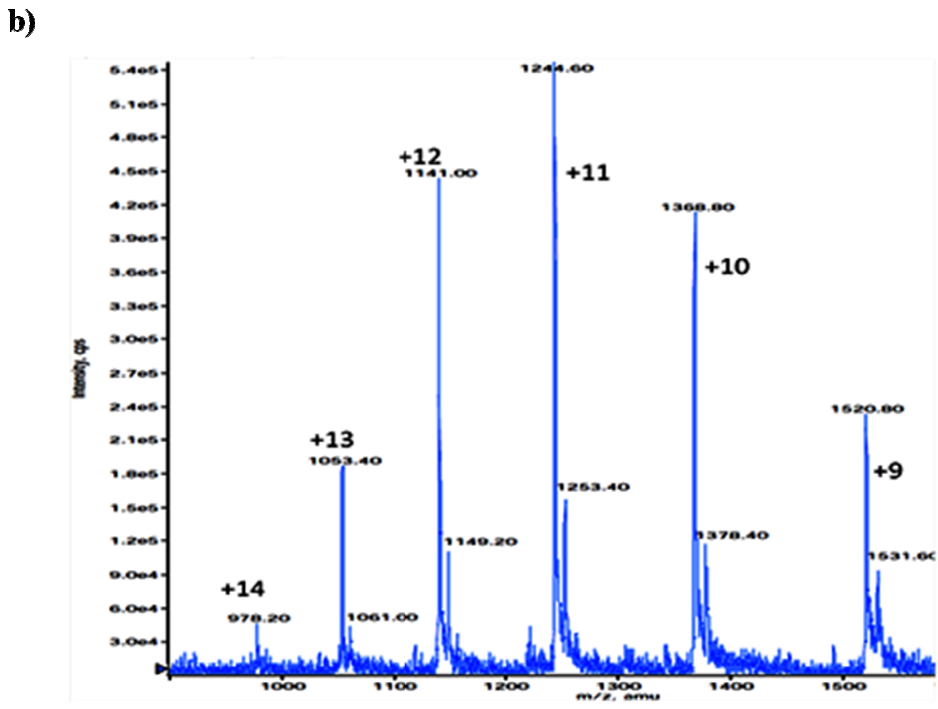
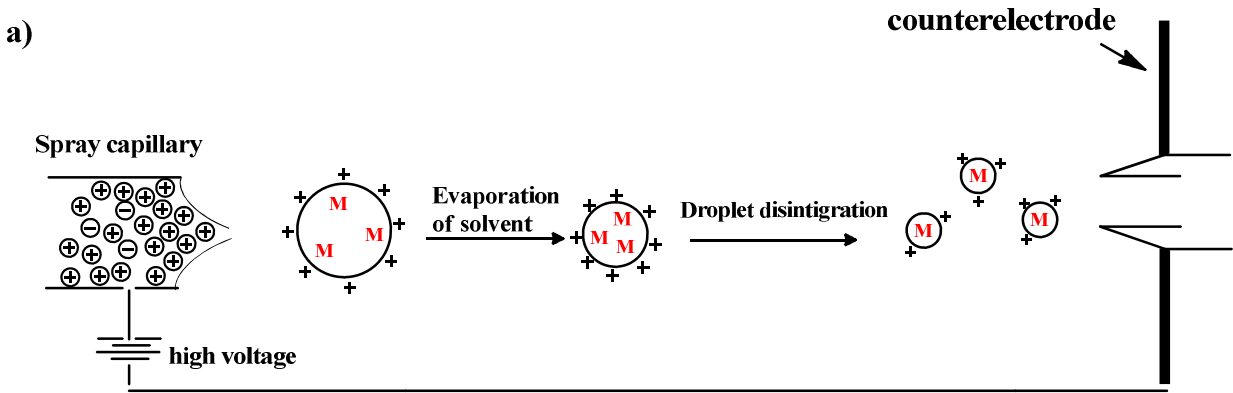


Figure 2.3: The ESI MS (a) ionization process (adapted from Ref. [50]), (b) spectrum of a ribonuclease from bison pancreas.

2.1.1.2 Mass analyzer

Mass analyzers are generally central to the technology and most mass spectrometers are named after their analyzers. The ions formed by MALDI or ESI are directed to the analyzer for filtering according to their m/z ratio. Generally, there are five types of mass analyzers based on the nature of m/z filtering [50]. These are magnetic-sector instruments (B, BE), time-of-flight (TOF), quadrupole ion trap (QIT), fourier transform ion cyclotron resonance (FTICR) and Orbitrap. Each analyzer has its own advantage over the other. The FTICR MS gives a high-resolution mass spectra and excellent mass accuracy measurements [51]. It also allows ions of all m/z to be detected simultaneously. The Orbitrap is the latest m/z analyzer with high resolving power and mass accuracy capabilities [52]. It has a comparable performance to FTICR without the need for cryogenic cooling. A TOF mass analyzer filters ions based on the difference in transit time of the ions through a field-free hollow tube under extremely high vacuum [53]. Lighter ions fly faster in the tube and thus hit the detector earlier than the heavy ones. A review of the methods used to improve the resolving power of TOF mass spectrometers is presented in ‘The new Time-of-Flight mass spectrometry’ [54]. Quadrupole ion trap (QIT) mass analyzers are composed of four parallel metal rods that have opposing time-varying AC voltages. For a given voltage, only ions of a certain m/z ratio are allowed to pass through the quadrupole filter. The sensitivity of Quadrupole mass spectrometers have been improved in the 3D QIT - manufactured jointly by Bruker Daltonics and Agilent Technologies [55].

2.1.1.3 Detector

The most commonly used detectors in MS are photomultiplier, microchannel plate and electron multiplier. After separation of the ions in the mass analyzer, the beam of ions generate a signal, which is detected electronically on a detector and stored in a data system together with their relative abundance for presentation in the format of an m/z spectrum. For instance, in Time-of-Flight mass spectrometers, the time taken by an ion to travel a certain distance in the field-free tube and strike a detector is used to calculate the m/z ratio seen in the mass spectrum.

2.1.2 Tandem mass spectrometry (MS/MS)

Multiple stage mass analysis of an analyte can be made in either more than one mass analyzer aligned in tandem (tandem-in-space MS) or in a single mass analyzer with each process separated in time (tandem-in-time MS) [56]. In typical tandem-in-space MS, the first analyzer is used to filter precursor ions accelerated from the ionization source. The second analyzer filters fragments of selected precursor ions according to their m/z ratio. Tandem-in-time MS analyzers have the capability to store ions and analyze them selectively at different times. In FT-MS, one example of tandem-in-time MS, ions within a given m/z range are selected and excited by applying a radiofrequency. Due to a combined effect of magnetic and electric fields, the excited ions undergo a cyclotron motion (oscillatory motion in Orbitrap) for detection as they pass near the detecting electrodes. Once this process is completed, ions can be stored for a later use or can be ejected from the analyzer cell [51].

The configuration of multiple mass analyzers in tandem can be performed in two ways. One way is to align the same types of mass analyzers one after the other as in Triple Quadrupole (QQQ) mass spectrometer. The other is to align different types of mass analyzers in tandem, thereby creating a hybrid mass spectrometer. Examples of commonly used mass analyzers and their MS/MS configurations are listed in Table 2.1.

The other feature of tandem mass spectrometers is that they allow dissociation of a precursor ion in a specialized sector. This is particularly useful for protein analysis, as they have to be fragmented into smaller peptides for sequencing. The three basic ion dissociation methods are collision-induced dissociation (CID), electron capture dissociation (ECD) and electron transfer dissociation (ETD). Each method is discussed in detail in Ref. [50]. The ions that are subjected to fragmentation are referred to as the precursor (parent) ions. The ions that result from the fragmentation of each precursor ion are known as the product (daughter) ions. The type of product ions formed depends on the level of energy and the type of fragmentation technique used. The CID is by far the leading technique used for peptide dissociation for sequence information [57]. Here, precursor peptide ions are filtered in the first analyzer and directed to a collision cell where they are dissociated by colliding with a neutral gas such as argon or helium.

Table 2.1 Types of mass analyzers and their MS/MS configurations.

Mass analyzers for MS	MS/MS configurations (most common ionization method)
<ul style="list-style-type: none">• Time-of-Flight (TOF)• FTICR• Quadrupole Ion Trap (QIT)• Ion Cyclotron Resonance (ICR)• Orbitrap	<ul style="list-style-type: none">• Triple Quadrupole - QQQ (ESI)• Quadrupole Ion Trap - ITMS (ESI)• Time-of-Flight/Time-of-Flight - TOF/TOF (MALDI) <p>Hybrid MS/MS</p> <ul style="list-style-type: none">• Quadrupole/Time-of-Flight - QTOF (ESI)• Linear Quadrupole Ion Trap/FTICR - LTQ-FTMS (ESI)• Linear Quadrupole Ion Trap/Orbitrap-LTQ-Orbitrap (ESI)

2.1.3 Liquid chromatography/mass spectrometry (LC/MS)

High performance liquid chromatography (HPLC) has been used for many years as the best separation technique for mixtures of compounds. The coupling of this powerful chromatography method to the high specificity and very low detection limit of mass spectrometers is ideally suited for the analysis of complex samples. An LC/MS system renders three major advantages. First, fractions of the complex sample separated in the LC part are directly introduced into the mass spectrometer and detected with a high degree of confidence [58]. Second, the volume of liquid solvents that is sprayed into the MS can be kept at a flow rate in the microliter and submicroliter per minute scale, thereby maintaining the ultrahigh vacuum system required for mass analysis [59]. Third, ESI-LC/MS allows the analysis of nonvolatile and thermally labile biomolecules (such as proteins and nucleic acids) in solution form, which would have otherwise been difficult to ionize by hard ionization methods.

Separation of protein or peptide mixtures in a reversed phase LC/MS system takes place in a capillary column packed with porous silica gel particles [60]. Attached to the silica gel, there is a

non-polar liquid phase called the stationary phase. The stationary phase in Vydac MS columns (manufactured by Grace Davison Discovery Science) can have C4, C8 or C18 carbon hydrophobic units attached chemically to the silica packing material. The C18 columns often with i.d. (internal diameter) up to 200 μm are being widely used in LC/MS-based protein analyses. The ultra small diameter of the columns helps to improve sensitivity by reducing excess dilution of the analyte by a mobile solvent. During the separation process the molecules of interest are partitioned between the stationary and the mobile phase based on their hydrophobic characteristics. Analyte molecules that have greater hydrophobic character make a strong interaction with the stationary phase and hence are slowly eluted, whereas molecules that have less hydrophobic character are eluted faster from the column and analyzed earlier in the MS.

The data obtained from ESI-LC/MS/MS analysis have three dimensions. All data together generate the LC-MS total ion chromatogram (TIC) which is the plotted total ion current against time (or scan number). The second and the third dimension are the MS spectrum of all the collected precursor ions and the MS/MS spectrum of the fragments of a selected precursor ion, respectively (see Figure 2.4).

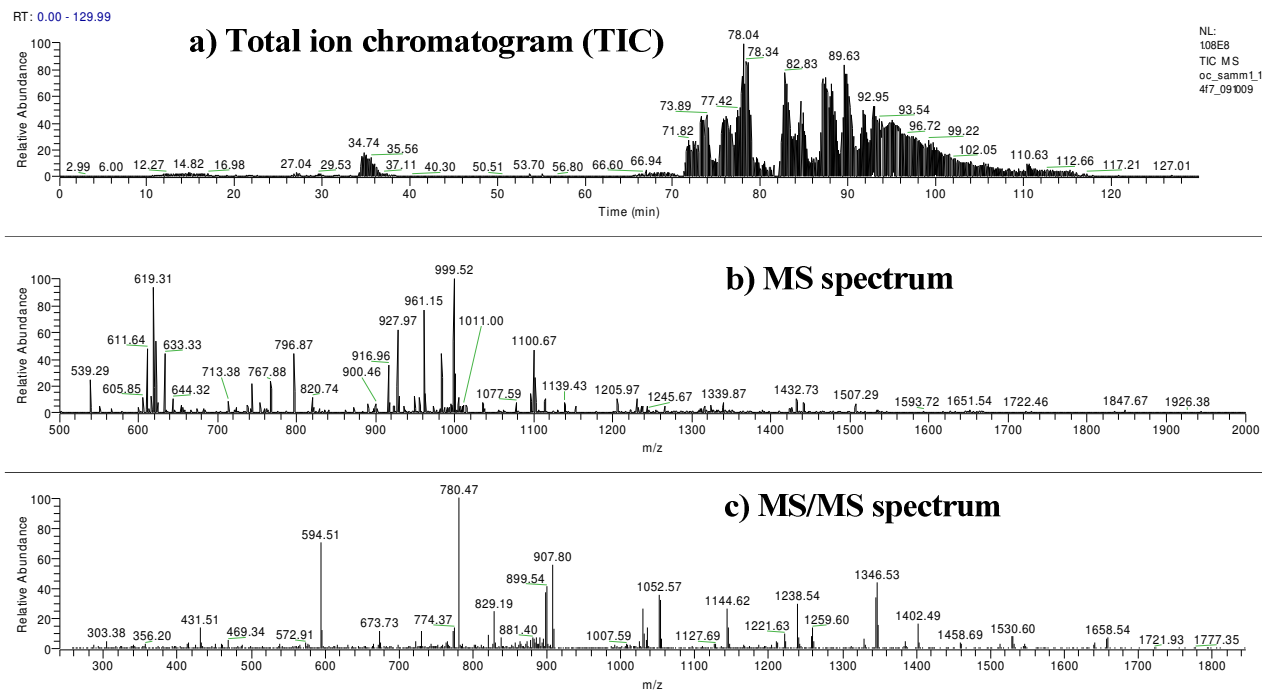


Figure 2.4: A representation of (a) TIC (b) MS spectrum of precursor ions (mass range 500 – 2000 m/z) and (c) MS/MS spectrum of the fragments of the precursor ion at m/z 916.96.

2.2 Peptide identification and characterization using MS

2.2.1 Peptide mass fingerprinting

Peptide mass fingerprinting (PMF) is one of the most common approaches used to determine the identity of peptides in a sample solution [61-62]. A typical PMF experiment involves three steps: Proteolytic digestion, mass measurement on a high-accuracy mass spectrometer and mass matching by performing a database search.

2.2.1.1 Proteolytic digestion

Enzymes that cleave peptide bonds in proteins are called proteases [63]. The two most commonly used serine proteases in proteolytic digestion are chymotrypsin and trypsin. Chymotrypsin from bovine pancreas selectively cleaves polypeptides at the C-terminal side of leucine, tryptophan, and tyrosine residues. A secondary cleavage may occur at the C-terminal side of phenylalanine, isoleucine and alanine. Trypsin cleaves polypeptides at the C-terminal side of arginine and lysine residues except when followed by a proline. Trypsin has the advantage of high proteolytic specificity and relatively lower auto-proteolysis (self-cleavage). As a result, it is widely used in PMF experiments. There are two major approaches for protein digestion [63]. The first approach is known as the in-gel digestion. Here, protein samples are run on polyacrylamide gel, then a Coomassie stained gel band, which corresponds to the protein of interest, is excised and cut into small pieces for enzymatic cleavage [64]. The second approach is known as the in-solution digestion. Here, the protease is directly added to the protein solution [65]. Disulfide bridges (S-S bonds) in globular proteins are usually resistant to proteolysis. As a result, this bonds are cleaved (for instance, by DTT) prior to adding the protease.

2.2.1.2 MALDI-TOF peptide mass mapping

Several types of mass spectrometers can be used for peptide mass measurement.

MALDI-TOF/TOF mass spectrometers such as the 4700 proteomics analyzer (from Applied Biosystems) and the FLEX series (from Bruker Daltonics) are well suited for this purpose. Both instruments have high sensitivity and provide fast measurement.

2.2.1.3 Database search

Peptide identification in PMF is performed by matching the experimental peptide masses obtained from MS analysis to the corresponding masses of *in silico*¹ generated peptides [66]. Prior knowledge of the protein sequence greatly simplifies the identification process by eliminating the need to look for possible matches in a database containing a large number of protein sequences. In general, the success of protein identification by PMF depends on the quality of the MS data, the power of the search algorithm used by the software tools and database accuracy. Some of the common search tools used in PMF are:

1. **Findpet tool** - ExPASy Proteomics Server (<http://au.expasy.org/tools/findpept.html>) [67]
2. **MS-Fit** - ProteinProspector (<http://prospector.ucsf.edu/prospector/mshome.htm>) [68] and
3. **Mascot** (http://www.matrixscience.com/search_form_select.html) [69]

Typical peptide mass matching using these software tools starts by setting workable search parameters on the user interface. As an example, the MS-Fit user interface from ProteinProspector is shown in Figure 2.5. The search begins by selecting a database (for instance - SWISS-PROT, NCBIInr or the protein sequence, if known) to be searched. Next, the enzyme used (e.g. trypsin, chymotrypsin etc.) is selected and the number of missed cleavages that may result from incomplete digestion are specified.

The variable and constant (if any) modifications that need to be considered in the search can be selected. Then, a desirable mass tolerance (in ppm or Da) is entered to determine the limit of mass error. Finally, a list of experimental masses obtained from MS analysis are pasted in the

¹ *in silico* digestion- a computer-based digestion, whereby a software tool is used to generate a list of peptides by performing a realistic cleavage of a protein in a database.

‘Data paste area’ and are matched by clicking on the ‘start search’ button. The result contains a list of peptides that matched with the experimental masses for each scored protein hits.

A

MS-Fit

Database: SwissProt.2009.07.07
 DNA Frame Translation: 3
 Taxonomy: All, HUMAN MOUSE, HUMAN RODENT
 Output: HTML, Hits to file, Name: lastres
 Digest: Trypsin, Max. Missed Cleavages: 1
 Constant Mods: Biotin (N-term), Carbamidomethyl (C), Carbamidomethyl (N-term)
 [+ Pre-Search Parameters]
 Start Search
 Sample ID (comment):
 Display Graph:
 Maximum Reported Hits: 5
 Sort By: Score Sort
 Min. # peptides required to match: 4
 Report MOWSE Scores: Pfactor: 0.4
 Masses are: monoisotopic
 Tol: 200 ppm, Sys Err: 0
 Contaminant Masses:
 Possible Modifications: Peptide N-terminal Gln to pyroGlu, Oxidation of M, Protein N-terminus Acetylated, Acrylamide Modified Cys
 User Def Mod 1: Acetyl (K)
 User Def Mod 2: Acetyl (K)
 User Def Mod 3: Acetyl (K)
 User Def Mod 4: Acetyl (K)
 OR
 Unknown Amino Acid: Single Base Change: Homology:
 Max Mods: 1, Min. # match with NO AA subs: 1
 Instrument: ESI-Q-TOF, Data Format: PP M/Z Charge
 Data Paste Area:
 915.2257141
 1013.296797
 1045.374821
 1070.313517
 1133.342687
 1389.492308

B

MS-Fit Search Results

[+] Parameters

[+] Pre Search Results

Fraction-Spot-Run ID: 1-1-1
 MS-Fit search selects 427 entries (results displayed for top 5 matches).

[-] Results Summary

Protein Hit Number	MOWSE Score	# pep # mat	% % Cov	% % TIC	Mean Data Err ppm	Data Tol ppm	# Hom Prot	MS-Digest Index	# Protein MW (Da)/pI	Accession #	Species	Protein Name
1	52571	4/4/20	56.5	20.0	-50.4	201	No	322484	13690/8.6	P61824	BISBI	Ribonuclease pancreatic
2	4317	7/6/30	14.5	30.0	-92.8	157	No	375128	94354/5.3	Q4Z014	PSEU2	Alanyl-tRNA synthetase
3	4133	5/5/25	13.3	25.0	-104	61.8	No	249936	87607/5.2	Q59934	STRMU	Formate acetyltransferase
4	2726	4/4/20	38.9	20.0	-99.6	118	No	322963	25521/5.4	Q4KHT1	PSEF5	Ribonuclease 3

322484 x x . x . . . x .
 375128 x x . . x x . . . x .
 249936 x . x . x x
 322963 x . x . . . x . x

[+] Detailed Results

1. 4/20 matches (20%).
 Acc. #: [P61824](#) Species: BISBI Name: Ribonuclease pancreatic
 Index: [322484](#) MW: 13690 Da pI: 8.6

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start End	Missed Cleavages	Sequence
915.2257	915.4029	100.0	-194		92 98	0	(K)YPNC(Carbamidomethyl)AYK(T)
2223.9981	2224.0856	100.0	-39.3		105 124	0	(K)HIIVAC(Carbamidomethyl)EGNPYVPVHFDAVS(-)
2364.9155	2364.9312	100.0	-6.63		11 31	0	(R)QHMSSTSAASSSNYC(Carbamidomethyl)NQMMK(S)

Figure 2.5: The MS-Fit (A) user interface, (B) representation of search results.

2.2.2 Tandem MS of peptides and database searching

2.2.2.1 Nomenclature of peptide fragment ions

The dissociation of a peptide precursor ion in an MS/MS collision cell generates two categories of fragment ions. Those that retain the charge on the N-terminal side of the peptide are termed the a-, b- and c- ions, whereas those that retain the charge on the C-terminal side are the x-, y- and z- ions.

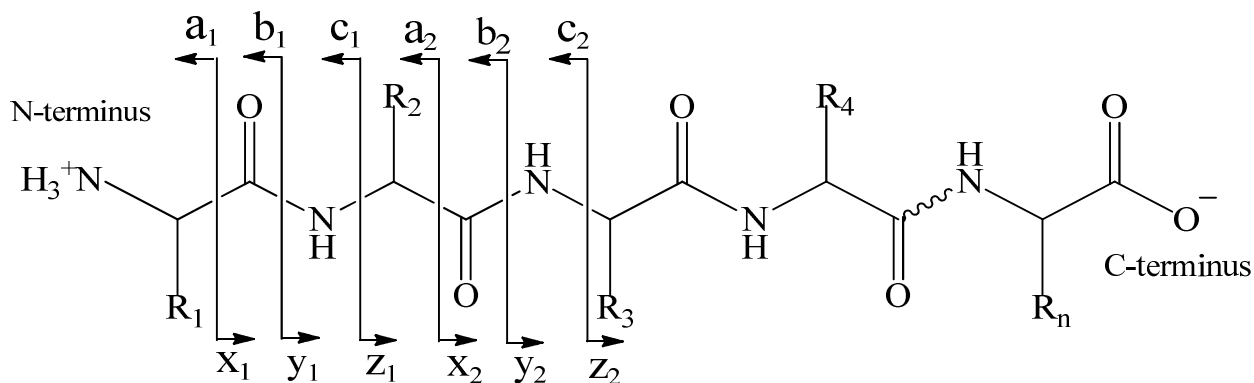


Figure 2.6: The Roepstorff nomenclature for peptide fragment ions [70].

Low energy CID of peptides generates primarily b- and y-ion fragments. Loss of the 'C=O' (≈ 28 Da) group from the b-ions results in the a-ions. The mass difference $b_n - b_{n-1}$ or $y_n - y_{n-1}$ (where, $n = 2, 3, 4, \dots$) is equal to a residue mass. The molecular mass of the original peptide is the sum of all the residue masses in the peptide plus H₂O (one -H at the N-terminus and one -OH at the C-terminus). Post-translation modifications at any site in the amino acid sequence are observed as an increase or decrease in the total peptide mass.

2.2.2.2 MS/MS Database search

Further analysis of the identity of peptides in a sample solution is carried out by sequencing the corresponding tandem mass spectra. This method of peptide identification has the advantage of characterizing post-translational modification sites at the residue level. However, assigning amino acid sequences to each MS/MS spectrum manually is an extremely slow and difficult process. Thus, in the same way as PMF, MS/MS-based peptide identification involves the use of software tools to automatically assign MS/MS spectra [71]. There are three categories of software tools that are used to match peptide amino acid sequences to MS/MS spectra. These are database-dependent search programs, *de novo* sequencing tools and hybrid search tools.

Database-dependent search programs: **Mascot** and **Sequest** are common examples of this category. For a given set of product ion spectra, these programs extract a list of scored candidate peptides from a sequence database. The search result is limited to user-entered parameters such as the protease used, allowed post-translational modifications, mass tolerance and missed cleavage. In general, the scores associated with candidate peptides entail how good the experimental and the calculated mass values are matched. Most frequently, the best scoring peptide matches are considered in the statistical validation of the result. A very good score, however, does not always mean that the identified peptide is a true match (although it is the best match). Thus, discriminating between the true and the false matches is one of the challenging tasks when protein identification is solely based on database search tools [72]. The challenge is even more pronounced when a test sample is multiply modified. Overall, the rate of false positive identification depends on the quality of the mass spectra, the discriminative power of the scoring algorithm and the size of the database.

***De novo* sequencing:** **PEAKS** is one example of software packages that are used for *de novo* sequencing of peptides. In cases where a protein is not contained in a database, sequencing can be performed using *de novo* methods. The product ion spectra obtained from MS/MS analysis are used for sequencing. Manual assignment of peptides to the MS/MS spectra is obviously a difficult and time consuming process. PEAKS makes sequencing much easier even without the

use of databases [73]. The program performs *de novo* sequencing in four basic steps: preprocessing, candidate computation, refined scoring and finally global and functional confidence scoring. This sequencing method, however, requires a high quality and complete fragment ion spectrum for successful identification.

Hybrid search tools: ByOnic, for instance, is a new hybrid software tool that combines the *de novo* sequencing and database search methods for peptide identification by MS/MS [74]. The program performs three major tasks during the search: First, it recalibrates the m/z measurements and computes lookup peaks². The lookup peaks are computed in pairs (m , m^1); the mass difference ($m^1 - m$) between the individual peaks is equal to a residue mass. Second, using either the peak pairs or individual lookup peaks, ByOnic searches candidate b- or y-ion theoretical peptide ions generated from a sequence database. Third, successfully matched values are scored. The scorer gives “benefits” for every theoretical peak found and “penalties” for peaks not found. The user interface in ByOnic, in the same way as the other search tools, has user-manageable search parameters. One advantage of ByOnic is that all possible amino acid modifications listed in Table 2.3 can be automatically considered in the search just by selecting the ‘ON’ option for the oxidation mode on the user interface. Moreover, a user can define the number of allowed modifications per peptide.

² Lookup peaks : A b- or y-ion peak identified by *de novo* method in the initial pass through the MS/MS spectra.

2.3 Characterization of protein-ligand interactions using oxidative footprinting and MS

In addition to X-ray crystallography and NMR, MS-based methods are also used for studying molecular interactions in protein-ligand complexes. The following are a few examples of such methods.

1. **SUPREX (stability of unpurified proteins by rates of H/D exchange)** – this MALDI-MS based method calculates the dissociation constant (K_d value) of a protein-ligand complex using a measured change in the thermodynamic stability of a protein upon ligand binding [75].
2. **PLIMSTEX (Protein-ligand interactions in solution by MS, titration and H/D exchange)** - this method examines the binding affinity (K_i) and stoichiometry, and protection against H/D exchange (ΔD_i) by plotting protein-ligand titration curves. Here, the MS is used to quantify changes that occur during H/D exchange [76].
3. **FPOP (fast photochemical oxidation of proteins)** is a hydroxyl radical mediated oxidative footprinting method coupled to MS that has recently been developed to examine structural features in protein-ligand complexes [77].

Several research groups have reported the use of the radical footprinting approach together with mass spectrometry in obtaining high-resolution information about binding interactions. Although it started initially for probing DNA-protein interfaces [78-80], over the years the method has been extended to the probing of other biomolecular complexes such as protein-protein and protein-carbohydrate [81-82].

The first step in the protein footprinting approach is the oxidation of the amino acid side chains by hydroxyl radicals. The rate of oxidation of each amino acid in the sequence is highly dependent on its side chain reactivity and solvent accessibility [83]. Generally, aromatic, aliphatic, sulphur-containing and charged amino acid side chains are highly susceptible to modification. In the MS-based footprinting approach, structural information is inferred by comparison of the level of modification of tryptic peptides in the free and ligand-bound protein. Furthermore, specific modification sites are identified by analyzing MS/MS spectra manually and by using automated software tools such as Mascot and/or ByOnic.

Quantitation of the oxidative footprinting data can be performed by two approaches. The first approach is to determine the percentage of oxidation at each amino acid by calculating the ratio of the LC-MS signal intensities of the peptides containing the specific oxidized amino acid to the total of all intensities (oxidized and non-oxidized) associated with that amino acid [84].

Mathematically,

$$\% AA_{oxid.} = \frac{\sum I_{oxid.}}{\sum I_{oxid.} + \sum I_{non-oxid.}} \times 100 \dots\dots \text{Equation 1}$$

Where, $\% AA_{oxid.}$ is the percentage oxidation of each amino acid, $\sum I_{oxid.}$ is the sum of the signal intensities of all the peptides containing the specific oxidized amino acid and $\sum I_{non-oxid.}$ is the sum of the signal intensities of all the non-oxidized peptides containing the amino acid of interest.

The second approach is to calculate the level of oxidation based upon a measure of the area under the SIC for each oxidized peptide over the sum of areas for the total peptide (oxidized and non-oxidized) across all charge states [82].

Mathematically,

$$\%oxid.peptide = \frac{A_{oxid.}}{A_{non-oxid.} + A_{oxid.}} \times 100 \dots\dots\dots \text{Equation 2}$$

Where, $A_{oxid.}$ is the peak area under the SIC of the oxidized form of the peptide and $A_{non-oxid.}$ is the peak area under the SIC of the non-oxidized form of the peptide.

Thus, a reduced level of oxidation in a peptide after ligand binding may suggest shielding of residue side chains from oxidation by hydroxyl radicals. In some cases, the reduction in the level of oxidation at a particular site may also infer a conformational change induced due to ligand binding.

2.3.1 Hydroxyl radicals

Hydroxyl radicals ($\cdot\text{OH}$) are highly reactive oxidants, which during reaction with amino acids abstract a hydrogen from side chain carbon atoms [81]. Thus, the radical leaves the side chain carbon with an unpaired electron. In aerobic conditions, the carbon with unpaired electron is susceptible to oxidation. The shift in peptide mass due to the modified amino acid can then easily be detected by mass spectrometry [85-86]. The advantages of using hydroxyl radicals in oxidative footprinting are that they can be generated in many different ways, can be produced abundantly on a short time scale and they induce irreversible side chain modifications in biomolecules [87-90]. Moreover, hydroxyl radicals are sufficiently small for resolving structural changes at the residue level.

2.3.2 Methods to generate hydroxyl radicals for protein footprinting

Laser photolysis of hydrogen peroxides is one method of generating enough hydroxyl radicals for protein footprinting experiment. Exposure of the peroxide to a high-energy laser source in aqueous solution produces two radicals (see Figure 2.7) [91-93]. UV laser sources that can generate multiple laser shots at microsecond [87] and nanosecond [94] timescales were used for footprinting studies. A single shot from such sources can create enough hydroxyl radicals to map the surface of a protein before it undergoes large conformational change or unfolding [84].

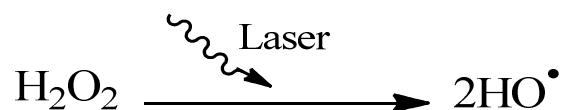


Figure 2.7: A representation of laser dissociation of a H_2O_2 to two hydroxyl radicals.

One of the challenges in oxidative footprinting experiments is that hydroxyl radicals might be produced in excess of what is required. This may lead to unwanted reactions in the protein solution. At increased doses, in addition to oxidizing the amino acid side chains, the chance for protein backbone cleavage to occur is high. Methods have been developed to address this

problem. In one approach, the protein solution is flash frozen in liquid nitrogen immediately after laser exposure to quench further action of the hydroxyl radicals [94]. Then the sample is freeze-dried to remove excess hydrogen peroxides and water molecules in the solution. In another approach, in addition to the flash freezing, chemical scavengers such as catalase and glutamine are used to mask the effect of excess hydroxyl radicals in the protein solution [77]. A recent footprinting report indicated that there is a high probability for ‘cold chemical oxidation of proteins’ to occur when freeze-drying oxidized protein solutions [95]. The report pointed out the importance of using chemical scavengers in reducing the effect of uncontrolled oxidation subsequent to laser exposure.

Fenton Chemistry is another method used to generate hydroxyl radicals from hydrogen peroxides. Here, the radicals are generated through the oxidation of Fe(II) to Fe(III) by hydrogen peroxides [96] (see Figure 2.8a). The method uses a chemical method of oxidation of amino acids using commonly available and affordable reagents. Inhibition of uncontrolled oxidation of amino acids side chains can be achieved by chelating of the iron (II) with EDTA. In the absence of EDTA, quenching or inhibition can be achieved by having excess unchelated Fe(II) in the sample solution (Figure 2.8b).

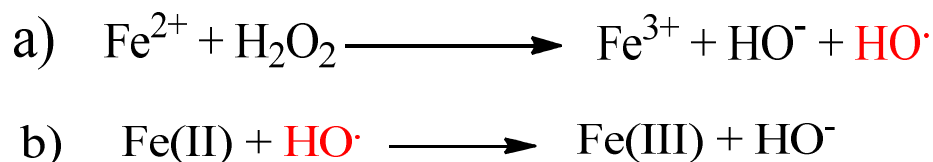


Figure 2.8: A representation of hydroxyl radical (a) generation (b) quenching in a sample solution [96].

2.3.3 Amino acid oxidation by hydroxyl radicals

Based on their side chain, amino acids can be classified into aliphatic, aromatic, basic, acidic, sulfur containing and proline.

All amino acids side chains have different sensitivity for hydroxyl radicals. Those with highly reactive side chains undergo a quick reaction with the radical and their modification is easily detected by mass spectrometry. On the contrary, amino acid residues with less reactive side chains take longer time to react with the radicals and hence are not good footprinting probes. The pseudo-first-order reactivity rate constant for each amino acid side chain is shown in Table 2.2.

Table 2.2- A Pseudo-first-order rate constant for the reaction of amino acid side chains with hydroxyl radicals [97-99].

Side chain	Rate ($M^{-1} s^{-1}$)	pH
Cys	3.4×10^{10}	7.0
Trp	1.3×10^{10}	6.5–8.5
Tyr	1.3×10^{10}	7.0
Met	8.3×10^9	6–7
Phe	6.9×10^9	7–8
His	4.8×10^9	7.5
Arg	3.5×10^9	6.5–7.5
cystine	2.1×10^9	6.5
Ile	1.8×10^9	6.6
Leu	1.7×10^9	~6
Val	8.5×10^8	6.9
Pro	6.5×10^8	6.8
Gln	5.4×10^8	6.0
Thr	5.1×10^8	6.6
Lys	3.5×10^8	6.6
Ser	3.2×10^8	~6
Glu	2.3×10^8	6.5
Ala	7.7×10^7	5.8
Asp	7.5×10^7	6.9
Asn	4.9×10^7	6.6
Gly	1.7×10^7	5.9

2.3.3.1 Oxidation of aliphatic amino acids

This group includes nonpolar amino acids such as glycine, alanine, valine, leucine, and isoleucine. In aerobic conditions, hydrogen abstraction by hydroxyl radicals from the hydrocarbon side chain of these residues results in addition of a hydroxyl group or formation of a keto group with detectable mass changes of +16 or +14 Da, respectively. Usually, the +16 Da oxidation product (OH addition) is more abundant than the +14 Da keto-product. Glycine is rarely oxidized by hydroxyl radicals. A reaction mechanism showing oxidation of aliphatic amino acid side chains by hydroxyl radicals is shown in Figure 2.9.

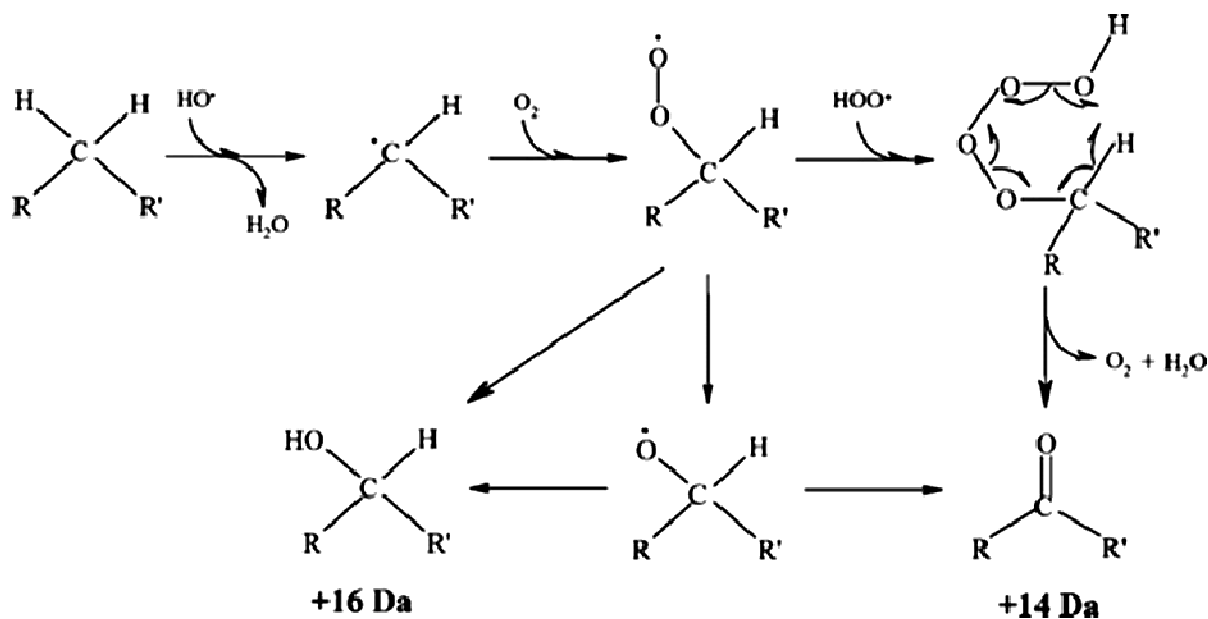


Figure 2.9: Oxidation of aliphatic amino acids by hydroxyl radicals [97].

2.3.3.2 Oxidation of aromatic amino acids

The aromatic amino acids phenylalanine, tyrosine, and tryptophan contain a phenyl ring in their side chain structure. The reaction of a hydroxyl radical with the side chain involves the addition of a hydroxyl group to the phenyl ring. One or multiple +16 Da mass shifts are a characteristic major product of this group of amino acids. A representative reaction mechanism of phenylalanine oxidation is shown in Figure 2.10 [100].

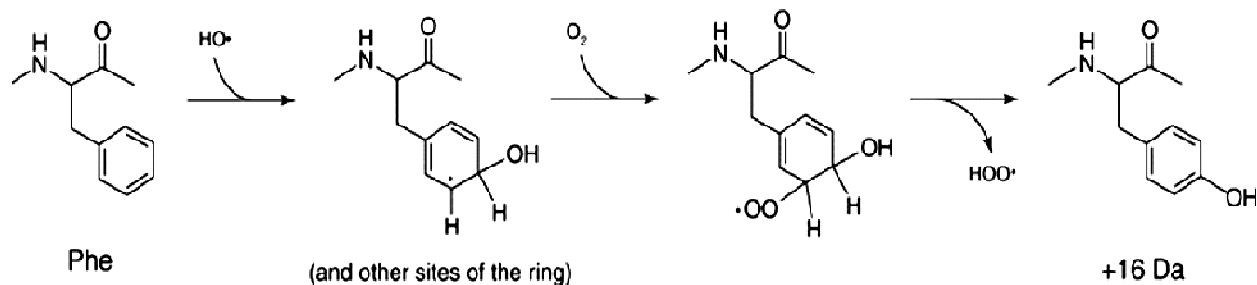


Figure 2.10: Oxidation of phenylalanine by hydroxyl radicals.

2.3.3.3 Oxidation of basic amino acids

Arginine and lysine have strongly basic side chains and histidine is a weakly basic amino acid [101]. At neutral pH, arginine and lysine are fully ionized. The arginine side chain in particular reacts readily with hydroxyl radicals (good footprinting probe) and loses a guanidino group upon oxidation, resulting in a mass loss of 43 Da (see Figure 2.11a). Lysine undergoes a reaction similar to aliphatic side chains. Histidine is unique in that it gives complex oxidation products. Attack by hydroxyl radicals of the side chain may cause ring opening. In mass spectrometry, a peptide containing modified histidine may give a mixture of products with +16, -22, -23, +5 or -10 Da mass shifts.

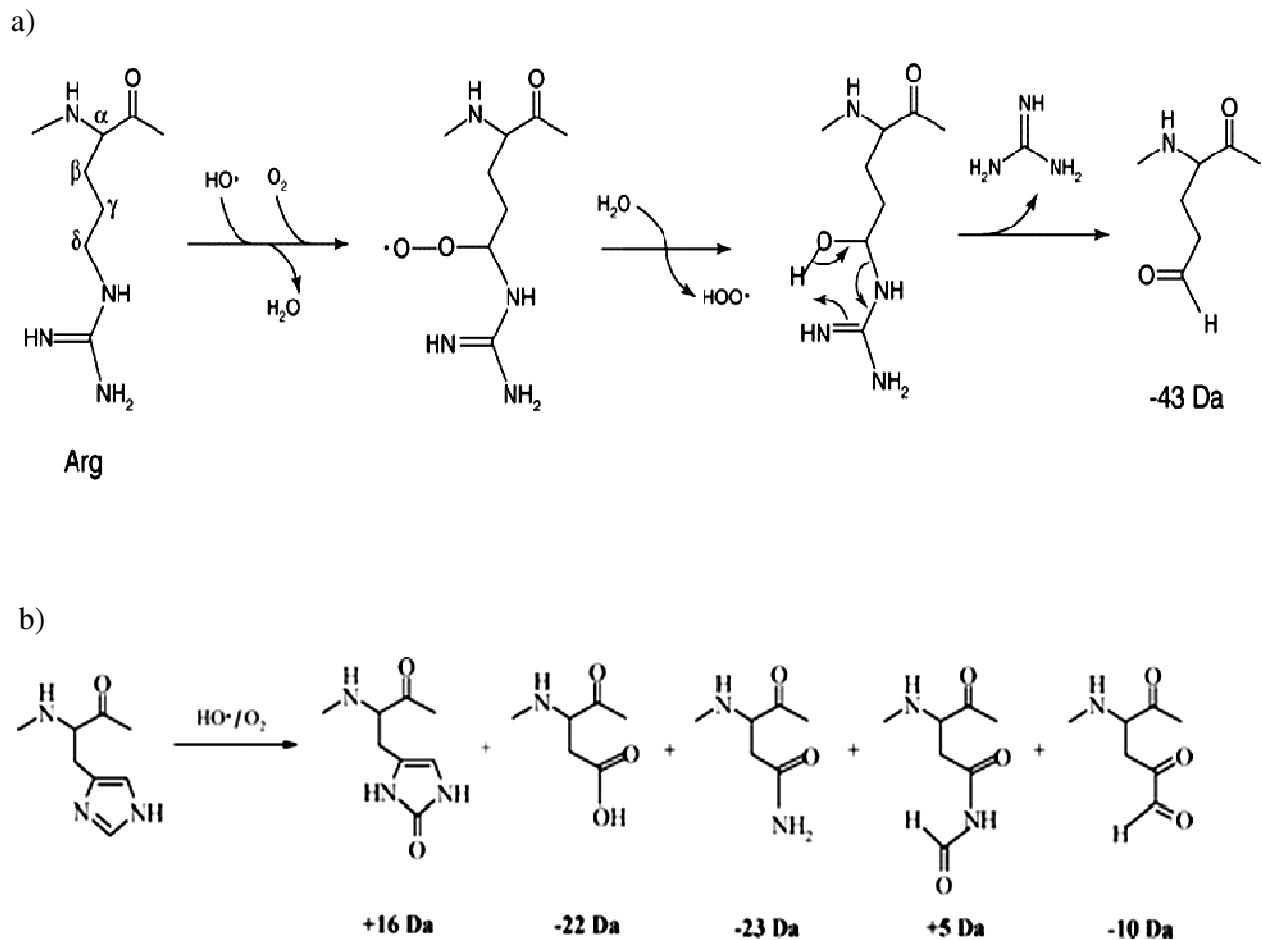


Figure 2.11: Oxidation of (a) arginine, (b) histidine by hydroxyl radicals [97].

2.3.3.4 Oxidation of acidic amino acids

Aspartate and glutamate belong to this group. Oxidative decarboxylation is a typical reaction of acidic amino acids in footprinting experiments. Thus, reaction of hydroxyl radicals with acidic residues is accompanied by decarboxylation and formation of an aldehyde group at the adjacent carbon atom. A characteristic mass loss of 30 Da is a common observation for these amino acids.

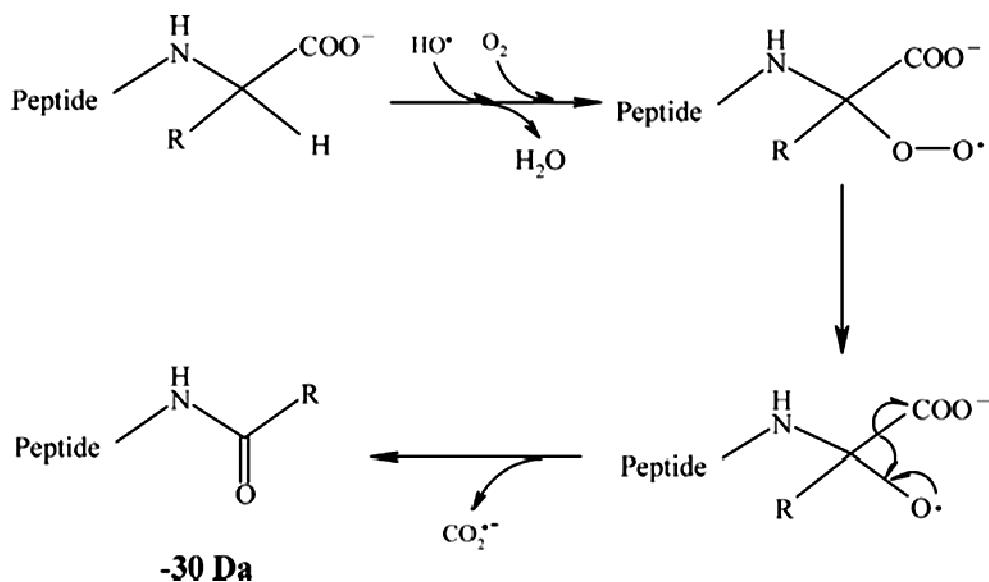


Figure 2.12: Oxidative decarboxylation of acidic amino acid side chains by hydroxyl radicals[97].

2.3.3.5 Oxidation of sulfur containing amino acids

Sulfur containing amino acids are methionine and cysteine. Their reaction with hydroxyl radicals may produce multiple oxidation products. Methionine side chain oxidation generates +16 Da sulfoxides and +32 Da sulfones as major and minor products, respectively. Further oxidation of methionine sulfoxide may result in a mass loss of 32 Da. Cysteine side chain oxidation may give products with +32 Da and/or +48 Da mass shifts.

effect (due to ligand binding) of the reactive and solvent accessible sites of the protein can easily be analyzed by mass spectrometry. Several oxidative footprinting results have confirmed the relative reactivity order of amino acids with hydroxyl radicals as Cys > Met > Trp > Tyr > Phe > His > Leu ~ Ile > Arg ~ Lys ~ Val > Ser ~ Thr ~ Pro > Gln ~ Glu > Asp ~ Asn > Ala > Gly. 14 of the 20 amino acids are good footprinting probes covering about 65% of the sequence of a typical protein [102-104]. Table 2.3 summarizes the possible mass shift due to radical oxidation for 19 amino acid residues. Glycine is too unreactive to be used as a probe in footprinting experiments.

Table 2.3 - Observed mass shifts due to side chain oxidation of residues by hydroxyl radicals [102-103, 105].

Residue	Side chain modification and mass change
Cys	sulfonic acid (+48), sulfuric acid (+32), hydroxy (-16)
Met	sulfoxide (+16), sulfide (+32), aldehyde (-32)
Trp	hydroxy- (+16, +32, +48, etc), pyrrol ring-open (+32)
Tyr	hydroxy- (+16, +32)
Phe	hydroxy- (+16, +32)
His	oxo-(+16), ring-open (-22, -10, +5)
Leu	Leu hydroxy- (+16), carbonyl (+14)
Ile	hydroxy- (+16), carbonyl (+14)
Val	hydroxy- (+16), carbonyl (+14)
Pro	hydroxy- (+16), carbonyl (+14)
Arg	deguanidination (-43), hydroxy- (+16), carbonyl (+14)
Lys	hydroxy- (+16), carbonyl (+14)
Glu	decarboxylation (-30), hydroxy- (+16), carbonyl (+14)
Gln	hydroxy- (+16), carbonyl (+14)
Asp	decarboxylation (-30), hydroxy- (+16)
Asn	hydroxy- (+16)
Ser	hydroxy- (+16), carbonyl (-2, or +16-H ₂ O)
Thr	hydroxy- (+16), carbonyl (-2, or +16-H ₂ O)
Ala	hydroxy- (+16)

3. EXPERIMENTAL

3.1 Materials and Reagents

Acetonitrile, 100 %, from J.T. Baker (U.S.A.).

Ammonium bicarbonate, 99 %, from Sigma-Aldrich (U.S.A.).

Blotting grade blocker non-fat dry milk from Bio-Rad Labs (U.S.A.).

Catalase from Sigma-Aldrich (U.S.A.).

Coomassie brilliant blue R-250 staining solution from Bio-Rad Labs (U.S.A.).

α -Cyano-4-hydroxy cinnamic acid from Fluka (U.S.A.).

DTT (Dithiothreitol), Cleland's reagent, from Bio-Rad Labs (U.S.A.).

ELISA 96-well plates from Costar (U.S.A.).

Formic acid, 88 %, from J.T. Baker (U.S.A.).

Goat anti-mouse IgG, alkaline phosphatase conjugated, from Sigma-Aldrich(U.S.A.).

H-Met-NH₂ – HCL from Bachem Americas Inc. (U.S.A.).

Hydrogen peroxide solution, 30 %, from Sigma-Aldrich (U.S.A.).

Laemmli sample buffer from Bio-Rad Labs (U.S.A.).

L-Glutamine, 99.5 %, from Fluka Biochemica (U.S.A.).

Methanol, 99.9 %, from Fisher Scientific (U.S.A.).

14F7 mAb from collaboration partners at CIM.

N-glycol GM3 ganglioside from collaboration partners at CIM.

Nitrocellulose membrane from GE Healthcare (U.S.A.).

pHydrion paper from Micro Essential Laboratory (New York, U.S.A.).

pNPP (*p*-Nitrophenyl Phosphate) tablets – source unknown

Precision Plus Protein standards from Bio-Rad Labs (U.S.A.).

Purified water (18.2 M Ω) from an in-house NANOpure DIamond system, Barnstead International (U.S.A.).

Sequencing grade modified trypsin from Promega (U.S.A.).

Sequencing grade chymotrypsin from Promega (USA).

Tween 20 – source unknown.

Tris base from EMD chemicals.

10xTris/Glycine/SDS buffer from Bio-Rad Labs (U.S.A.).

ZipTip C18 pipette tips from Millipore (U.S.A.).

3.2. Equipments

Agilent 1100 capillary liquid chromatography system from Palo Alto (U.S.A.).

ELISA microplate reader from BioTek.

ThermoFinnigan LTQ-FT MS, ESI source, from Thermo Electron Cor. (U.S.A.).

Model Compex 110 KrF excimer laser source, 248 nm UV wavelength, from Coherent Inc.,
(Santa Clara, U.S.A.).

4700 Proteomics analyzer, MALDI-TOF/TOF, from Applied Biosystems (U.S.A.).

192-well MALDI sample plate from Applied Biosystems (U.S.A.).

Voyager DE-Pro, MALDI-TOF, from Applied Biosystems (U.S.A.).

Vydac MS C18 300A HPLC column, 150 x 0.150-mm, from Grace Davison Discovery Sciences
(U.S.A.).

3.3 Methods

3.3.1 Sample preparation

3.3.1.1 SDS-PAGE

Gel electrophoresis was used to check sample purity for enzymatic digestion. 10 μ l of a 2.2 mg/ml antibody solution was mixed with 15 μ l of Laemmli sample buffer in a 1.5 ml Eppendorf tube (final conc. 0.88 μ g/ μ l). The mixture was then heated in a boiling water bath for 15 minutes. After loading the sample on a Bio-Rad ready gel, SDS-PAGE was run for 70 minutes at electrophoretic current and voltage of 150 mA and 100 V, respectively. The gel was stained using Coomassie brilliant blue for 1 hour. Then the staining solution was discarded and the gel was destained by washing with water in a container left on a shaker overnight.

3.3.1.2 ELISA

Briefly, each well of a polysorb 96-well assay plate was coated with NeuGc-GM3 ganglioside by adding 50 μ l of a 10 μ g/ml ganglioside solution in methanol. The plate was put in an incubator at 37°C for 2 h. After successive washes by rinsing with TBST to remove unbound gangliosides, the plate was blocked overnight (at room temperature) with 0.5 % non-fat dry milk blocking solution. On the next day, the blocking solution was discarded and 50 μ l of a dilution series of the antibody solution (in 0.05 M Tris-HCl buffer) ranging from 0 to 12 μ g/ml was added to the wells and incubated for 90 minutes at 37°C. Unbound antibodies were removed by washing each well with 100 μ l of TBST. Then, 100 μ l of alkaline phosphatase conjugated goat anti-mouse IgG diluted in 0.05 M Tris-HCl buffer (IgG: buffer ratio of 1:5000) was added to each well and incubated for 1 hr at 37°C. After the incubation, the plate was washed 6 times with TBST. For color development and detection, a substrate solution was prepared by dissolving two pNPP tablets in a 15 ml solution of 0.05 M Tris base and 0.15 M NaCl, pH 8.5. 200 μ l of the substrate solution was added to each well and incubated for 30 minutes. Then, 50 μ l of 1N NaOH was added to each well to stop the reaction. Finally, the absorbance value of the antibody-antigen complex solution and a blank solution was measured at 405 nm on an ELISA microplate reader.

3.3.2 Laser photolysis

Laser photolysis was conducted on the free and ligand-bound 14F7 mAb. Practically, immediately before laser exposure, 0.8 μl of hydrogen peroxide (5 % w/w, dilution of 30 % stock solution) and 1.25 μl of 200 mM DTT were added to 10 μl of a 2.2 mg/ml mAb solution in an Eppendorf tube. The final volume was brought to 30 μl by adding purified water (final conc. of the protein 0.73 $\mu\text{g}/\mu\text{l}$). The mixture was then passed at a flow rate of 4.4 $\mu\text{l}/\text{min}$ through a thin capillary tube exposed to a KrF excimer laser pulse at a wavelength of 248 nm. The oxidized sample was collected in new Eppendorf tube containing a mixture of catalase and H-Met-NH₂ – HCL in a 1:1 ratio to quench and mask the excess hydrogen peroxide and hydroxyl radicals left un-reacted in the solution. In addition, the oxidized sample was flash frozen in liquid nitrogen immediately after photolysis to prevent any un-controlled oxidation. Finally, the sample was lyophilized for 2 h to remove any remaining H₂O₂ and water. The same experimental procedure was followed in laser photolysis of the antibody-ganglioside complex except that the antibody was incubated with the ganglioside on ice for 2 h before irradiation.

3.3.3 Proteolytic digestion

In-solution tryptic digestion of the antibody was performed. To prepare the sample for tryptic digestion, the lyophilized sample was resuspended in 15 μl of 50 mM ammonium bicarbonate digest buffer (for non-oxidized samples, in addition to the digest buffer, DTT was added to a final DTT concentration of 10 mM). The mixture was heated for 90 minutes in a water bath at 67°C. Then, 2.2 μl of a sequencing grade modified trypsin was added. This amount of trypsin gives a protease to protein ratio of 1:50. The pH of the mixture was adjusted to be between 6 and 8 using a pHydration paper. Finally, the sample was incubated for 24 h at 37°C.

3.3.4 Peptide mass fingerprinting

Peptide mass fingerprinting was performed on a 4700 MALDI-TOF/TOF proteomics analyzer (See Figure 3.1). The CID was turned off for this purpose. First, the matrix α -cyano-4-hydroxy cinnamic acid (10 mg/ml) was prepared by dissolving 10 mg of α -cyano powder in a mixture of 0.5 ml acetonitrile and 0.5 ml ethanol. On a MALDI plate, 0.6 μ l of the tryptic peptide solution obtained from the proteolytic digestion and 0.8 μ l of the matrix were spotted and allowed to dry at room temperature before mass mapping. Then the plate was inserted into the 4700 MS. Per sample spot, 100 laser shots with fixed laser intensity of 5000 Hz were taken for MS acquisition in positive mode. Mass spectra were acquired for peptides with m/z values between 500 and 3000.

A)



B)



Figure 3.1: AB 4700 MALDI-TOF/TOF proteomics analyzer (A) and AB 192-well MALDI plate (B).

3.3.5 LC-MS/MS analysis

Reverse phase (C18) liquid chromatographic MS/MS analyses of proteolytic peptides were performed using a HPLC system consisting of an Agilent 1100 HPLC binary pump with corresponding autosampler. This LC system was coupled via a nanoelectrospray ion source to a LTQ-FTICR mass spectrometer (Thermo Fisher Scientific). For the analyses, 1-2 μl of peptide solution was injected onto the 150 x 0.150-mm C18, 5 μm resin, column. The mobile phase consisted of acetonitrile and MS grade water, both containing 0.1 % formic acid. Chromatographic separation was achieved using a binary gradient from 5 to 95 % of acetonitrile in 120 min (details in Table 3.1). The flow rate was 0.8 $\mu\text{l}/\text{min}$. Mass spectra were acquired in the positive ion mode applying a data-dependent automatic switch between survey scan and tandem mass spectra (MS/MS) acquisition. Peptide samples were analyzed by collision induced dissociation (CID) in the LTQ ion trap (low resolution) by acquiring one FTICR survey scan in the mass range of m/z 500 – 2000. The raw data were acquired in centroid format on Thermo scientific XCalibur 2.0 software. MS/MS spectra were manually inspected by Qual Browser version 2.0.7.

Table 3.1: Elution gradient

Time (min)	Flow Rate ($\mu\text{l}/\text{min}$)	Composition (in %)
0.00	2.00	A = 95.0, B = 5.0
2.00	2.00	A = 95.0, B = 5.0
3.00	0.80	A = 95.0, B = 5.0
40.00	0.80	A = 40.0, B = 60.0
55.00	0.80	A = 20.0, B = 80.0
60.00	0.80	A = 95.0, B = 5.0
130.00	0.80	A = 95.0, B = 5.0



Figure 3.2: LTQ-FT MS interfaced with Agilent 1100 capillary liquid chromatography (Ron Orlando's Lab., CCRC, University of Georgia).

3.3.6 Database searching and spectral assignment

The experimental masses obtained from the MALDI-MS analysis of the oxidized (O-14F7) and non-oxidized (N-14F7) samples were searched against a database containing the 14F7 Fab amino acid sequences. ExPASy proteomics server has software tools that can be used to identify matching peptide sequences and potential post-translational modifications (PTM) for a given set of experimental masses. For the N-14F7 sample, the software tool called 'Findpept' was used for identification of matching peptide sequences. Methionine oxidation was selected as a variable modification. The mass tolerance was set to ± 0.5 Da and maximum two missed cleavages were allowed. For the oxidized 14F7 sample, the software tool called 'FindMod' was used for the identification of matching peptide sequences and modification sites. The mass tolerance was set to ± 0.5 Da, maximum two missed cleavages and three modifications per peptide were allowed.

The identification of peptide matches for the MS/MS spectra was performed *via* a database searching on ByOnic. Before performing the search the raw file tandem mass spectra was converted into an mzXML file (a compressed file format) to quicken the identification. Based on the selection made from a list of MS analyzers (ion trap in this case), the program automatically assigned a parent ion mass tolerance of 0.4 Da and fragment ion mass tolerance of 2.3 mmu. The program also automatically identified the precursor ion charges. The search was performed against a database containing the 14F7 Fab sequence (forward and reverse). High specificity for tryptic peptides and medium filtering for semi-tryptic peptides were allowed. For the O-14F7 sample, all possible amino acid PTMs listed in Table 2.3 were considered in the search. The minimum score set for a good peptide match in this identification was 400.

3.3.7 Docking simulations of the 14F7 - NeuGc-GM3 complex

The program AutoDock 3.05 [106-108] was used for predicting the complex structure of the 14F7 Fab with the trisaccharide moiety of the N-glycolyl GM3, NeuGc(α 2-3)Gal(β 1-4)Glc. The structure of the trisaccharide was built on the GLYCAM website, <http://www.glycam.com/>. The docking program predicts how a ligand binds to a known 3D structure by employing different search methods, one of which applies a Lamarckian genetic algorithm. Using this algorithm, 100 docking runs were performed using a grid map of $81 \text{ \AA}^3 \times 81 \text{ \AA}^3 \times 101 \text{ \AA}^3$ (points in the X, Y and Z directions). The grid map was centered on the complementarity determining regions of the 14F7 crystal structure. The phi (Φ) torsion angle (C2-O3-C3) was -60° . The glycosidic bond around C2-O3 was freely rotatable. The ceramide group of the N-glycolyl GM3 ganglioside was replaced by a methyl ($-\text{CH}_3$) group.

Overview of a peptide mass fingerprinting experiment

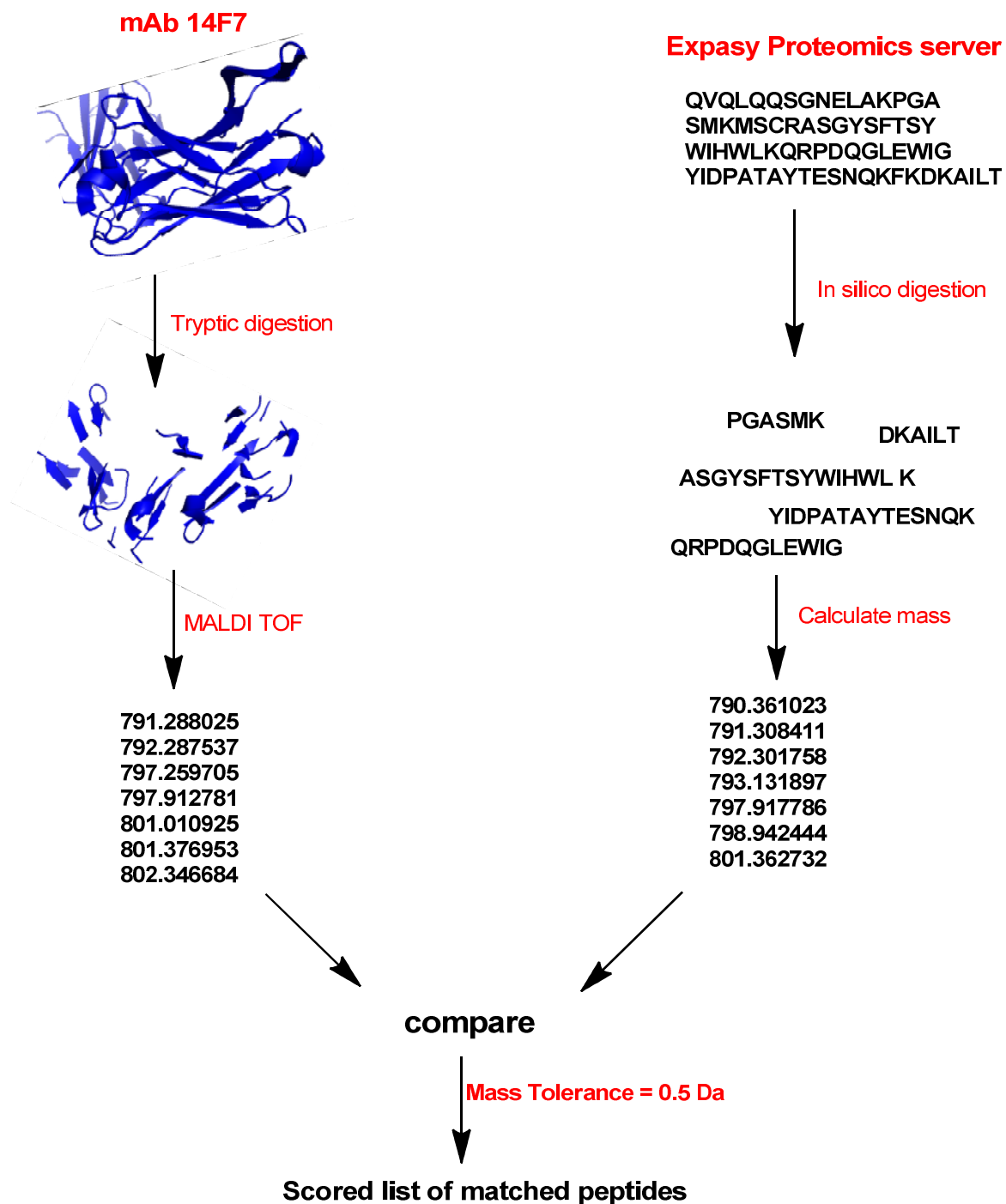


Figure 3.3: A schematic diagram showing a summary of the peptide mass fingerprinting experiment performed on the 14F7 mAb. Picture made with ChemBioDraw Ultra 12.0.

Overview of oxidative footprinting and MS/MS experiments

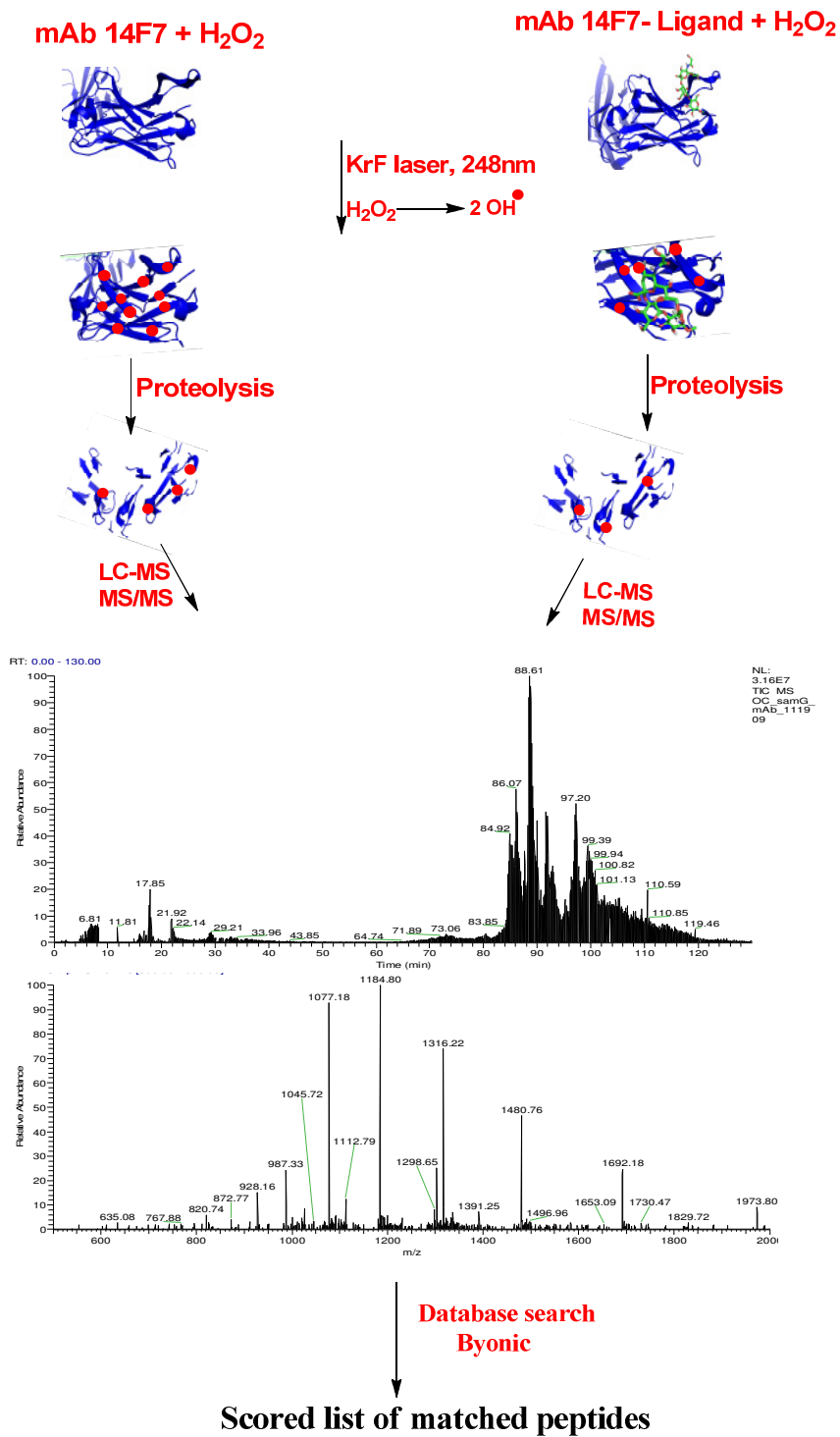


Figure 3.4: A schematic diagram showing a summary of the oxidative footprinting experiment performed on the free and ligand-bound 14F7 samples. Picture made with ChemBioDraw Ultra 12.0.

4. RESULTS AND DISCUSSION

4.1 Sample preparation

4.1.1 SDS-PAGE

The purity of the 14F7 mAb solution was confirmed with SDS-PAGE (see Figure 4.1). The single band in well-2 represents the intact mAb solution with an approximate molecular mass of 150 kDa. Well-3 is a representative of the 14F7 Fab solution. In addition to the Fab fragment (\approx 50 kDa), other unknown fragments (for instance, light bands at 75 kDa and 100 kDa) were observed. Therefore, only the intact mAb solution was used in all experimental set-ups in this project.

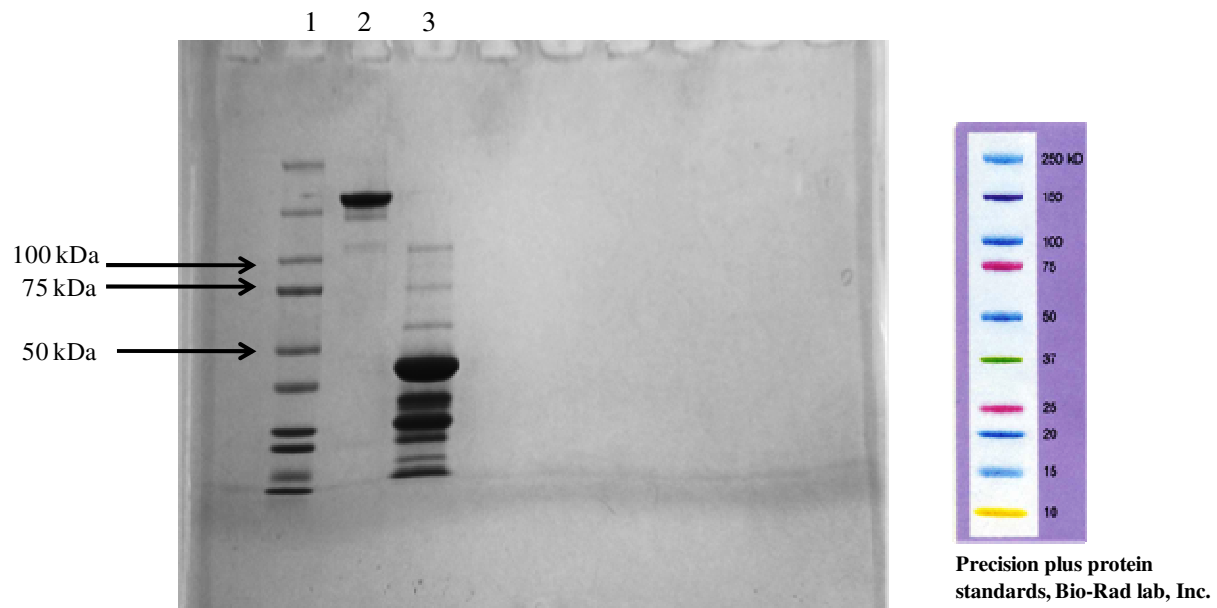


Figure 4.1: An SDS-PAGE of 14F7. Well-1- molecular weight marker, well-2 – 14F7 mAb, Well-3 – 14F7 Fab. In well 2, the band at \approx 150 kDa represents the 14F7 mAb. In well 3, in addition to the 14F7 Fab at \approx 50 kDa, two light bands at 75 and 100 kDa were observed, suggesting the impurity of the Fab solution.

4.1.2 ELISA

The background absorbance of the plate was subtracted from the absorbance of the antibody-antigen complex in each well to get the corrected sample reading. The corrected absorbance value is proportional to the amount of antibody-antigen complex present in each well [109]. The relationship between the absorbance and the concentration of the 14F7 mAb is presented in Figure 4.2. The line graph deviates slightly ($R^2 = 0.9614$) from the expected straight line. This could be due to excessive washing and/or trace amounts of unbound antibodies left unreacted in the wells. The graph shows that the absorbance at 405 nm of the complex solution increases with increasing concentration of the antibody. This result agrees with the ELISA binding assay performed on the same antibody-antigen complex by Carr *et al.* [30]. The conclusion is that the 14F7 mAb is active and recognizes the NeuGc-GM3 ganglioside.

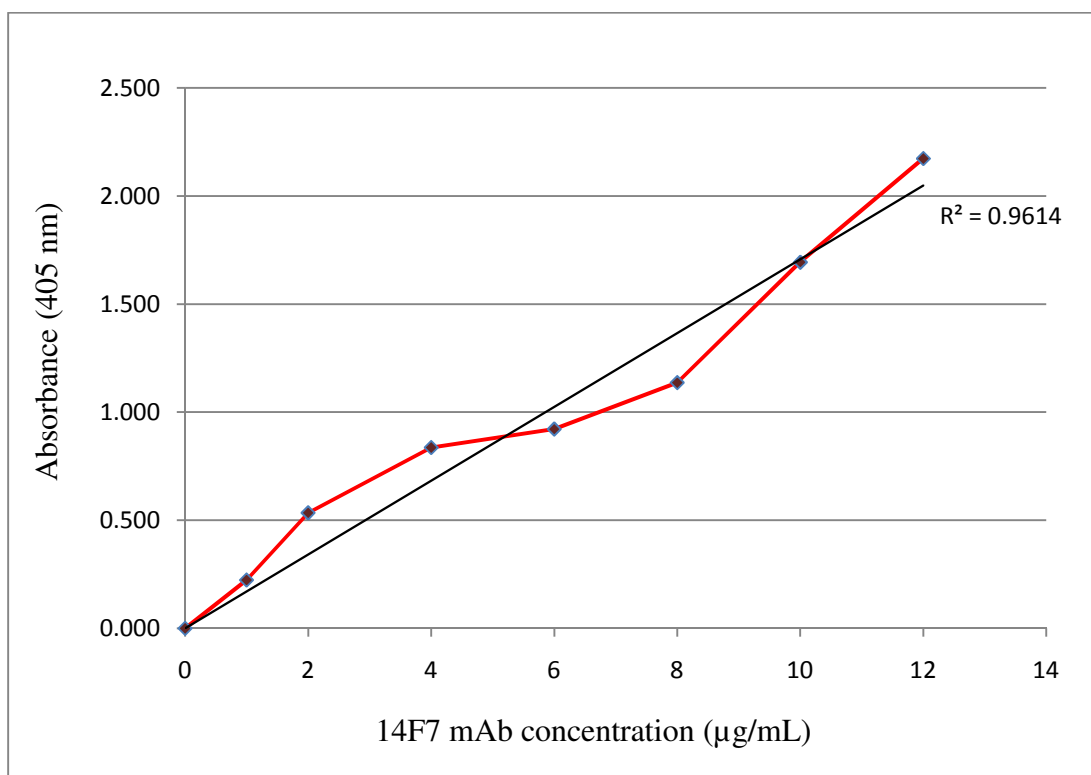


Figure 4.2: Binding assay of 14F7 mAb against NeuGc-GM3 ganglioside by ELISA. The absorbance value increased proportionally with increasing concentration of the antibody solution, suggesting that the ganglioside binds to the antibody.

4.2 MALDI-MS analysis of N-14F7 and O-14F7 mAb digests

4.2.1 Peptide mass matching using the ExPASy database search

The identity of the experimental masses obtained from the MALDI analyses was determined by performing a database search against the 14F7 Fab amino acid sequences. Table 4.1 shows a list of identified peptide matches for the N-14F7 mAb digest.

For the O-14F7, as compared to the non-oxidized form, the number of identified peptide matches is fewer (see Table 4.2). Moreover, the database search found six oxidized tryptophan (TPO) residues – three TPOs at position 33 (VH-CDR1) and one TPO at positions 266 (VL-CDR1), 379 and 394, the latter not being part of the CDRs.

Table 4.1: A list of peptide matches identified by ExPASy for the N-14F7 digest. The CDR peptides are highlighted by red boxes. The arrows indicate the three VH-CDR3 peptides identified by the database searching.

Matching peptides for specific cleavage:						
User mass	DB mass	Δ mass (daltons)	peptide	position	modifications	missed cleavages
523.000	523.262	0.262	(K)/SFNR/(N)	439-442		0
523.000	523.262	0.262	(K)/SFNR/(N)	439-442		0
523.000	523.262	0.262	(K)/SFNR/(N)	439-442		0
537.000	537.303	0.303	(K)/FKDK/(A)	64-67		1
588.000	588.335	0.335	(K)/DINVK/(W)	374-378		0
588.000	588.335	0.335	(K)/DINVK/(W)	374-378		0
612.000	612.419	0.419	(K)/KIVPR/(D)	221-225		1
612.000	612.419	0.419	(K)/KIVPR/(D)	221-225		1
658.000	658.425	0.424	(K)/LELKR/(A)	335-339		1
658.000	658.425	0.424	(K)/LELKR/(A)	335-339		1
676.000	676.326	0.326	(K)/IDGSER/(Q)	381-386		0
711.000	711.294	0.294	(K)/DEYER/(H)	415-419		0
726.000	726.353	0.352	(R)/THESPR/(L)	271-276		0
759.000	759.436	0.435	(K)/AILTADR/(S)	68-74		0
832.000	832.477	0.477	(K)/TSTSPIVK/(S)	431-438		0
869.000	869.357	0.357	(K)/SFNRNEC	439-445		1
991.000	990.500	-0.499	(K)/WKIDGSER/(Q)	379-386		1
1003.000	1002.558	-0.442	(K)/DKAILTADR/(S)	66-74		1
1078.000	1077.573	-0.427	(R)/WPLTFGAGTK/(L)	325-334		0
1078.000	1077.573	-0.427	(R)/WPLTFGAGTK/(L)	325-334		0
1266.000	1265.648	-0.351	(K)/YASQSISGIPSR/(F)	281-292		0
1291.000	1290.553	-0.446	(R)/HNSYTCEATHK/(T)	420-430		0
1560.000	1559.818	-0.182	(K)/DINVKWKIDGSER/(Q)	374-386		2
1592.000	1591.735	-0.265	(R)/QNGVLNSWTDQDSK/(D)	387-400		0
1832.000	1831.883	-0.116	(R)/ASQSI>NNLHWYQQR/(T)	256-270		0
1846.000	1845.896	-0.104	(R)/ASGYSFTSYWIHWLK/(Q)	24-38		0
2227.000	2227.007	0.007	(K)/DSTYSMSSLTLTKDEYER/(H)	401-419		1
→ 2461.000	2461.138	0.138	(R)/GIYYAMDYWGQGTTVTVSS AK/(T)	106-127		0
→ 2880.000	2880.380	0.379	(K)/QRPDQGLEWIGYIDPATAYT ESNQK/(F)	39-63		0
→ 2886.000	2886.424	0.424	(R)/LRRGIYYAMDYWGQGTTVT VSSAK/(T)	103-127		2
→ 2902.000	2902.419	0.419	(R)/LRRGIYYAMDYWGQGTTVT VSSAK/(T)	103-127	MSO	2

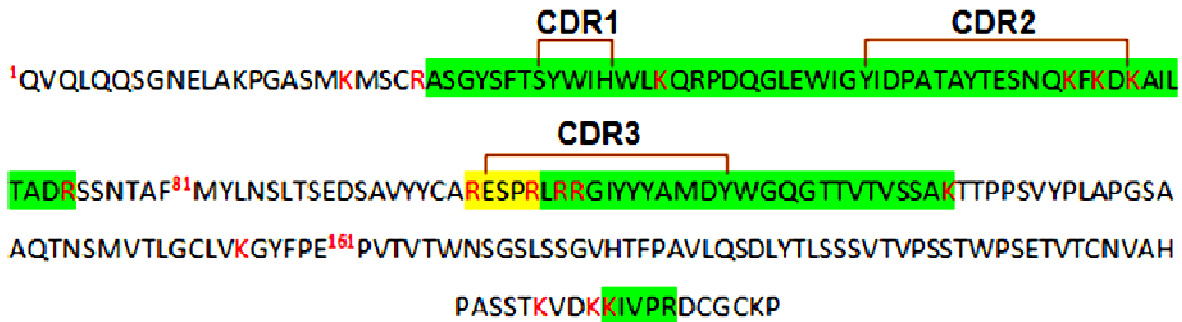
Table 4.2: A lists of peptide matches identified by ExPASy for the O-14F7 digest. The identified CDR peptides are highlighted by red boxes. Notice the oxidation of tryptophan residue (TPO) by hydroxyl radicals.

Matching peptides:						
User mass	DB mass	Δ mass (Dalton)	#MC	peptide	position	known modifications
523.000	523.2623	0.262	0	SFNR	439-442	
537.000	537.3031	0.303	1	FKDK	64-67	
588.000	588.3351	0.335	0	DINVK	374-378	
612.000	612.4191	0.419	1	KIVPR	221-225	
658.000	658.4246	0.425	1	LELKR	335-339	
676.000	676.3260	0.326	0	IDGSER	381-386	
711.000	711.2944	0.294	0	DEYER	415-419	
726.000	726.3529	0.353	0	THESPR	271-276	
757.000	757.4315	0.432	1	ESPRLR	99-104	
759.000	759.4359	0.436	0	AILTADR	68-74	
832.000	832.4774	0.477	0	TSTSPIVK	431-438	
914.000	913.5326	-0.466	2	ESPRLRR	99-105	
1006.000	1006.4952	0.495	1	WKIDGSER	379-386	(TPO: 379)
1078.000	1077.5727	-0.426	0	WPLTFGAGTK	325-334	
1266.000	1265.6484	-0.351	0	YASQSISGIPSR	281-292	
1291.000	1290.5531	-0.446	0	HNSYTCEATHK	420-430	
1592.000	1591.7347	-0.264	0	QNGVLNSWTDQDSK	387-400	
1608.000	1607.7296	-0.269	0	QNGVLNSWTDQDSK	387-400	(TPO: 394)
1832.000	1831.8834	-0.116	0	ASQSISNNLHWYQQR	256-270	
1846.000	1845.8958	-0.103	0	ASGYSFTSYWIHWLK	24-38	
1846.000	1845.8958	-0.103	0	ASGYSFTSYWIHWLK	24-38	(1xTPO)
1846.000	1845.8958	-0.103	0	ASGYSFTSYWIHWLK	24-38	(2xTPO)
1848.000	1847.8783	-0.121	0	ASQSISNNLHWYQQR	256-270	(TPO: 266)
1983.000	1982.8297	-0.169	1	DEYERHNSYTCEATHK	415-430	
2014.000	2014.0386	0.039	0	QVQLQQSGNELAKPGASK	1-19	
2227.000	2227.0071	0.007	1	DSTYSMSSTLTLTKDEYER	401-419	

4.2.2 Percentage sequence coverage of 14F7 digest

The percentage sequence coverage is calculated by dividing the number of amino acids observed by the 14F7 Fab amino acid length. The observed amino acids are highlighted in green in Figure 4.3. The fragments cover about 44.5 % of the Fab amino acid sequences. This is equivalent to 70 % of the expected coverage (64.5 %) that can be achieved by a theoretical tryptic digest using the ExPASy 'Peptidmass' tool within an m/z range of 500 to 3000 Da and with a ± 0.5 Da mass error. Four residues within both the VH-CDR3 and VL-CDR3 regions were not covered by any of the fragments (highlighted in yellow in Figure 4.3). Several other proteases such as chymotrypsin and V8 were tested in order to optimize the sequence coverage, however, without any improvement (see Appendix 3, Figure A1 for MALDI spectrum of the chymotryptic digest).

VH domain



VL domain

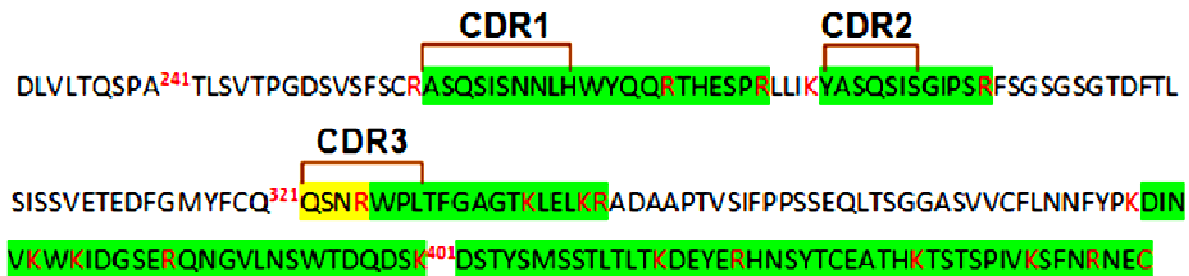


Figure 4.3: Peptides identified by ExPASy in the VH and VL regions of the 14F7 mAb digest. The observed residues are highlighted in green. The residues labeled in red are potential tryptic cleavage sites.

4.2.3 MALDI-TOF spectra - why are some peptides less abundant?

Figure 4.4 displays a representative of the high-resolution MALDI spectrum for the N-14F7 mAb digest. All the peaks corresponding to the CDR peptides are assigned. The abundance of the light chain CDR peptides is higher than that of the heavy chain CDR. There could be several reasons for the lower abundance of the latter peptides. First, it could be that the enzymatic cleavage in this region of the 14F7 mAb was too specific and did not generate many fragments. Meanwhile, an attempt made to digest the 14F7 mAb using relatively less specific proteases such as chymotrypsin and V8 did not improve the abundance of the heavy chain CDRs (see appendix 3, Figure A1 for the mass spectrum of the chymotryptic digest).

A second possible explanation could be that the intact antibody may not have unfolded completely before adding the enzyme. To improve this, several parameters of the digestion protocol including the concentrations of the reducing agent and digest buffer and the denaturation temperature were varied. The third and the most likely reason could be the stability of the peptides within this region. Docking and molecular dynamics simulations performed by Krenzel *et al.* [24] suggest that the VH-CDR3 is quite stable. It was proposed that the three consecutive tyrosine residues within the VH-CDR3 amino acid sequence might interact with neighboring aromatic residues, thereby stabilizing the region. A fourth possible explanation for the low abundance of the heavy chain CDR peptides is that gas-phase basicity of the peptides might be very low, thereby affecting their ionization efficiency in MALDI MS. Obviously basic peptides are more likely to absorb a proton and ionize. Therefore, attaching a quaternary amine to the peptides may enhance ionization by providing a permanent positive charge[110].

To check for any differences in the relative abundance of the VH-CDR peptides when using the Fab and intact 14F7 mAb, separate digests were performed for analysis on the 4700 MALDI MS. A representative of the mass spectrum for the Fab digest is shown in Appendix 3, Figure A2. The Fab MALDI spectrum is similar to that of the intact mAb. This suggests that, as expected, the abundance of the VH CDR peptides is independent from if we use the Fab or mAb.

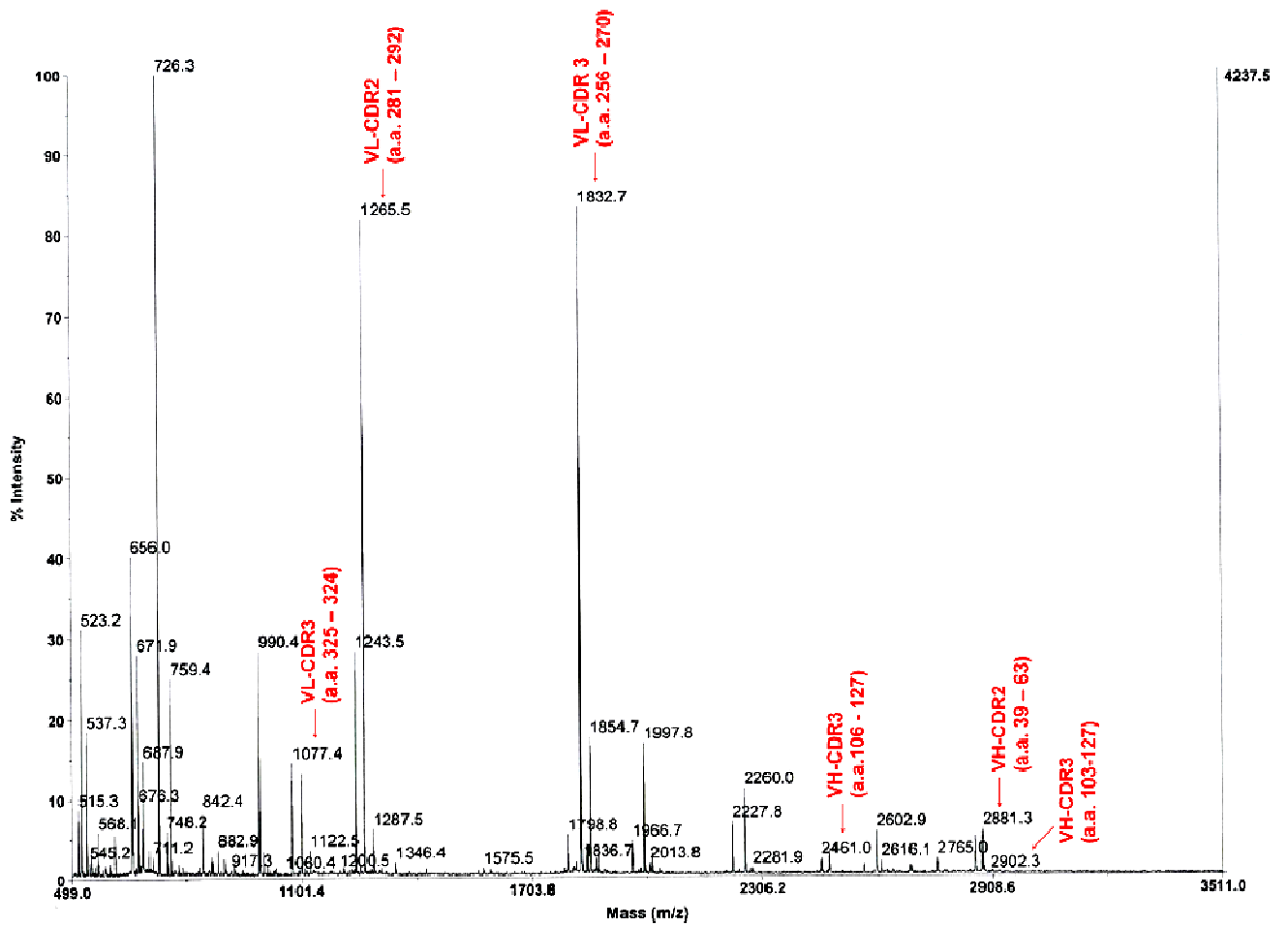


Figure 4.4: A representative MALDI spectrum of N-14F7 mAb digest. The observed CDR peptides are assigned as shown at the apex of the corresponding peaks.

Figure 4.5 displays a representative MALDI spectrum of the O-14F7 digest. As a result of the oxidation, the intensity of the non-oxidized peptides (for instance, m/z 1265.4 and m/z 1832.6) is significantly reduced. A new peak appeared at m/z 1281.4, which represent the oxidized form of the peptide ion at m/z 1265.4. The same observation was reported in Charvatova *et al.* [84], where oxidation of a protein called galectin-1 resulted in a reduction in the abundance of the non-oxidized tryptic peptides. Meanwhile, the heavy chain CDRs (which already have a low abundance in the mass spectrum of the non-oxidized form) have completely disappeared after oxidation of the antibody. This has added a significant challenge to the oxidative footprinting investigation of the heavy chain 14F7 CDR fragments, which are the most important regarding ligand binding. In general, the oxidation of amino acids (for instance, aliphatic amino acids, section 2.3.3.1) results in removal of hydrogen from their side chain. This might have a negative influence on the gas-phase basicity of peptides containing the modified amino acids.

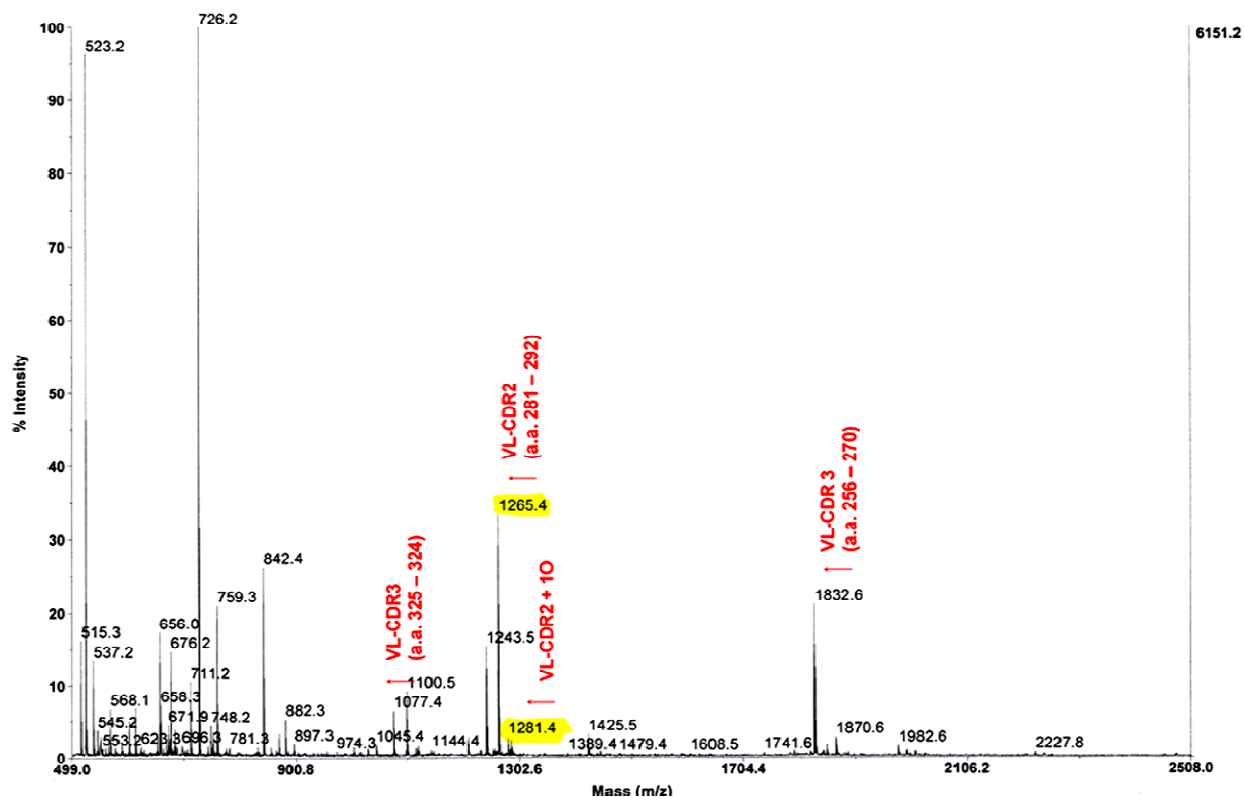


Figure 4.5: A representative MALDI spectrum of O-14F7 mAb digest. The intensities of the non-oxidized peptides further decreased, while the oxidized form of the VL-CDR2 is observed at m/z 1281 (VL-CDR2 + 10Oxygen atom). The peaks for the heavy chain CDRs were not observed.

Table 4.3 shows the comparison of the intensities of the six CDR peptides before and after oxidation of the 14F7 mAb. In almost all of the mass fingerprinting experiments performed, the relative abundance of the non-oxidized heavy chain CDR peptides remained less than 15%. Figure 4.6 shows a bar graph representation of the intensity values for the oxidized and non-oxidized forms.

Table 4.3: Comparison of the relative intensities of CDR peptides before and after oxidation of 14F7 mAb.

[M+H] ⁺	Position	Sequence (CDR residues shown in bold)	Relative abundance	
			Before oxidation *	After Oxidation *
1846	24-38 (VH-CDR1)	ASGYSFTSY WIHWL K	5.26 ± 0.4	0.34 ± 0.5
2881	39-63 (VH-CDR2)	QRPDQGLEWIGYIDPATAYTESN QK	7.50 ± 0.5	0
2461	106-127 (VH-CDR3)	GIYYYAMDY WGQGTTVTVSSAK	0.62 ± 0.9	0
1832	256-270 (VL-CDR1)	ASQ SISNNLHWYQQR	100.0 ± 0.0	21.54 ± 0.6
1266	281-292 (VL-CDR2)	YASQ SISGIPSR	32.95 ± 8.6	19.72 ± 1.8
1078	325-334 (VL-CDR3)	WPL TFGAGTK	13.10 ± 5.8	3.94 ± 0.0

* - Values represent the average value obtained from two experiments. Errors are measured to one decimal place.

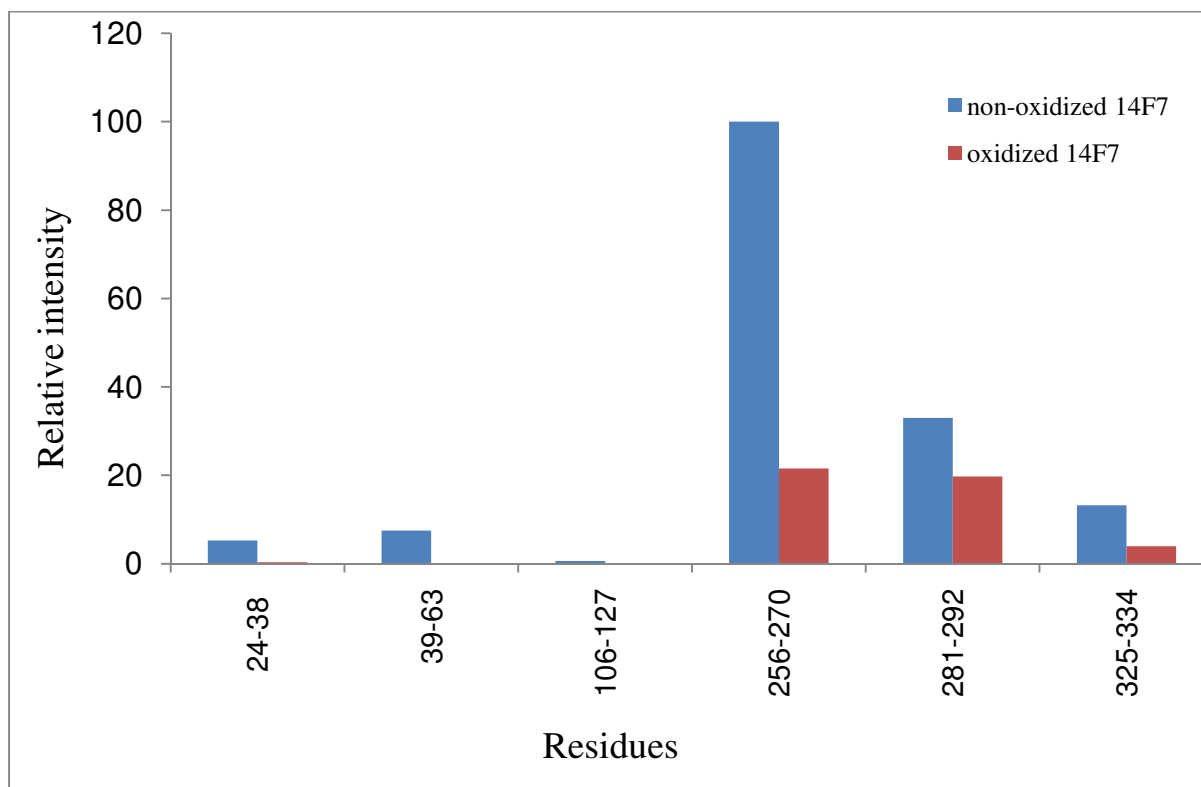


Figure 4.6: A bar graph representation of the relative intensities of the six CDR peptides before and after oxidation of the 14F7. The intensities of the heavy chain CDR peptides are clearly lower than that of the light chain.

4.3 ESI-LC/MS/MS analysis of O-14F7 and OL-14F7 digests

4.3.1 LC-MS total ion chromatograms

Figure 4.7 displays the total ion chromatograms (TIC) for a full-scan analysis between 500 and 2000 m/z of the N-14F7 mAb (top), O-14F7 mAb (middle) and the OL-14F7 complex (bottom).

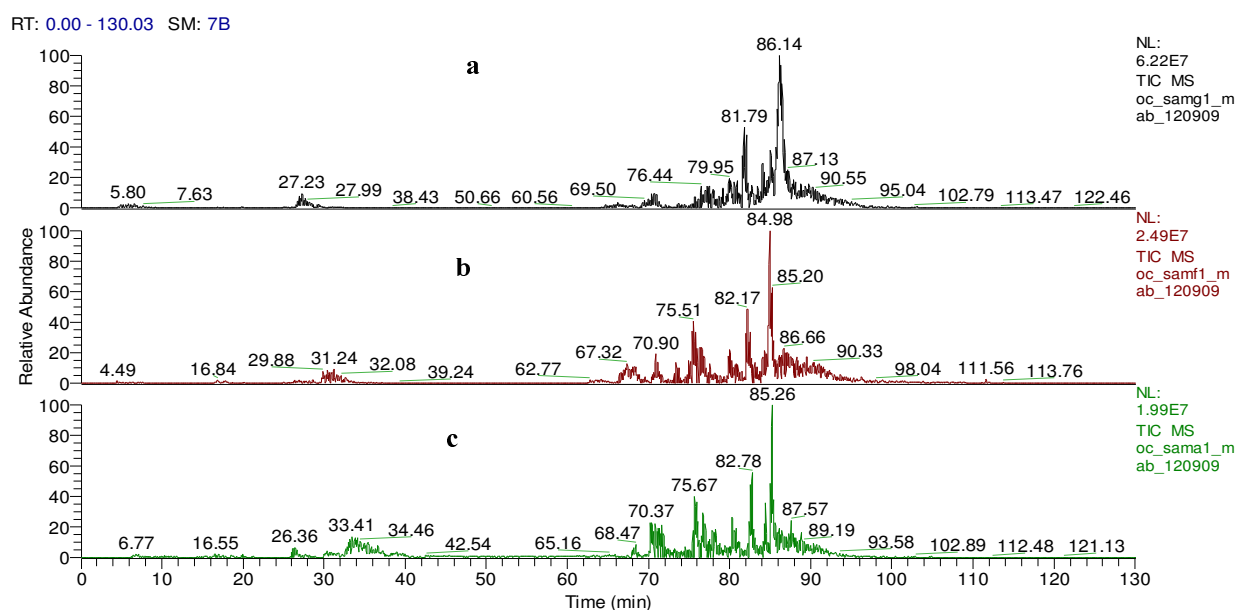


Figure 4.7: A representative of LC/MS TIC of (a) N-14F7 mAb (b) O-14F7 mAb and (c) the OL-14F7 complex.

4.3.2 Peptides and modification sites identified using ByOnic

Table 4.4 summarizes the O-14F7 tryptic peptides and oxidation sites identified by database searching against 14F7 Fab sequence on ByOnic. All the light chain CDR-related tryptic peptides (the last three in the Table) had undergone a significant oxidation at the indicated sites. The observed amino acid mass changes are in agreement with the mass changes characterized in the previous footprinting studies [84, 97], except for the 1 mass unit shift of the residue 258Q[+1] in the peptide 256-270. The heavy chain tryptic peptide segments 24-38 (at m/z 923.45) and 39-63 (at m/z 1440.69) contain the amino acid sequences of the VH-CDR1 and VH-CDR2, respectively. However, none of the amino acid residues within these peptide segments were found to be oxidized. Previous study shows that Asp52 within the peptide segment 39-63 has critical importance for ligand binding to the antibody [24]. However, only one peptide segment containing oxidized Asp52 was identified (score less 300) by ByOnic. The same peptides and oxidation sites were identified by ByOnic for the ligand-bound 14F7 tryptic digest, except Tyr114 of the peptide segment 106-127.

Table 4.4: A list of tryptic peptides (score > 400) and oxidation sites identified for O-14F7 by ByOnic database searching. Highlighted in bold are amino acid residues that belong to CDRs. Modified amino acids are underlined.

Residues	Sequence	[M+H] ⁺	m/z	Identified oxidation sites
24-38	ASGYSFTSYW II HWLK	1845.90	923.45	-
39-63	QRPDQGLEWIGYIDPATAYTES NQ K	2880.37	1440.69	-
106-127	GI YYYA <u>MDY</u> WGQGT <u>I</u> VTVSSAK	2459.12	1230.06	112M[+16] 114Y[+16] 120T[-2]
256-270	AS Q SIS NNLHWYQQR	1831.88	916.44	257S[-2] 258Q[+1]
281-292	YAS Q SIS GIPSR	1265.65	633.32	281Y[+16] 283S[+16] 286I[+14]/[16]
325-334	WPL TFGAGTK	1077.56	539.28	325W[16]/[+32] 326P[+16]

4.3.2.1 Oxidation of the tryptic peptide segments 106-127 and 325-334

The peptide segment 106-127 (Table 4.4) contains nine amino acids that belong to the VH-CDR3. A list of peptide matches within this segment for the O-14F7 and OL-14F7 are shown in Appendix 4, Table A3 and A4. In the case of the O-14F7, the search identified 32 potential peptide matches, 11 of which have oxidized methionine residue (Met112[+16]). The ion score and the probability value associated with the oxidized peptides suggest that the oxidation at Met112 has most likely occurred. In addition to the oxidized Met112, the search also identified three potential peptide matches with oxidized tyrosine residues (Tyr114[+16]) at position 114. Overall, the number of oxidatively modified peptides within residues 106-127 is 14 (11 oxidations at Met112 and 3 at Tyr114).

The equivalent analysis was undertaken for the OL-14F7 sample. Here, the database search identified 17 potential peptide matches. The total number of significant matches, as compared to that of O-14F7, was reduced by more than half. Furthermore, only six oxidized peptide matches were identified. The oxidations occurred solely at Met112 and not at Y114 (see Appendix 4, Table A4). Thus, the oxidative footprinting results obtained for the two states of the 14F7 mAb (with and without the ligand) suggest that there is a significant change in the overall oxidation pattern for peptide segment 106-127 after ligand binding.

Figure 4.8 shows an ESI mass spectra of the peptide segment 106-127 in the oxidized and non-oxidized forms. In the upper spectrum of the N-14F7 sample, the monoisotopic peaks at m/z 1230.61 and at m/z 1238.62 represent the doubly charged ($[M+2H]^{+2}$) non-oxidized and oxidized forms of the peptide segment 106-127, respectively. In the middle spectrum of the O-14F7 sample, the $[M+2H]^{+2}$ non-oxidized and oxidized forms of the peptide are located at m/z 1230.61 and at m/z 1238.61, respectively. In the lower spectrum of the OL-14F7 sample, the $[M+2H]^{+2}$ non-oxidized and oxidized forms of the peptide are located at m/z 1230.61 and at m/z 1230.60, respectively. The mass spectra differ mainly by the abundance (or intensity) of the oxidized form of the peptide ion. The peak that correspond to the oxidized peptide (at m/z 1238.62) is located in the noise region for the N-14F7 sample, whereas in the O-14F7 and OL-14F7 samples the abundance of this peptide is higher by more than two fold.

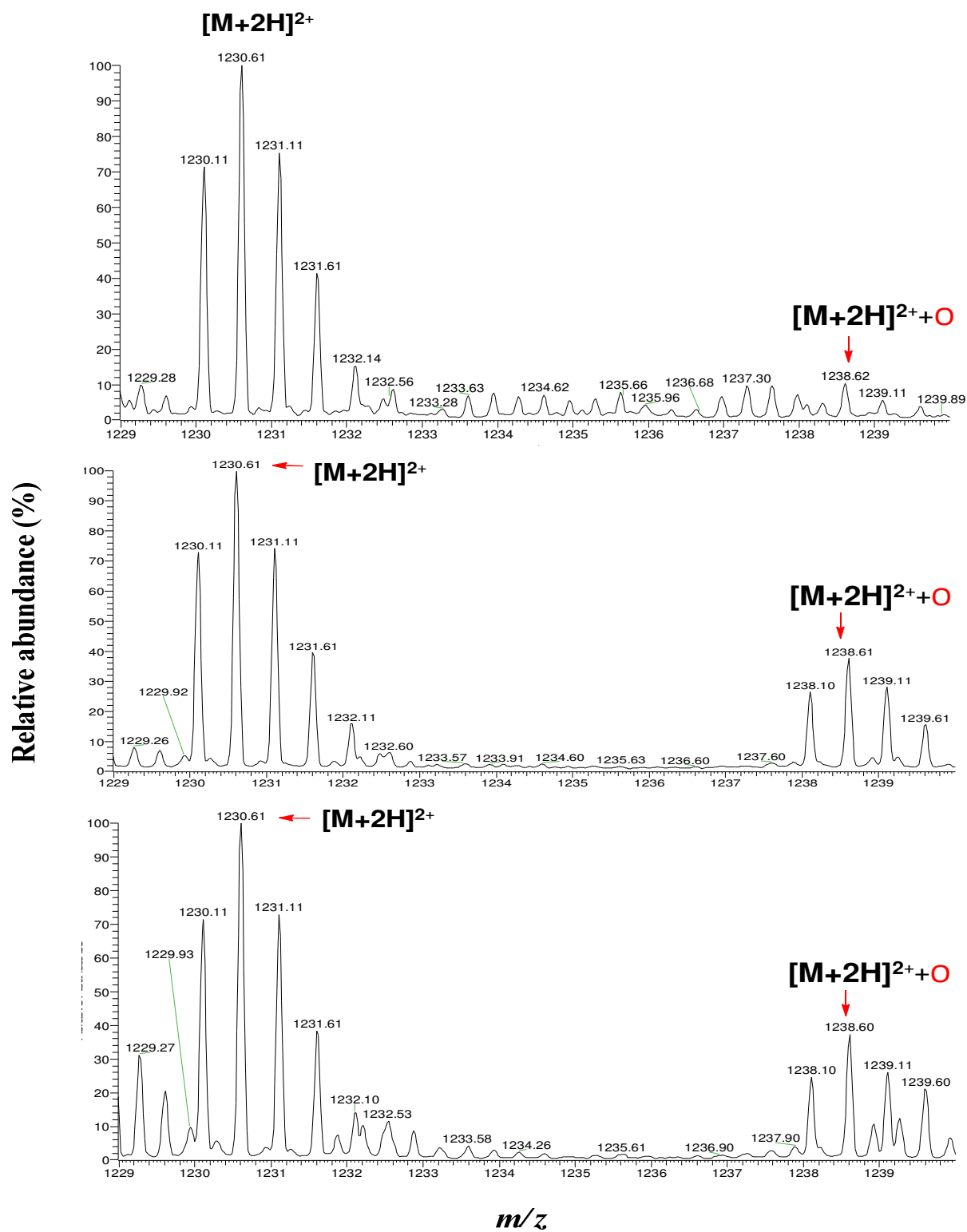


Figure 4.8: ESI-MS spectra of the VH peptide comprising residues 106-127 (GIYYAMDYWGQGTTVTVSSAK) of (a) N-14F7, (b) O-14F7 and (c) the OL-14F7 complex. The spectra were averaged over the time range of 0.01-130 min.

Figure 4.9 shows the comparison of the absolute intensities of the oxidized and the non-oxidized forms of the peptide segment 106-127 in the O-14F7 and OL-14F7 (intensity values are obtained from the mass spectra in Figure 4.8). The intensities of both the oxidized and non-oxidized forms of the peptide is lower by half in the OL-14F7 complex, as compared to that of the O-14F7. The low abundance of the peptide with a m/z 1238.60 in the OL-14F7 could be due to the fact that the residue side chain responsible for the radical-induced oxidation is protected upon ligand binding. The most likely reason for the decrease in intensity of the peptide with a m/z 1230.61 in the OL-14F7 could be due to the inaccessibility of the tryptic sites around this particular amino acid sequence after the antibody has made binding interactions with the ligand.

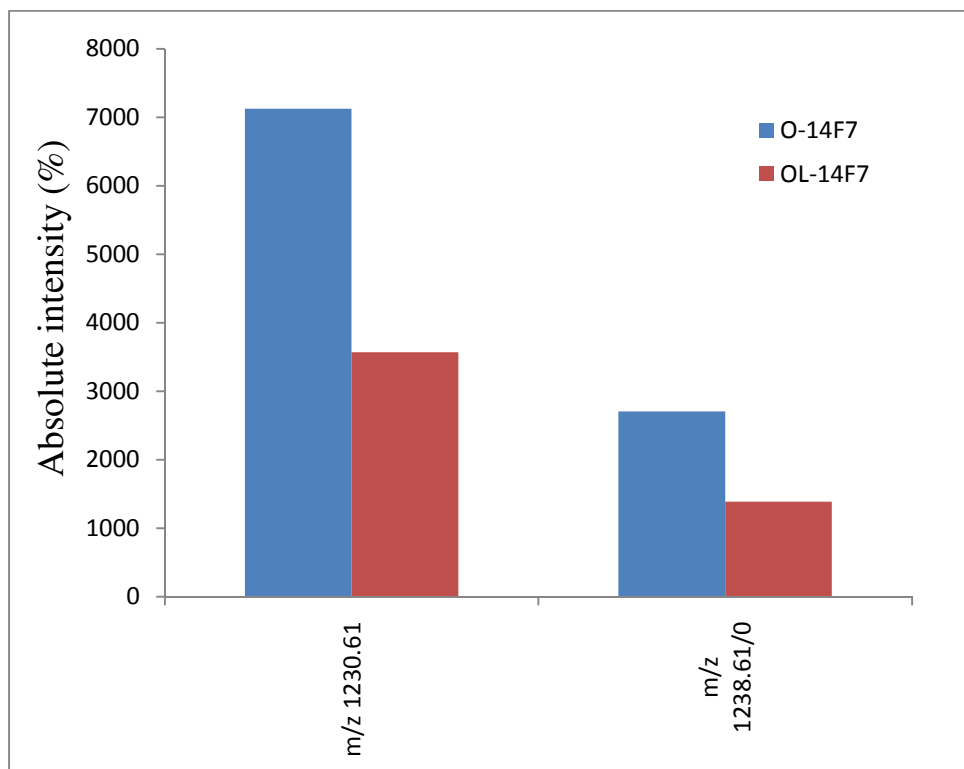


Figure 4.9: A bar graph representation of the intensities of peptide segment 106-127 (oxidized and non-oxidized forms) for the O-14F7 and OL-14F7 samples. The intensity of both the non-oxidized and the oxidized forms of the peptide ion is lowered by half in the ligand-bound 14F7 digest.

The equivalent analysis was undertaken for the tryptic peptide segment 325-334. Here, the peptide contains four amino acids that correspond to the VL-CDR3 (see Table 4.4). A list of peptide matches within this segment identified for the O-14F7 and OL-14F7 are shown in Appendix 4, Table A5 and A6. The search identified 81 potential peptide matches in the O-14F7 sample, of which 18 are modified. Multiply oxidized peptides at Trp325 (W[+14]/[+16]/[+32]) and Pro326 (P[+16]) were observed.

In the case of OL-14F7 sample, ByOnic identified 96 potential peptide matches, of which only 6 are modified by hydroxyl radicals. The oxidations occurred at Trp325 ([+14]/[+16]) and Pro326 ([+16]). The low number of oxidized peptides in this particular sample (as compared to that of O-14F7) suggests that there is a change in the oxidation pattern within the peptide after ligand binding to the antibody.

4.3.3 MS/MS spectra – sequence and oxidation site confirmations

The amino acid sequences and oxidation sites within the tryptic peptide fragments 106-127 and 325-334 were manually confirmed from the ESI-MS/MS product ion spectra. The same amino acid sequences and oxidation sites were determined for both the O-14F7 and OL-14F7 samples. Therefore, here, only the analysis for the O-14F7 MS/MS spectra is discussed.

For the tryptic fragment 106-127, the MS/MS spectra of the doubly charged oxidized and non-oxidized forms of the peptide are shown in Figure 4.10. This peptide would have a theoretical y_n ($n= 1-21$) and b_n ($1-21$) fragment ion series shown at the top of Figure 4.10. *De novo* peptide sequencing via manual interpretation of MS/MS spectra was performed by comparing the m/z values of each peak in the spectra with the m/z values of the theoretical fragment ion series. The y_n ions ($n= 4-17$) were identified and assigned in the upper and lower spectra. The m/z values of these ions gave mass differences that correspond to the monoisotopic masses of residues 111-123 (AMDYWGQGTTVTV). Moreover, the m/z values for the y_n ions ($n = 4-5$ and $8-15$) that contain the last 15 amino acids from the C-terminal side are identical in both the upper and lower spectra. However, the m/z values of two of the y_n ions- y_{16} and y_{17} , were observed to increase by 16 mass units in the MS/MS spectrum of the oxidized form (Figure 4.10b, lower spectrum). This led to the conclusion that oxidation in this peptide occurred at Met112. The presence of oxidized

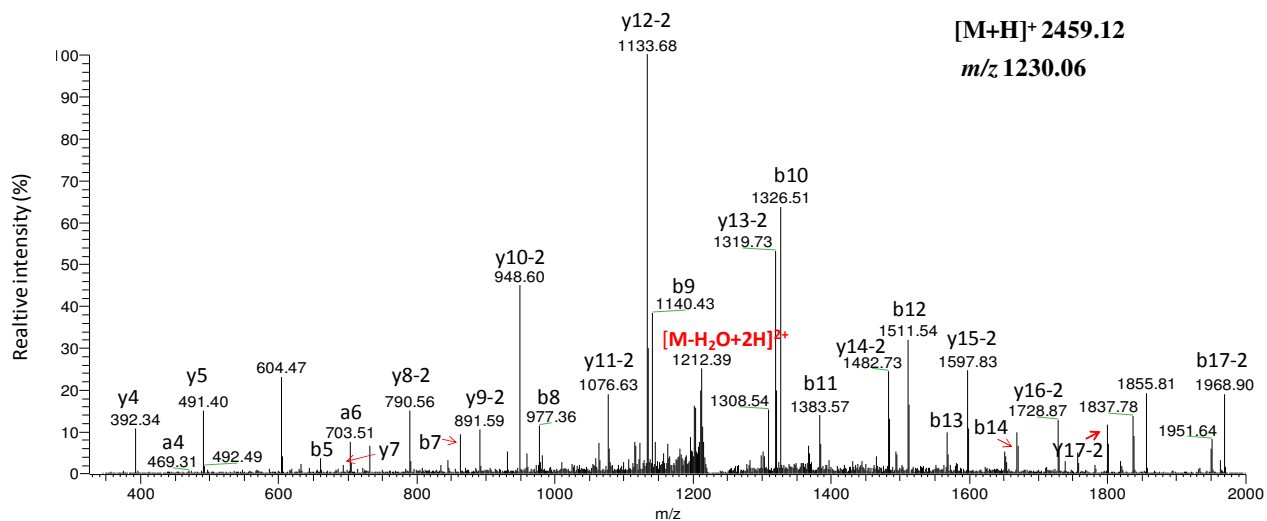
b_n ions ($n = 7-17$) that contain the first 17 residues from the N-terminus confirms this assignment.

In both spectra, the m/z values of the y_n ions ($n = 8-16$) and b_{17} were shifted by -2 unit from the predicted m/z values. This indicates that a second oxidation occurred within residues TTVTVSSAK towards the C-terminus. Meanwhile, the presence of the unmodified y_7 ion in the upper mass spectrum indicates the most probable oxidation site to be at the second threonine (Thr120) from the N-terminus. The side chain of the Thr120 is first oxidized to a carbonyl group (+16) followed by loss of one water molecule [97].

A product ion spectrum with oxidized Tyr114 was not identified in both the O-14F7 and OL-14F7 MS/MS scans, although ByOnic found three oxidized peptides at this residue.

58.0	171.1	334.1	497.2	660.2	731.3	862.3	977.3	1140.4	1326.51	1383.5	1511.5	1568.6	1669.6	1770.6	1869.7	1970.7	2069.8	2156.8	2243.8	2314.8	
b1	b2	b3	b4	b5	b6	b7	b8	b9	b10	b11	b12	b13	b14	b15	b16	b17	b18	b19	b20	b21	
G	I	Y	Y	Y	A	M	D	Y	W	G	Q	G	T	T	V	T	V	S	S	A	K
y21	y20	y19	y18	y17	y16	y15	y14	y13	y12	y11	y10	y9	y8	y7	y6	y5	y4	y3	y2	y1	
2403.9	2290.8	2127.8	1964.7	1801.7	1730.6	1599.6	1484.6	1321.5	1135.5	1078.5	950.4	893.4	792.3	691.3	592.2	491.2	392.1	305.1	218.1	147.1	

a)



b)

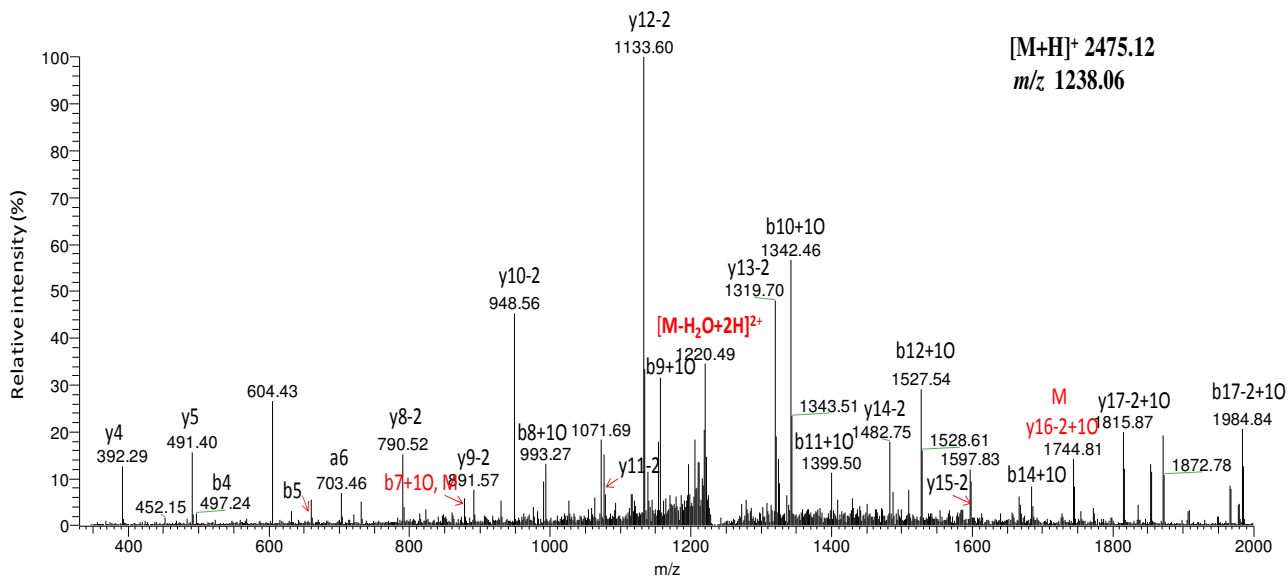
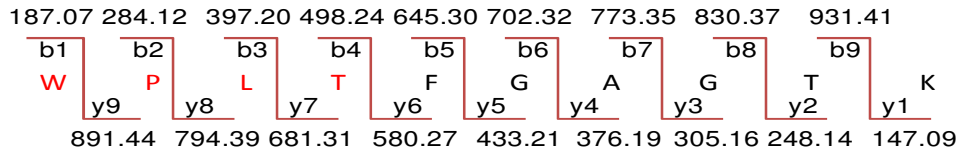


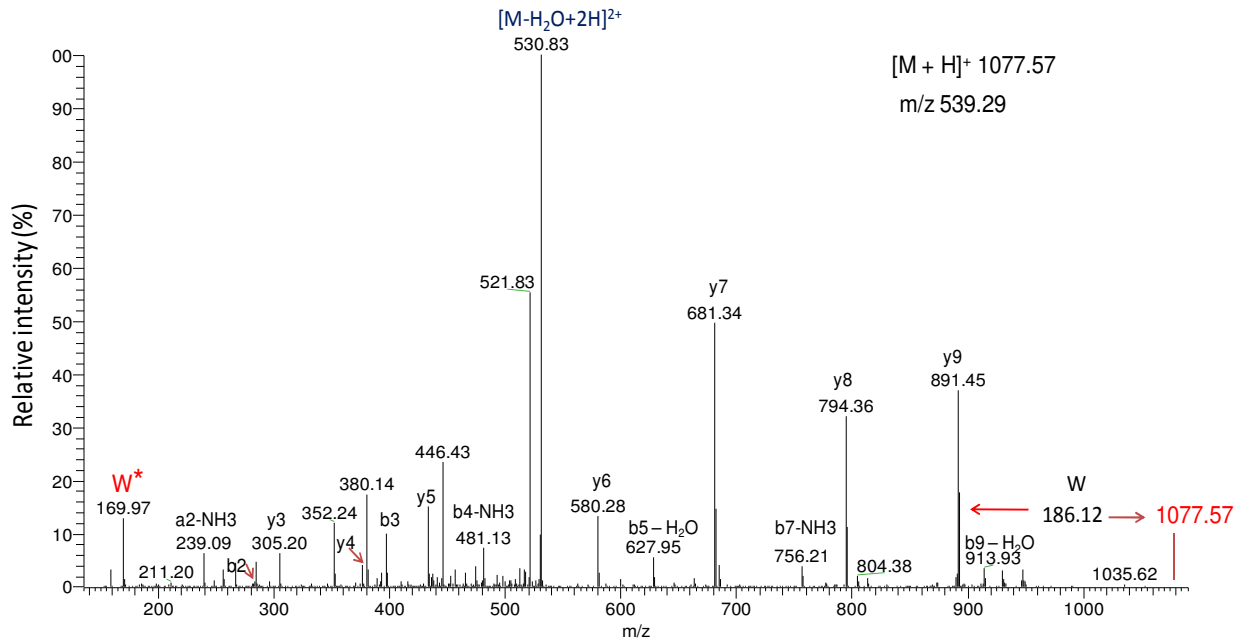
Figure 4.10: ESI-MS/MS spectra of the peptide segment 106-127 (GIYYAMDYWGQGTTVTVSSAK) for the (a) non-oxidized form (b) oxidized form. Sequencing of the spectra showed that Met112 (b7 ion from the N-terminus) is oxidized. The same oxidation site was observed in the OL-14F7 MS/MS spectra.

A similar procedure was followed for sequencing the tandem mass spectra of the oxidized and non-oxidized forms of the peptide segment 325-334. Each peak in the mass spectra was manually assigned by matching its m/z value with the predicted y - and b -series m/z values shown at the top of Figure 4.11. All the identified y_n ions ($n = 3-9$) have identical m/z values in both spectra. Meanwhile, a shift of +16 mass units occurred at the b_n ions ($n = 2-3$ and 6) in the mass spectrum of the oxidized form (Figure 4.11b). However, the y_9 ion (at m/z 891.45), which also contains the proline residue as the b_2 ion, is not modified. Therefore, the oxidation must have occurred at Trp325 on the N-terminal side of the proline. This is confirmed by the increase in mass of the tryptophan fragment by 16 mass units (from m/z 169.97 to m/z 186.06).

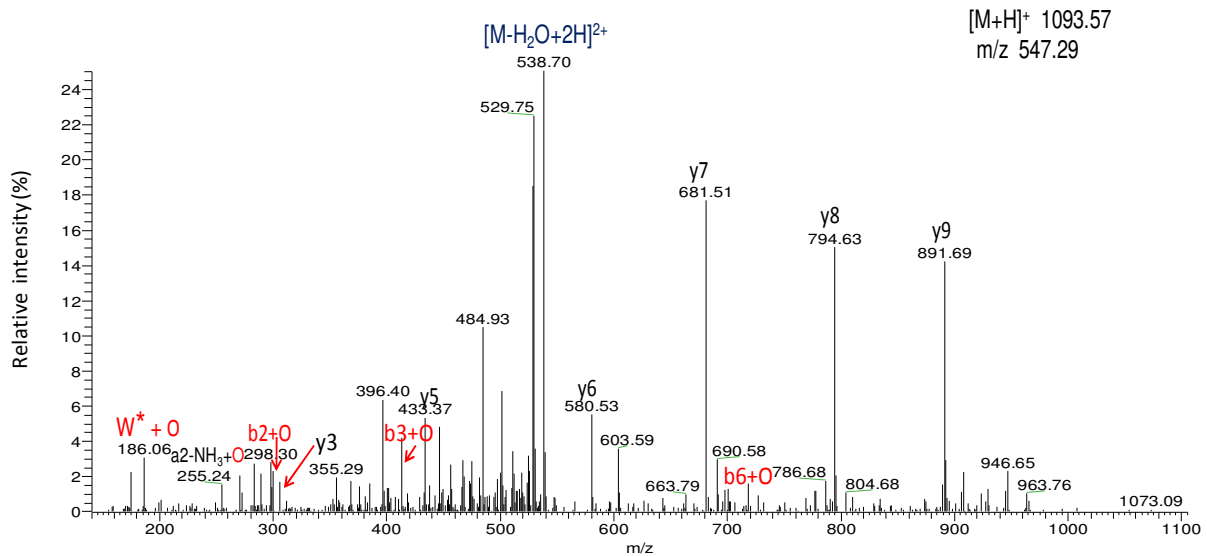
Detailed MS/MS analysis of multiply oxidized peptides at Trp325 and Pro326 were not possible due to few numbers of MS/MS scans and high level of noise within their product ion spectra.



a)



b)



* m/z 169.97- Tryptophan fragment mass [111].

Figure 4.11: ESI-MS/MS spectra of the peptide segment 325-334 (WPLTFGAGTK) for the (a) non-oxidized (b) oxidized forms. Sequencing of the spectra showed that Trp325 (b1 ion from the N-terminus) is oxidized. The same oxidation site was identified in the OL-14F7 MS/MS spectra.

4.3.4 Percentage oxidation and level of protection of modified peptides

To determine the percentage of oxidation, selected ion chromatograms (SIC) were generated for oxidized and non-oxidized forms of tryptic peptides from the LC-MS signals of both the O-14F7 and OL-14F7. The level of oxidation of each peptide was then calculated based upon the areas under the selected chromatograms (see section 2.3, equation 2, page 35). Figure 4.12-15 shows the integrated chromatogram peak areas for the oxidized and non-oxidized forms of the peptide segments 106-127 and 325-334. Furthermore, the level of protection (after ligand binding) of the amino acid side chains within each peptide was determined as in a previous study [82] (also see appendix 5 for an example of how to calculate the percentage oxidation and levels of protection). Table 4.5 summarizes the calculated percentages of oxidation and levels of protection for the peptide segments 106-127 and 325-334 derived from a replicate experiment. The errors recorded for most of the experiments performed were quite high. This is due to malfunctioning of the liquid chromatography system in the duration of sample analysis for this project.

As shown previously in Figure 4.9, page 74, the spectral intensity of the oxidized form of the peptide segment 106-127 was lowered by half in the OL-14F7 sample. In Table 4.5, the percentage of oxidation of the residue side chains within the same peptide is shown to be lower (as compared to the O-14F7) by a factor of approximately two upon ligand binding to the 14F7. This is equivalent to a 45 % protection. This finding matches the database search result already discussed under section 4.3.2.1, page 72, concerning the lower number of oxidized Met112 observed in the OL-14F7.

In addition, a significant degree of protection (over 85 %) is observed for the side chains of the light chain peptide residues 325-334. One reason for such a large percentage of protection could be that one or several amino acid side chains within this region interact directly with the ligand. Meanwhile, in this study the percentage oxidation of the peptide containing only oxidized Trp325 was found to decrease after ligand binding. This, however, is contrary to the phage display and 14F7 light-chain shuffling study, which showed the light chain region to have no ligand binding significance [32]. An alternative explanation could be that the overall 3D

structure of the light chain CDRs might have undergone a sizable conformational change during complex formation. However, in the present and previous docking simulations of the complex such a large conformational change has not been observed.

Table 4.5: Percentage oxidation of the tryptic peptide segments 106-127 and 325-334 of 14F7 mAb in the presence and absence of the ligand.

Residues	Peptide sequence	m/z [M+2H] ²⁺	m/z [M+O+2H] ²⁺	Percentage of oxidation before ligand binding*	Percentage of oxidation after ligand binding*	Level of protection of the amino acid side chains within the segment after ligand binding (%)
106-127	GIYYAMDYWGQGTTVVSSAK	1230.60	1238.60	28.9 ± 6.6	15.8 ± 4.1	45
325-334	WPLTFGAGTK	539.29	547.29	5.51 ± 3.1	0.75 ± 0.3	86

*values represent average values from two experiments.

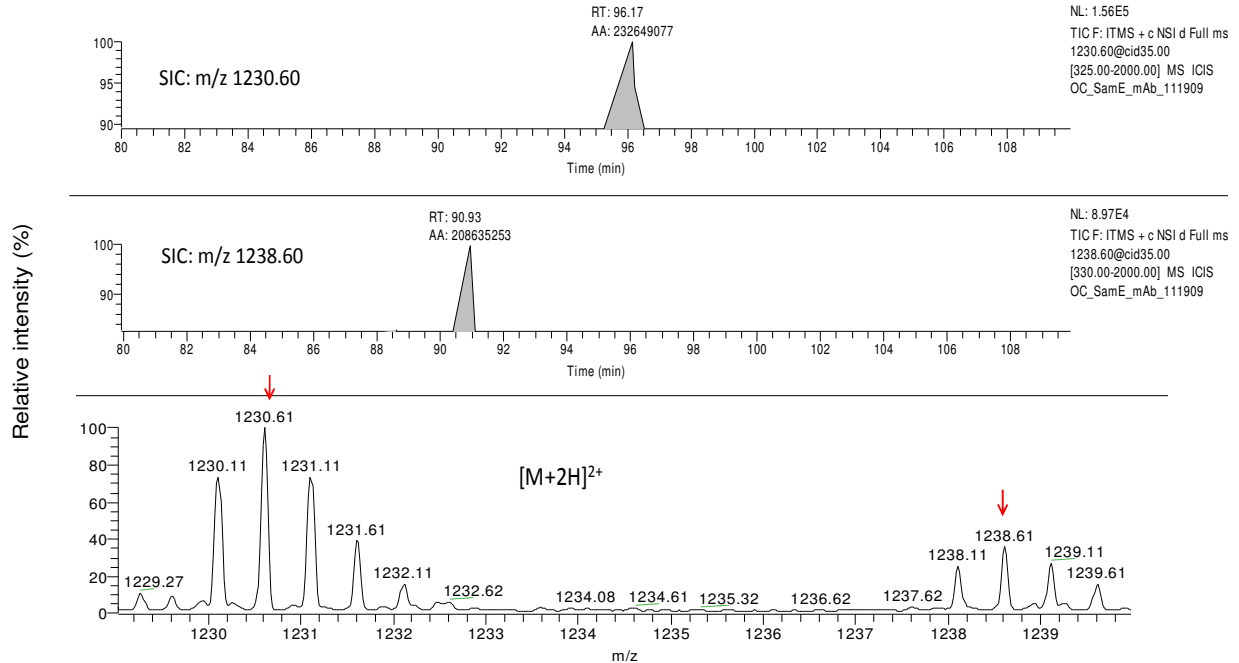


Figure 4.12: A representative SIC with integrated peak area for the oxidized (at m/z 1238.61) and non-oxidized (at m/z 1230.61) forms of the peptide segment 106-127 (GIYYAMDYWGQTTVTVSSAK) driven from the O-14F7 MS scans.

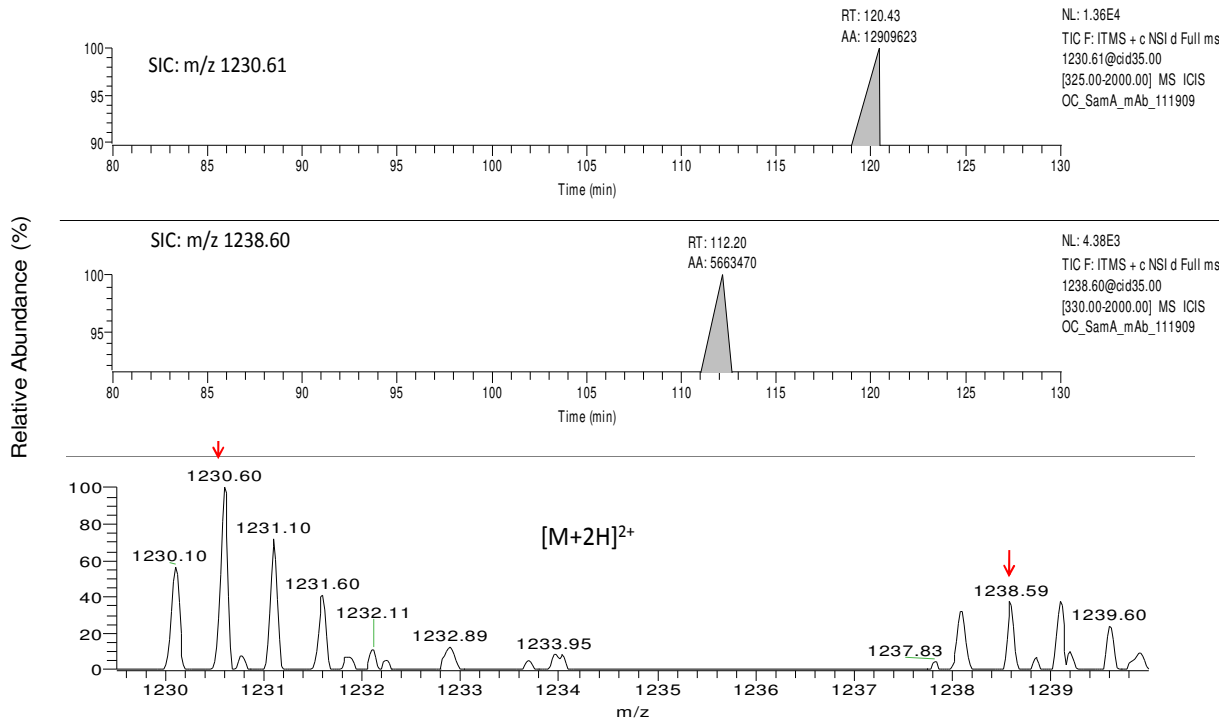


Figure 4.13: A representative SIC with integrated peak area for the oxidized (at m/z 1238.60) and non-oxidized (at m/z 1230.59) forms of the peptide segment 106-127 (GIYYAMDYWGQTTVTVSSAK) driven from the OL-14F7 MS scans.

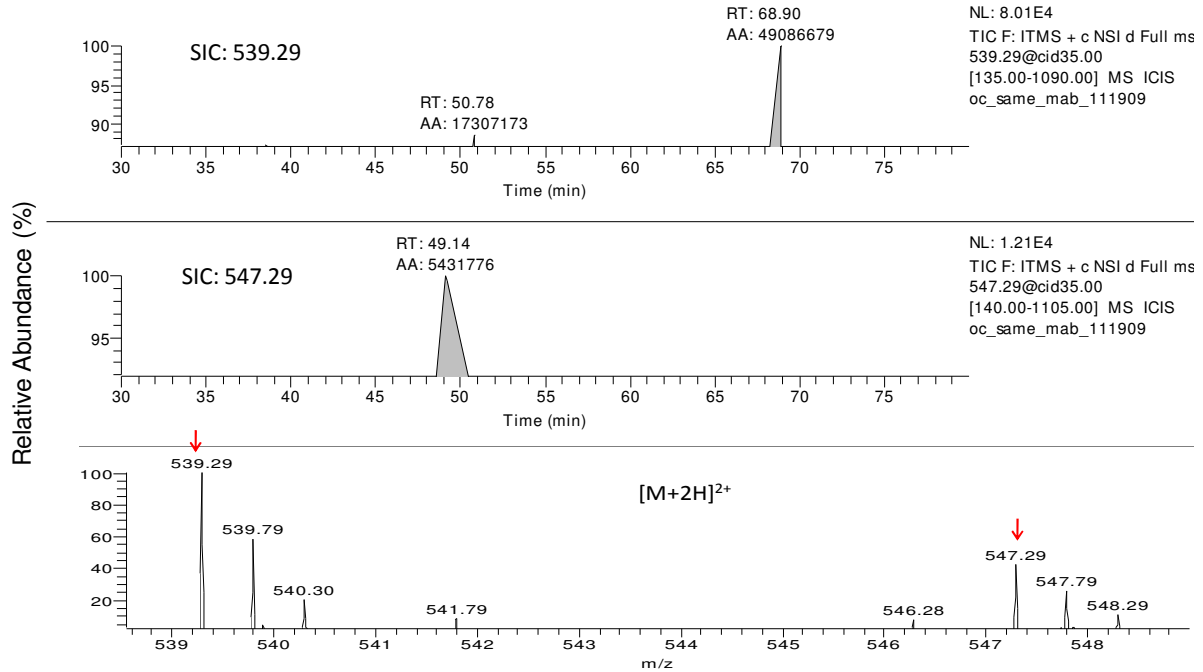


Figure 4:14: A representative SIC with integrated peak area for the oxidized (at m/z 547.29) and non-oxidized (at m/z 539.29) forms of the peptide segment 325-334 (WPLTFGAGTK) extracted from the O-14F7 MS scans.

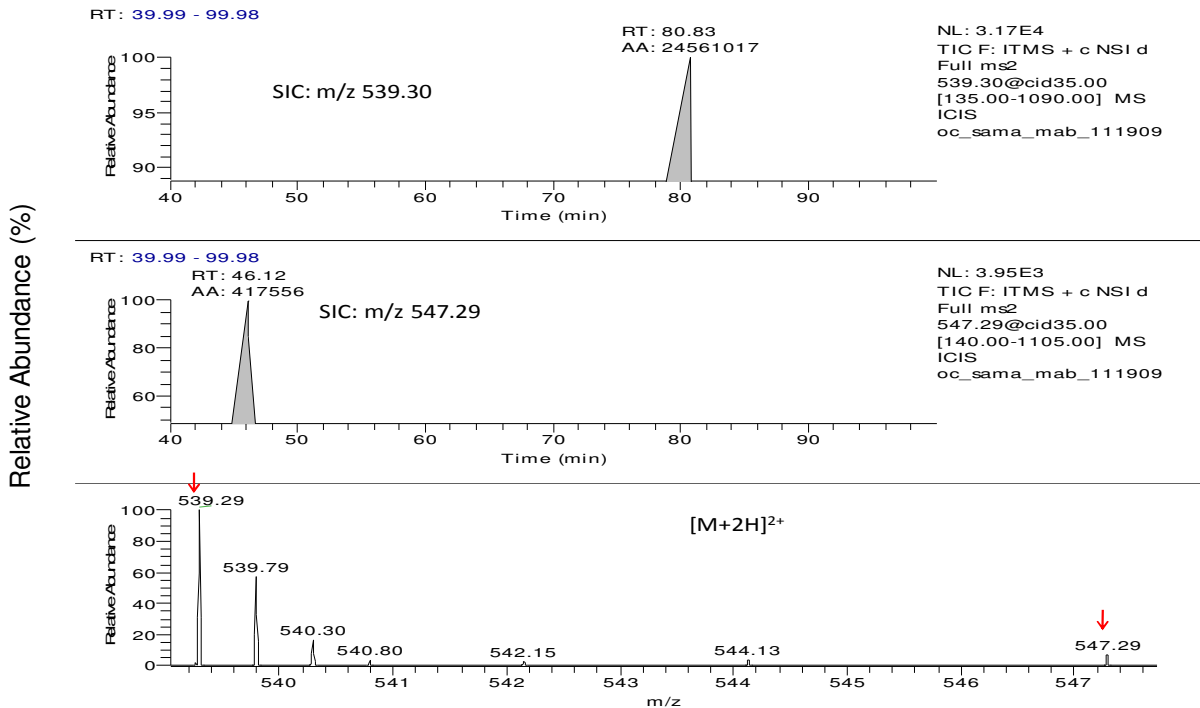


Figure 4:15: A representative SIC with integrated peak area for the oxidized (at m/z 547.29) and non-oxidized (at m/z 539.29) forms of the peptide segment 325-334 (WPLTFGAGTK) extracted from the OL-14F7 MS scans.

4.4 A new structural model of 14F7-NeuGc-GM3 complex

Several years ago, the structural model of the 14F7 Fab fragment in complex with the terminal disaccharide of N-glycolyl GM3, NeuGc α 3Gal β , was proposed as describe in Ref. [24]. Docking and molecular dynamics simulations were used to predict the complex model and characterize the binding interactions.

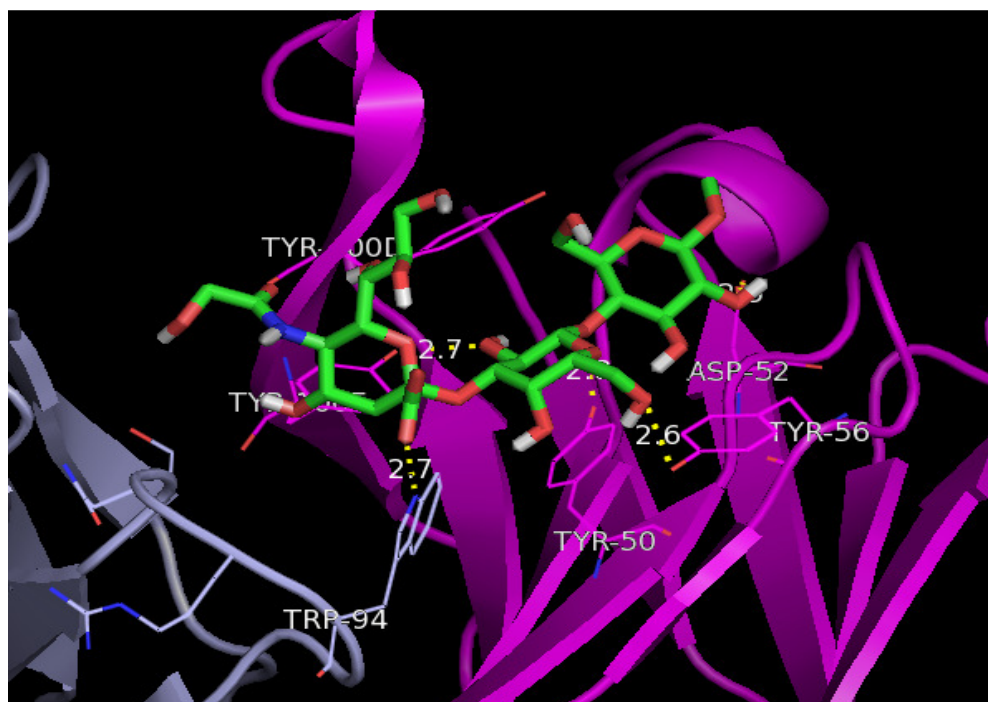
In the present docking simulations, each run resulted in one conformation of the ligand complex. Table 4.6 shows 14 distinct conformational clusters obtained using an r.m.s.d. tolerance of 2.0 Å. Run 30, ranked 6th, has the highest number of clusters (ca. 66). The lowest and mean docked energy for this run are -6.47 and -6.03 kcal/mol, respectively. The ligand conformation for this particular run was visualized with PyMOL. Figure 4.16 (a) and (b) shows the conformation of the antibody-antigen complex obtained from run 30. The complete list of the potential hydrogen bonding interactions within the complex are presented in Table 4.7. The new ligand conformation is quite different from the previous model, in two main aspects. First, in the former model, the ligand is hydrogen bonded to the 14F7 Fab mainly through the polar head (NeuGc) group, whereas in the new ligand complex the binding interactions are distributed over all three sugar residues. Second, in the former model, the binding interactions are limited to the CDRs of the VH region. In the new model, a tryptophan residue (Trp325) in the VL region engages in a hydrogen bond (through its indole ring) with the NeuGc carboxylate. This agrees well with the footprinting result. As derived in section 4.3.4 (Table 4.5), the peptide containing the Trp325 exhibits a reduced level of oxidation after ligand binding, which could be due to protection by the ligand.

Moreover, the strong hydrogen bonding interaction of the aspartate residue (VH Asp52) in this docking model agrees well with previous site-directed mutagenesis results, that shows mutation at this residue heavily affects ligand binding. With this new model, we can also explain why the mutation of Asp52 to similarly sized hydrophobic residues (valine and isoleucine) completely abolished binding, instead of inducing a specificity switch to NeuAc-GM3, as could be expected from the previous model (unpublished data). An attempt has been made to validate the new complex model by molecular dynamics simulations. However, a completed result could not be achieved within the time limit of this thesis work.

Table 4.6: Docking conformational clusters ranked according to energy.

Clus	Lowest	Run	Mean	Num	Histogram
-ter	Docked		Docked	in	
Rank	Energy		Energy	Clus	5 10 15 20 25 30 35
					: : : : :
1	-8.60	53	-8.31	5	#####
2	-7.07	61	-6.82	3	###
3	-6.95	94	-6.95	1	#
4	-6.92	71	-6.65	2	##
5	-6.57	83	-6.48	3	###
6	-6.47	30	-6.03	66	#####
7	-6.34	35	-6.10	5	#####
8	-6.33	95	-6.33	1	#
9	-6.20	49	-5.97	2	##
10	-6.00	26	-6.00	1	#
11	-5.88	42	-5.57	6	#####
12	-5.88	69	-5.88	1	#
13	-5.87	75	-5.83	3	###
14	-5.59	12	-5.59	1	#

a)



b)

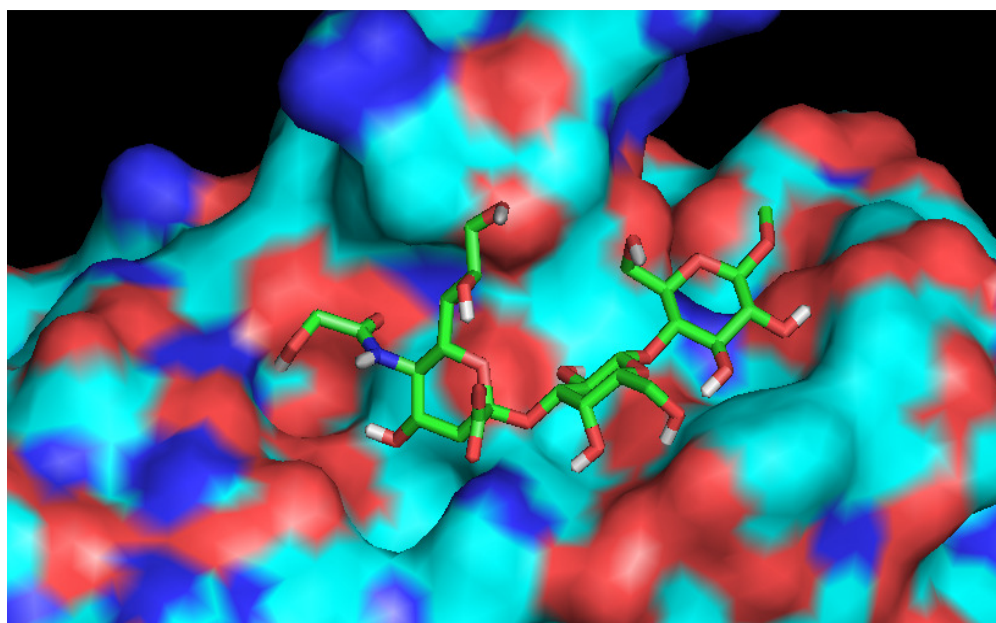


Figure 14.16: A docking model of Fab 14F7-NeuGc-GM3 trisaccharide complex with (a) Selected hydrogen bonding interactions, (b) an electrostatic surface representation.

Table 4.7: Potential hydrogen bonding interactions between the 14F7 Fab and the NeuGc(α 2-3)Gal(β 1-4)Glc ganglioside, as predicted by docking simulations.

Amino acid	sugar	Distance (Å)
Tyr H50 OH	Gal β 4 O4 (glycosidic bond)	2.8
Try H109 OH	Gal β 4 O6 (C2-O-C6) or Gal β 4 H2O	2.9 2.7
Tyr H57 OH	Gal β 4 H6O	2.6
Glu H58 CO (carbo. Acid)	Gal β 4 H4O	3.3
Asp H52 1O	Glc β 1 H2O or Glc β 1 O1 (O-ceramide)	2.5 3.2
Thr H55 OH	Glc β 1 H2O	3.3
Gly H100B C=O	NeuGc α 3 H7O	2.9
Try H108 C=O	NeuGc α 3 O5N	3.1
W L94 NH	NeuGc α 3 O1A (carb. Acid)	2.7

4.5 Correlation of the footprinting result with SASA of residue side chains

The extent of solvent accessibility of amino acid side chains is a critical factor in the hydroxyl-radical based footprinting analysis of macromolecular complexes [112]. The change in the percentage of oxidation before and after complex formation can easily be traced when the potential probe sites are fully accessible and highly reactive. The solvent accessibilities of amino acid side chains of 14F7 Fab were calculated using the NACCESS program [113]. Although the VH-CDR1 and VH-CDR2 are highly solvent accessible (see Appendix 6), none of the residues within this region are found oxidized, both in the O-14F7 and OL-14F7 samples. Even without oxidation, these residues were hardly seen in the MALDI-TOF MS spectra of the 14F7 mAb and Fab digests. One reason for this could be a high stability of the heavy chain CDR structure, thereby resisting tryptic digestion. Therefore, the footprinting analysis was not successful in characterizing the effect of ligand binding on the residues within the VH-CDR1 and VH-CDR2.

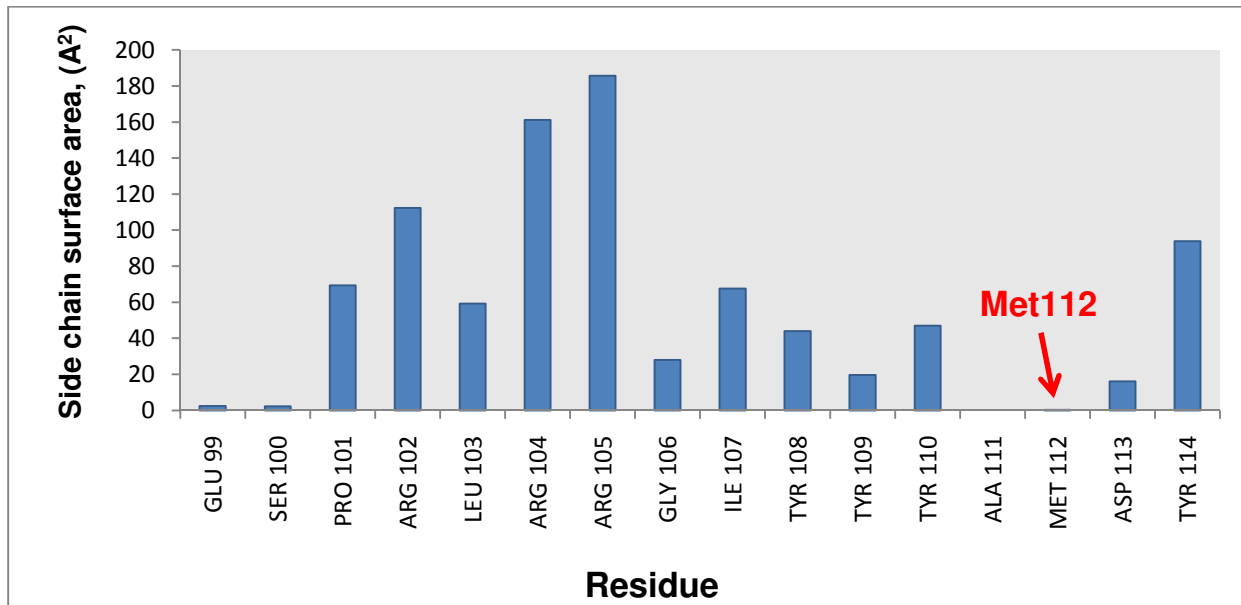
Meanwhile, the levels of oxidation of a segment of the VH-CDR3 peptide (residues 106-127) in the free and ligand bound forms of the 14F7 were amenable to detailed analysis. Peptides containing oxidized Met112 were found abundantly in the O-14F7. This oxidation site has been located in the MS/MS spectrum of the fragment ion mass with good certainty. On the contrary, the solvent accessible surface area of the Met112 residue side chain (as calculated from the crystallography data) is surprisingly less than 1 \AA^2 (see Figure 4.17a). From this point of view, one would expect that only a very small amount of Met112 would be oxidized. Moreover, in the crystal structure the residue is buried inside a cavity which the ligand can hardly reach. Therefore, the question is how did it get oxidized so abundantly in the unliganded 14F7? A similar observation was reported in previous studies by Hambly *et al.* and J.S. Sharp *et al.* [87, 114]. Both groups found out that a methionine residue with essentially no solvent accessible surface area was oxidized by hydroxyl radicals. Kiselar *et al.* [115] suggested a different mechanism by which solvent inaccessible methionines could substantially be oxidized. The process takes place by an intermolecular radical transfer mechanism, in which a hydrogen is first abstracted from a solvent accessible reactive residue. Then a donor group proximate to the initial radical site transfer an electron to the radical. The solvent accessible residue would then pick up

a proton and transfers the radical to the electron donor. A molecular oxygen would then oxidize the radical without it being solvent accessible. Trp and Tyr residues were shown to participate strongly in this radical transfer mechanism.

In 14F7, there are four tyrosine (Tyr108, Tyr109, Tyr110 and Tyr114) side chains close to the Met112 residue within peptide 106-127, which could speed up the radical transfer, thereby causing a substantial oxidation of the residue. The 45 % decrease in the percentage of oxidation of residues 106-127 (Table 4.5) after ligand binding could then be attributed to the protection of the intermolecular transfer of the radicals that are responsible for oxidizing Met112. This suggests that there is in fact a protection conferred by the trisaccharide ligand in the vicinity of the VH-CDR3 region of the antibody. An alternative explanation could be that 14F7 in the crystals adopts a different structure than in solution, and indeed the VH-CDR3 is modeled in substantial crystal contacts. On the other hand, docking simulation suggest that the conformation of this loop is rather stable, also in solution.

The absolute solvent accessible surface area of the Trp325 side chain is more than 80 percent in the free 14F7 Fab crystal structure (Figure 4.17b) as calculated by NACCESS. The 86 percent protection (Table 4.5) of the peptide containing Trp325 after ligand binding could either be due to a large conformation change or shielding conferred by the ligand. Meanwhile, the solvent accessible surface area calculation of the docked complex model suggest that Trp325 is less accessible in the ligand bound 14F7 complex (data not shown).

a)



b)

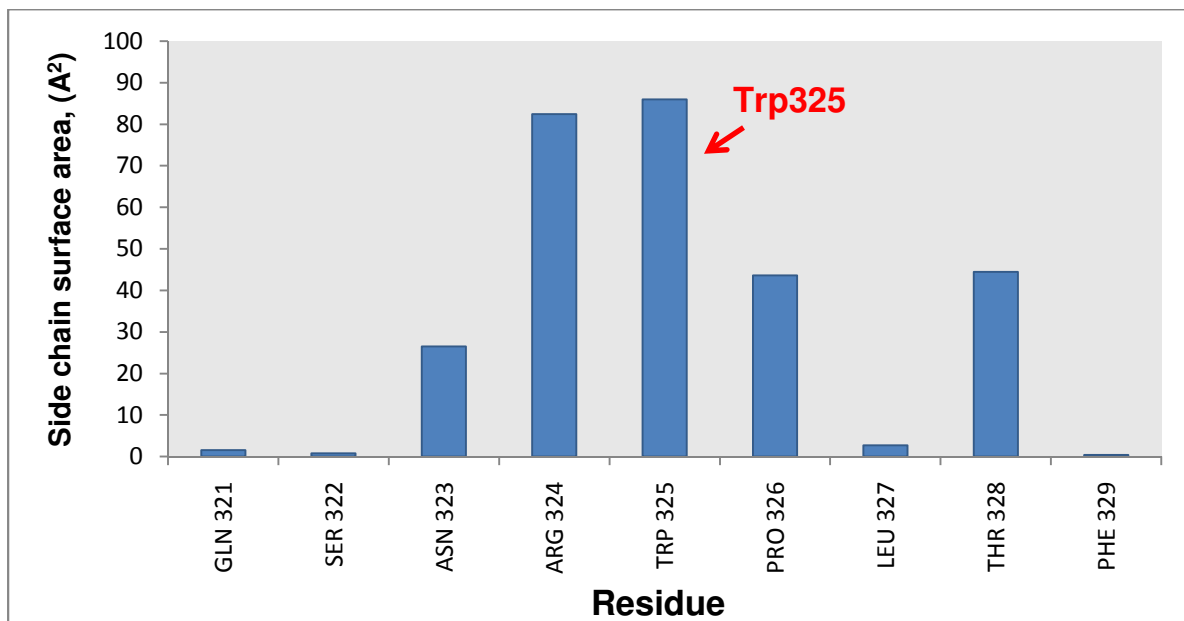


Figure 4.17: Side chain surface area of (a) VH-CDR3 residues (b) VL-CDR3 residues, as calculated by NACCESS using the crystal structure of the 14F7 Fab.

5. CONCLUSION AND FURTHER WORKS

Conclusion

This work represents the first experimental investigation of the binding interactions between the 14F7 mAb and N-glycolyl GM3 ganglioside using oxidative footprinting and mass spectrometry methods.

The amino acids important to the interaction are characterized based on experimental and computational analysis. The side chains of Met112 (of VH-CDR3) and Trp325 (of VL-CDR3) were protected from radical-induced oxidation upon ligand binding to the antibody. This result agrees well with a new model of the ligand complex, generated by docking simulations. In the new model, the ganglioside is bound mainly to the heavy chain CDRs with high surface complementarity and strong hydrogen bonding interactions. However, in contrast to the previous findings, the results obtained from both the footprinting analysis and molecular docking strongly suggest that the ligand also binds to the light chain residue Trp325.

Further work

While the oxidative footprinting coupled to MS is a promising approach, a more detailed characterization of the binding interactions requires optimization of some of the experimental conditions.

One of the challenges during this thesis was the low abundance of the heavy chain CDRs in the MS spectra. Several proteases have been tested to improve coverage, although with little success. The search for suitable enzymes should be extended and it might be helpful to perform the digestions using a combination of proteases or separately by targeting one or two CDRs at a time. In this regard, the V8 protease was successful in cleaving a full peptide sequence containing the VH-CDR3 residues and chymotrypsin was found to give a good coverage of the VH-CDR2. Also chemical cleavage may be considered.

Building a more precise model of the antibody-antigen complex structure is also crucial. This model should be based on all the experimental data and further refined by molecular dynamics

simulations. Most importantly, however, the new model – refined or not – can now be taken as a basis for further experimental validation, for example by site directed mutagenesis of selected 14F7 residues.

APPENDIX

Appendix 1

Buffers

A. SDS-PAGE

1. 1XTris/Glycine/SDS running buffer (1 L)

100 ml of 10X Tris/Glycine/SDS from Bio-Rad
900 ml Nanopure water

B. ELISA

1. TBST (Tris buffered saline with Tween 20)

50 μ L Tween 20
100 mL 1X TBS, pH 7.4

2. 0.05M Tris-HCl buffer (1 L)

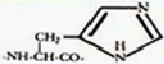

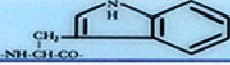
6.057 g of Tris base
800 mL of Nanopure water
pH adjusted to 7.8 by HCl
Final volume brought to 1 L with Nanopure water

3. 0.5 % Blocking solution (100 mL)

0.5 g of non-fat dry milk
95 mL of 0.5 M Tris-HCl buffer, pH 7.4

Appendix 2

Table 1A: The 20 naturally existing amino acid residues.

The Masses and Compositions of the Twenty Commonly Occurring Amino Acid Residues				
Symbols	Name and Composition	Residue Structure	Monoisotopic Mass	Average Mass
Ala A	Alanine C ₃ H ₅ NO	$\begin{array}{c} \text{CH}_3 \\ \\ \text{-NH-CH-CO-} \end{array}$	71.03711	71.0788
Arg R	Arginine C ₆ H ₁₂ N ₄ O	$\begin{array}{c} \text{CH}_2\text{-(CH}_2\text{)}_2\text{-NH-C-NH}_2 \\ \\ \text{NH} \\ \\ \text{-NH-CH-CO-} \end{array}$	156.10111	156.1876
Asn N	Asparagine C ₄ H ₆ N ₂ O ₂	$\begin{array}{c} \text{CH}_2\text{-CONH}_2 \\ \\ \text{-NH-CH-CO-} \end{array}$	114.04293	114.1039
Asp D	Aspartic Acid C ₄ H ₅ NO ₃	$\begin{array}{c} \text{CH}_2\text{-COOH} \\ \\ \text{-NH-CH-CO-} \end{array}$	115.02694	115.0886
Cys C	Cysteine C ₃ H ₅ NOS	$\begin{array}{c} \text{CH}_2\text{-SH} \\ \\ \text{-NH-CH-CO-} \end{array}$	103.00919	103.1448
Glu E	Glutamic Acid C ₅ H ₇ NO ₃	$\begin{array}{c} \text{CH}_2\text{-CH}_2\text{-COOH} \\ \\ \text{-NH-CH-CO-} \end{array}$	129.04259	129.1155
Gln Q	Glutamine C ₅ H ₈ N ₂ O ₂	$\begin{array}{c} \text{CH}_2\text{-CH}_2\text{CONH}_2 \\ \\ \text{-NH-CH-CO-} \end{array}$	128.05858	128.1308
Gly G	Glycine C ₂ H ₃ NO	$\text{-NH-CH}_2\text{-CO-}$	57.02146	57.0520
His H	Histidine C ₆ H ₇ N ₃ O		137.05891	137.1412
Ile I	Isoleucine C ₆ H ₁₁ NO	$\begin{array}{c} \text{CH(CH}_3\text{)CH}_2\text{-CH}_3 \\ \\ \text{-NH-CH-CO-} \end{array}$	113.08406	113.1595
Leu L	Leucine C ₆ H ₁₁ NO	$\begin{array}{c} \text{CH}_2\text{CH(CH}_3\text{)}_2 \\ \\ \text{-NH-CH-CO-} \end{array}$	113.08406	113.1595
Lys K	Lysine C ₆ H ₁₂ N ₂ O	$\begin{array}{c} \text{CH}_2\text{-(CH}_2\text{)}_3\text{-NH}_2 \\ \\ \text{-NH-CH-CO-} \end{array}$	128.09496	128.1742
Met M	Methionine C ₅ H ₉ NOS	$\begin{array}{c} \text{CH}_2\text{-CH}_2\text{-S-CH}_3 \\ \\ \text{-NH-CH-CO-} \end{array}$	131.04049	131.1986
Phe F	Phenylalanine C ₉ H ₉ NO	$\begin{array}{c} \text{CH}_2\text{-Ph} \\ \\ \text{-NH-CH-CO-} \end{array}$	147.06841	147.1766
Pro P	Proline C ₅ H ₇ NO		97.05276	97.1167
Ser S	Serine C ₃ H ₅ NO ₂	$\begin{array}{c} \text{CH}_2\text{-OH} \\ \\ \text{-NH-CH-CO-} \end{array}$	87.03203	87.0782
Thr T	Threonine C ₄ H ₇ NO ₂	$\begin{array}{c} \text{CH(OH)CH}_3 \\ \\ \text{-NH-CH-CO-} \end{array}$	101.04768	101.1051
Trp W	Tryptophan C ₁₁ H ₁₀ N ₂ O		186.07931	186.2133
Tyr Y	Tyrosine C ₉ H ₉ NO ₂		163.06333	163.1760
Val V	Valine C ₅ H ₉ NO	$\begin{array}{c} \text{CH(CH}_3\text{)}_2 \\ \\ \text{-NH-CH-CO-} \end{array}$	99.06841	99.1326

Appendix 3

Mass spectra

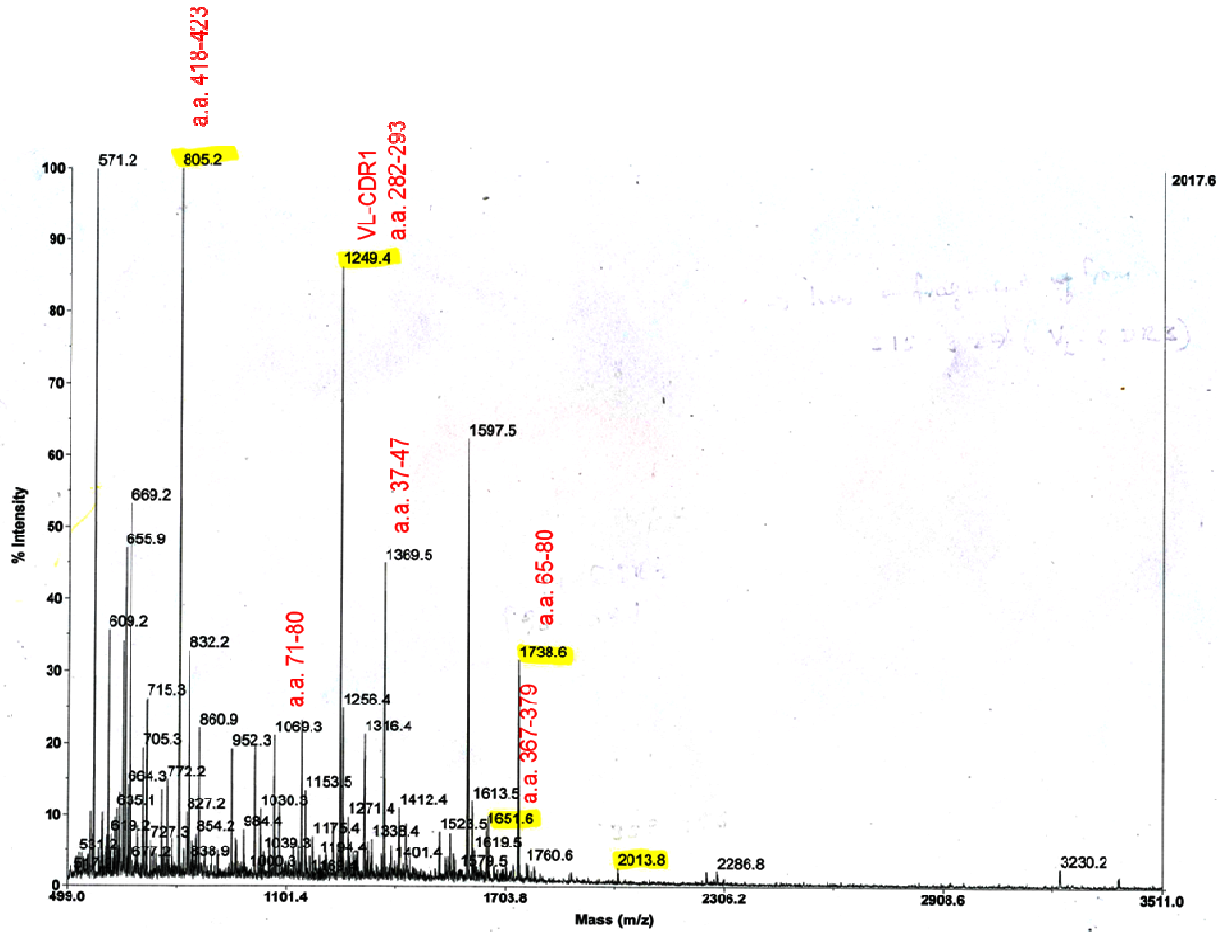


Figure A1. MALDI spectrum of 14F7 mAb chymotryptic digests.

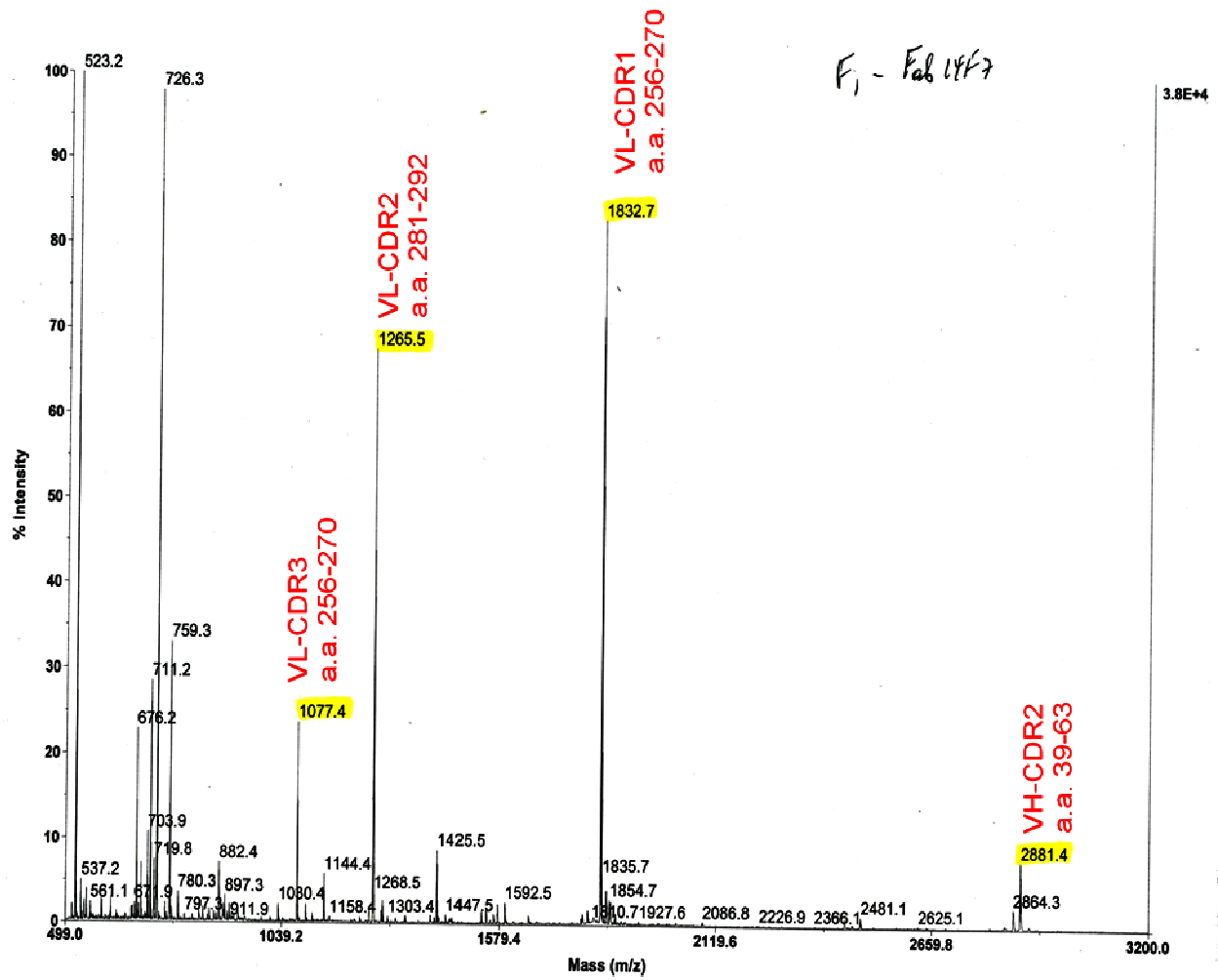


Figure A2. MALDI spectrum of 14F7 Fab tryptic digests.

Appendix 4

Table A2: VH peptides (residues 106-127) identified by database searching on ByOnic for the N-14F7 MS/MS spectra (oxidized peptides at Met112 are highlighted in red).

Peptide	Mass (M+H)	m/z	Cleavage	Score	Delta	LogProb
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	447.6	19.7	-0.01
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	495.9	0	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	581.1	7.7	-5.49
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	528	0	-4.71
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	521.6	7.2	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	406.6	7.8	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	454.2	11.4	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	483.1	0	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	499.8	0	-4.28
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	460.6	17.9	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	437	9.8	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	522.8	13.3	-0.01

Table A3: VH peptides (residues 106-127) identified by database searching on ByOnic for the O-14F7 MS/MS spectra. (oxidized peptides at Met112 and Y114 are highlighted in red and purple, respectively).

Peptide	Mass (M+H)	m/z	Cleavage	Score	Delta	LogProb
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	608.8	7.2	-0.01
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	529.1	0	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	617	617.3	-9.95
R.GIYYYAMDY[+16]WGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	631.3	12.9	-6.49
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	564.8	7.1	-0.01
R.GIYYYAM[+16]DYWGQGTTVT[-2]VSSAK.T	2475.12	1238.064	Tryptic	641.5	13.4	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	539.3	0	-0.01
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	437.8	0	-0.01
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	619.8	0	-0.01
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	572.1	0	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	543.8	2.5	-0.01
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	634.2	0	-6.45
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	481	4.8	-0.01
R.GIYYYAM[+16]DYWGQGTTVT[-2]VSSAK.T	2475.12	1238.064	Tryptic	660.5	4.7	-7.05
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	606.9	21.4	-0.01
R.GIYYYAM[+16]DYWGQGTTVT[-2]VSSAK.T	2475.12	1238.064	Tryptic	615.7	12.7	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	429.8	430.1	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	446.2	0	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	470.1	6.5	-0.01
R.GIYYYAM[+16]DYWGQGTTVT[-2]VSSAK.T	2475.12	1238.064	Tryptic	607	6.2	-0.01
R.GIYYYAMDY[+16]WGQGTTVT[-2]VSSAK.T	2475.12	1238.064	Tryptic	417.9	8.3	-0.01
R.GIYYYAMDY[+16]WGQGTTVT[-2]VSSAK.T	2475.12	1238.064	Tryptic	446.8	7	-3.49
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	536	536.3	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	501.3	501.6	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	497.2	0	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	513.6	0	-0.01
R.GIYYYAM[+16]DYWGQGTTVT[-2]VSSAK.T	2475.12	1238.064	Tryptic	404.8	1.7	-0.01
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	526.4	21.5	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	582.4	9.1	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	561.3	561.6	-8.85
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	525.5	0	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	515.1	20.1	-0.01

Table A4: VH peptides (residues 106-127) identified by ByOnic database searching for the OL-14F7 complex MS/MS spectra (oxidized peptides at Met112 are highlighted in red).

Peptide	Mass (M+H)	m/z	Cleavage	Score	Delta	LogProb
R.GIYYYAM[+16]DYWGQGTTVT[-2]VSSAK.T	2475.12	1238.064	Tryptic	633.9	16.1	-6.53
R.GIYYYAM[+16]DYWGQGTTVT[-2]VSSAK.T	2475.12	1238.064	Tryptic	558.3	29.4	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	604.1	0	-5.86
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	636.8	0	-6.65
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	453.5	4.1	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	537.8	16.8	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	537.5	0	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	482.6	9.3	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	488.1	9.8	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	414	5.3	-0.01
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	420.4	0	-0.01
R.GIYYYAM[+16]DYWGQGTT[-2]VTVSSAK.T	2475.12	1238.064	Tryptic	414.7	4.7	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	562	18.9	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	441.8	0	-0.01
R.GIYYYAMDYWGQGTT[-2]VTVSSAK.T	2459.12	1230.064	Tryptic	423.3	5.4	-0.01
R.GIYYYAM[+16]DYWGQGTTVT[-2]VSSAK.T	2475.12	1238.064	Tryptic	423.2	47.1	-0.01
R.GIYYYAMDYWGQGTTVT[-2]VSSAK.T	2459.12	1230.064	Tryptic	559.9	11.8	-5.19

Table A5: VL peptides (residues 325-334) identified by database searching on ByOnic for the O-14F7 MS/MS spectra. (oxidized peptides are highlighted in red)

Peptide	(M+H) ⁺	m/z	Cleavage	Score	Delta	LogProb
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	568.5	467.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	610	610.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	616.4	616.7	-0.01
R.W[+32]P[+16]LTFGAGTK.L	1125.56	563.284	Tryptic	415.6	23.2	-3.2
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	588.4	588.7	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	590.5	590.8	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	469.7	282.7	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	449.2	255.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	650.2	650.5	11.55
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	488.1	488.4	-0.01
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	479.7	12.7	-0.01
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	520.8	17.1	-4.65
R.W[+14]P[+16]LTFGAGTK.L	1107.55	554.279	Tryptic	419.2	48.9	-3.5
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	459.4	297.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	609.5	609.8	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	426.5	260.9	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	638.3	638.6	-0.01
R.W[+32]PLTFGAGTK.L	1109.56	555.284	Tryptic	444.4	20.1	-3.45
R.W[+32]P[+16]LTFGAGTK.L	1125.56	563.284	Tryptic	412.2	15.1	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	406.6	259.3	-0.01
R.W[+16]PLTFGAGTK.L	1093.57	547.289	Tryptic	411.1	5.3	-2.81
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	554.2	554.5	-0.01
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	400.7	12.7	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	415.2	228.5	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	429.6	226.8	-0.01
R.W[+32]PLTFGAGTK.L	1109.56	555.284	Tryptic	411.8	5.6	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	592	592.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	587	587.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	600.2	600.5	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	617.7	618	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	464.7	252	-0.01
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	469.4	12.9	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	559.5	559.8	-0.01
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	482.2	13.4	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	544.7	545	-0.01
R.W[+16]P[+16]LTFGAGTK.L	1109.56	555.284	Tryptic	451.9	0.3	-3.44
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	436.2	11.6	-0.01
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	444.9	11.2	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	613.9	614.2	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	604.4	604.7	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	412.2	211.8	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	443.6	279.2	-0.01
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	429	5.9	-0.01

R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	495.8	12.3	-0.01
R.W[+16]P[+16]LTFGAGTK.L	1109.56	555.284	Tryptic	435.7	0.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	566	566.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	553.5	553.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	590.1	590.4	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	565.8	566.1	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	595.9	596.2	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	432.9	273.8	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	481.6	305.9	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	578.1	578.4	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	474.6	255.4	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	526.5	526.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	481.4	481.7	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	543.3	543.6	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	549.6	549.9	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	492.1	492.4	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	588.9	589.2	-0.01
R.W[+32]PLTFGAGTK.L	1109.56	555.284	Tryptic	416.9	4.1	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	494.7	335	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	407.9	224.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	612	447.7	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	530.7	410	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	456.2	298.2	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	469.3	240.9	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	431.2	274.7	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	436.1	268.1	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	429.5	231.6	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	502.9	310.2	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	427.6	213.2	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	444	264.5	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	450.6	281.6	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	517.4	322.6	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	506.9	350.6	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	425	225.1	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	452.4	264.1	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	437.6	437.9	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	402.7	258.9	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	454.1	454.4	-0.01

Table A6: VL peptides (residues 325-334) identified by database searching on ByOnic for the OL-14F7 complex MS/MS spectra. (oxidized peptides are highlighted in red)

Peptide	(M+H) ⁺	m/z	Cleavage	Score	Delta	logProb
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	420.5	420.8	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	450.8	253.1	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	424.2	281.4	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	409.5	213.3	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	444.9	299.4	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	441.4	289	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	470.7	296.4	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	455.8	314.9	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	515.5	515.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	621.2	621.5	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	562.2	562.5	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	401.2	198.9	-0.01
R.W[+14]PLTFGAGTK.L	1091.55	546.279	Tryptic	448.2	1.4	-3.44
R.W[+16]P[+16]LTFGAGTK.L	1109.56	555.284	Tryptic	421.2	0.3	-3
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	520	342.1	-0.01
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	474.5	20.1	-3.96
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	549.9	550.2	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	565	565.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	593	593.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	579.1	579.4	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	483.3	483.6	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	520	520.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	505.4	505.7	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	603.8	604.1	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	612.6	612.9	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	592.6	592.9	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	555.3	555.6	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	553	553.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	634.1	634.4	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	529.3	529.6	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	649.9	650.2	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	593.2	593.5	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	627.2	627.5	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	602.2	602.5	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	658.4	658.7	-11.75
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	538.8	539.1	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	507.4	507.7	-0.01
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	467	15.2	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	597	597.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	592.9	593.2	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	539.9	540.2	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	602.2	602.5	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	579.1	579.4	-0.01

R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	515.9	384.8	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	412.4	233.1	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	620.3	496	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	598.6	598.9	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	580.5	580.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	472.7	473	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	590	590.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	634.3	634.6	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	598.7	599	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	570.5	413.2	-0.01
R.W[+16]PLTFGAGTK.L	1093.57	547.289	Tryptic	408.9	5.3	-2.81
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	446.3	243.2	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	481.1	295.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	560.9	561.2	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	540.4	540.7	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	584.1	584.4	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	578.2	468.6	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	515.2	515.5	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	543.6	452.6	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	536.1	536.4	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	464.6	335.7	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	445.9	272.7	-0.01
R.WP[+16]LTFGAGTK.L	1093.57	547.289	Tryptic	400.9	6.8	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	434.6	278.3	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	421.3	232.6	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	455.5	455.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	521.2	521.5	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	574.5	574.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	552.9	553.2	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	480.5	313.8	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	422.3	260.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	559.9	560.2	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	535	535.3	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	444.7	245.4	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	454.4	454.7	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	418.1	203.9	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	473.6	234.8	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	434.1	294.1	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	496	330.3	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	449.3	449.6	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	466.5	466.8	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	556.4	556.7	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	581.2	581.5	-0.01
R.WPLTFGAGTK.L	1077.57	1077.57	Tryptic	540	540.3	-0.01
R.WPLTFGAGTK.L	1077.57	539.289	Tryptic	422.7	288.8	-0.01

Appendix 5

Calculation of percentage of oxidation

Example: calculation of percentage oxidation of the tryptic peptide segment 106-127 of 14F7 mAb.

Table A7: SIC peak areas for the peptide segment 106-127 in O-14F7 and OL-14F7.

Peptide modified	O-14F7		OL-14F7	
	Non-oxidized peptide SIC peak area ($A_{\text{non-oxid.}}$)	Oxidized peptide SIC peak area ($A_{\text{oxid.}}$)	Non-oxidized peptide SIC peak area	Oxidized peptide SIC peak area
GIYYAMDYWGQTTVTV SSAK	266,349,423.8	228,737,244.6	13,188,189.34	5663469.78

Percentage oxidation of the peptide in O-14F7,

$$\begin{aligned}
 \% \text{ oxid. peptide} &= \frac{A_{\text{oxid.}}}{A_{\text{oxid.}} + A_{\text{non-oxid.}}} \times 100 \\
 &= \frac{228737244.6}{266349423.8 + 228737244.6} \times 100\% \\
 &= 46.20\%
 \end{aligned}$$

Thus, the percentage of oxidation of the peptide before ligand binding is 46.20 %. Similarly, the percentage of oxidation of the peptide after ligand binding was found to be 30.04 %. The level of protection of amino acid side chains within the peptide upon ligand binding was calculated as follows.

$$\begin{aligned}
 \% \text{ protection} &= \frac{\% \text{ oxid. peptide in O-14F7} - \% \text{ oxid. peptide in OL-14F7}}{\% \text{ oxid. peptide in O-14F7}} \\
 \% \text{ protection} &= \frac{46.20 - 30.04}{46.20} \times 100\% \\
 &= 35\%
 \end{aligned}$$

Thus, up on ligand binding, 35 % of the amino acid side chains within the peptide are protected from oxidation by hydroxyl radicals.

Appendix 6

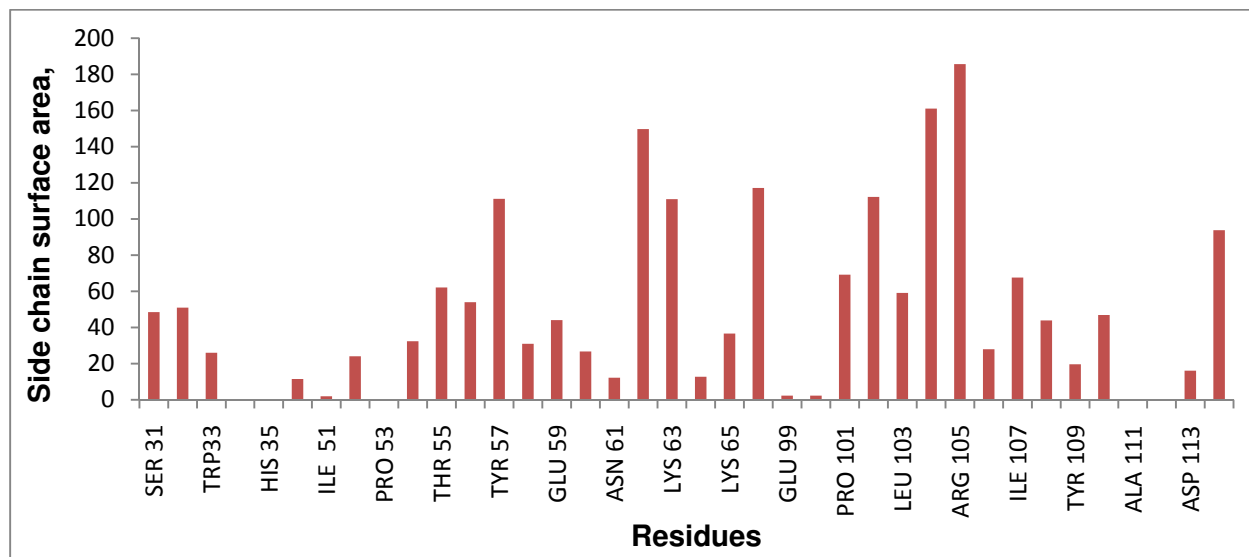


Figure A3: Side chain surface area of VH-CDR residues.

REFERENCES

1. Jemal, A., et al., *Cancer Statistics*. CA Cancer J Clin, 2009. **59**(4): p. 225-249.
2. Derycke, L., et al., *Molecular Targets of Growth, Differentiation, Tissue Integrity, and Ectopic Cell Death in Cancer Cells*. Cancer Biotherapy & Radiopharmaceuticals, 2005. **20**(6): p. 579-588.
3. Raff, M.C., *T and B Lymphocytes and Immune Responses*. Nature, 1973. **242**(5392): p. 19-23.
4. Gallup, D.G. and O.E. Talledo, *Benign and Malignant Tumors*. Clinical Obstetrics and Gynecology, 1987. **30**(3): p. 662-670.
5. Yeh, M.Y., et al., *Cell surface antigens of human melanoma identified by monoclonal antibody*. Proceedings of the National Academy of Sciences of the United States of America, 1979. **76**(6): p. 2927-2931.
6. Ueda, R., et al., *Cell surface antigens of human renal cancer defined by mouse monoclonal antibodies: identification of tissue-specific kidney glycoproteins*. Proceedings of the National Academy of Sciences of the United States of America, 1981. **78**(8): p. 5122-5126.
7. Birklé, S., et al., *Role of tumor-associated gangliosides in cancer progression*. Biochimie. **85**(3-4): p. 455-463.
8. Parish, C.R., *Cancer immunotherapy: The past, the present and the future*. Immunology and Cell Biology, 2003. **81**(2): p. 106-113.
9. Waldmann, T.A., *Monoclonal-antibodies in diagnosis and therapy*. Science, 1991. **252**(5013): p. 1657-1662.
10. Zanetti, M. and J.D. Capra, *The Antibodies*. 1 ed. Vol. 7. 2002, New York: Taylor and Francis Inc. p. 228.
11. Davies, D.R. and S. Chacko, *Antibody structure*. Accounts of Chemical Research, 1993. **26**(8): p. 421-427.
12. Alzari, P.M., M.B. Lascombe, and R.J. Poljak, *3-Dimensional structure of antibodies*. Annual Review of Immunology, 1988. **6**: p. 555-580.
13. Lesk, A.M., *Introduction to Protein Architecture: The Structural Biology of Proteins* 1st ed. 2001, New York: Oxford University Press Inc. p. 229-267.
14. Kerwin, S.M., *ChemBioOffice Ultra 2010 Suite*. Journal of the American Chemical Society, 2010. **132**(7): p. 2466-2467.
15. Kohler, G. and C. Milstein, *Continuous cultures of fused cells secreting antibody of predefined specificity*. Nature, 1975. **256**(5517): p. 495-497.
16. Kawashima, I., et al., *Generation of monoclonal antibodies specific for ganglioside lactones: Evidence of the expression of lactone on human melanoma cells*. International Journal of Cancer, 1994. **58**(2): p. 263-268.
17. VÁZQUEZ, A.M., et al., *Generation of a Murine Monoclonal Antibody Specific for N-Glycolylneuraminic Acid-Containing Gangliosides That Also Recognizes Sulfated Glycolipids*. Hybridoma, 1995. **14**(6): p. 551-556.
18. Kotani, M., et al., *Generation of one set of monoclonal antibodies specific for a-pathway ganglioside series gangliosides*. Biochimica et Biophysica Acta (BBA) - General Subjects, 1992. **1117**(1): p. 97-103.
19. Fishman, P. and R. Brady, *Biosynthesis and function of gangliosides*. Science, 1976. **194**(4268): p. 906-915.
20. Hakomori, S., *Glycosphingolipids in Cellular Interaction, Differentiation, and Oncogenesis*. Annual Review of Biochemistry, 2003. **50**(1): p. 733-764.
21. Roland Schauer, et al., *Biology of the sialic acids*. 1 ed. Biochemistry and Role of Sialic acids, ed. A. Rosenberg. 1995, New York: Springer.

22. Irie, A., et al., *The Molecular Basis for the Absence of N-Glycolylneuraminic Acid in Humans*. Journal of Biological Chemistry, 1998. **273**(25): p. 15866-15871.
23. Marquina, G., et al., *Gangliosides Expressed in Human Breast Cancer*. Cancer Res, 1996. **56**(22): p. 5165-5171.
24. Krengel, U., et al., *Structure and Molecular Interactions of a Unique Antitumor Antibody Specific for N-Glycolyl GM3*. Journal of Biological Chemistry, 2004. **279**(7): p. 5597-5603.
25. Colman, P.M., *Structure of Antibody-Antigen Complexes: Implications for Immune Recognition*, in *Advances in Immunology*, J.D. Frank, Editor. 1988, Academic Press. p. 99-132.
26. Fischmann, T.O., et al., *Crystallographic refinement of the three-dimensional structure of the FabD1.3-lysozyme complex at 2.5-Å resolution*. Vol. 266. 1991. p. 12915-20.
27. Brünger, A.T., et al., *2.9 Å resolution structure of an anti-dinitrophenyl-spin-label monoclonal antibody Fab fragment with bound hapten*. Journal of Molecular Biology, 1991. **221**(1): p. 239-256.
28. Sheriff, S., et al., *Three-dimensional structure of an antibody-antigen complex*. Proceedings of the National Academy of Sciences of the United States of America, 1987. **84**(22): p. 8075-8079.
29. Amit, A., et al., *Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution*. Science, 1986. **233**(4765): p. 747-753.
30. Carr, A., et al., *A Mouse IgG1 Monoclonal Antibody Specific for N-Glycolyl GM3 Ganglioside Recognized Breast and Melanoma Tumors*. Hybridoma, 2000. **19**(3): p. 241-247.
31. Carr, A., et al., *In Vivo and In Vitro Anti-Tumor Effect of 14F7 Monoclonal Antibody*. Hybridoma and Hybridomics, 2002. **21**(6): p. 463-468.
32. Rojas, G., et al., *Light-chain shuffling results in successful phage display selection of functional prokaryotic-expressed antibody fragments to N-glycolyl GM3 ganglioside*. Journal of Immunological Methods, 2004. **293**(1-2): p. 71-83.
33. Barford, D., A.K. Das, and M.-P. Egloff, *The structure and mechanism of protein phosphatases: Insights into Catalysis and Regulation*. Annual Review of Biophysics and Biomolecular Structure, 1998. **27**(1): p. 133-164.
34. Kuhn, P., et al., *The genesis of high-throughput structure-based drug discovery using protein crystallography*. Current Opinion in Chemical Biology, 2002. **6**(5): p. 704-710.
35. Clarkson, J. and I.D. Campbell, *Studies of protein-ligand interactions by NMR*. Biochem. Soc. Trans., 2003. **31**(Pt 5): p. 1006-1009.
36. Yates, J.R., *Mass spectrometry: from genomics to proteomics*. Trends in Genetics, 2000. **16**(1): p. 5-8.
37. Fenn, J., et al., *Electrospray ionization for mass spectrometry of large biomolecules*. Science, 1989. **246**(4926): p. 64-71.
38. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons*. Analytical Chemistry, 1988. **60**(20): p. 2299-2301.
39. Yinon, J., *Forensic Applications of Mass spectrometry*. 1 ed. Modern mass spectrometry. 1995, Florida: CRC Press, Inc.
40. Vekey, K., A. Telekes, and A. Vertes, *Medical Applications of Mass Spectrometry* 1st ed. 2007, Oxford: Elsevier Science.
41. Tanaka, K., et al., *Protein and polymer analyses up to 100 000 by laser ionization time-of-flight mass spectrometry*. Rapid Communications in Mass Spectrometry, 1988. **2**(8): p. 151-153.
42. Tanaka, K., *The Origin of Macromolecule Ionization by Laser Irradiation (Nobel Lecture)*. Angewandte Chemie International Edition, 2003. **42**(33): p. 3860-3870.
43. Macfarlane, R. and D. Torgerson, *Californium-252 plasma desorption mass spectroscopy*. Science, 1976. **191**(4230): p. 920-925.

44. Barber, M., B.N. Green, and K.R. Jennings, *The analysis of small proteins in the molecular weight range 10-24 kDa by magnetic sector mass spectrometry*. Rapid Communications in Mass Spectrometry, 1987. **1**(5): p. 80-83.
45. Imrie, D.C., J.M. Pentney, and J.S. Cottrell, *A Faraday cup detector for high-mass ions in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry*. Rapid Communications in Mass Spectrometry, 1995. **9**(13): p. 1293-1296.
46. Beavis, R.C., T. Chaudhary, and B.T. Chait, *alpha-Cyano-4-hydroxycinnamic acid as a matrix for matrix assisted laser desorption mass spectrometry*. Organic Mass Spectrometry, 1992. **27**(2): p. 156-158.
47. Beavis, R.C., B.T. Chait, and H.M. Fales, *Cinnamic acid derivatives as matrices for ultraviolet laser desorption mass spectrometry of proteins*. Rapid Communications in Mass Spectrometry, 1989. **3**(12): p. 432-435.
48. Mank, M., B. Stahl, and G. Boehm, *2,5-Dihydroxybenzoic Acid Butylamine and Other Ionic Liquid Matrixes for Enhanced MALDI-MS Analysis of Biomolecules*. Analytical Chemistry, 2004. **76**(10): p. 2938-2950.
49. Loo, J.A., H.R. Udseth, and R.D. Smith, *Peptide and protein analysis by electrospray ionization-mass spectrometry and capillary electrophoresis-mass spectrometry*. Analytical Biochemistry, 1989. **179**(2): p. 404-412.
50. Watson, J.T. and O.D. Sparkman, *Introduction to Mass Spectrometry: Instrumentation, Applications, and Strategies for Data Interpretation*. 4th ed. 2007, Chichester: John Wiley & Sons Ltd. p. 179-184.
51. Marshall, A.G., C.L. Hendrickson, and G.S. Jackson, *Fourier transform ion cyclotron resonance mass spectrometry: A primer*. Mass Spectrometry Reviews, 1998. **17**(1): p. 1-35.
52. Hu, Q., et al., *The Orbitrap: a new mass spectrometer*. Journal of Mass Spectrometry, 2005. **40**(4): p. 430-443.
53. Guilhaus, M., *Principles and instrumentation in time-of-flight mass spectrometry. Physical and instrumental concepts*. Journal of Mass Spectrometry, 1995. **30**(11): p. 1519-1532.
54. Cotter Robert, J., *The new Time-of-Flight Mass Spectrometer*. Analytical Chemistry, 1999. **71**: p. 445A-451A.
55. Sparkman, O.D., *Mass Spectrometry PittCon® 2006*. Journal of the American Society for Mass Spectrometry, 2006. **17**(6): p. 873-884.
56. McLuckey, S.A. and J.M. Wells, *Mass Analysis at the Advent of the 21st Century*. Chemical Reviews, 2001. **101**(2): p. 571-606.
57. Tang, X.J., P. Thibault, and R.K. Boyd, *Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation*. Analytical Chemistry, 1993. **65**(20): p. 2824-2834.
58. De Wit, J.S.M., et al., *Direct coupling of open-tubular liquid chromatography with mass spectrometry*. Analytical Chemistry, 1987. **59**(19): p. 2400-2404.
59. Carrier, A. and J. Parent, *Modification of a commercial electrospray nebulizer for operation in a liquid chromatography/mass spectrometry system at flow rates in the low $\mu\text{L}/\text{min}$ range*. Rapid Communications in Mass Spectrometry, 2001. **15**(18): p. 1681-1684.
60. Minakuchi, H., et al., *Octadecylsilylated Porous Silica Rods as Separation Media for Reversed-Phase Liquid Chromatography*. Analytical Chemistry, 1996. **68**(19): p. 3498-3501.
61. Henzel, W.J., et al., *Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases*. Proceedings of the National Academy of Sciences of the United States of America, 1993. **90**(11): p. 5011-5015.
62. Henzel, W.J., C. Watanabe, and J.T. Stults, *Protein identification: the origins of peptide mass fingerprinting*. Journal of the American Society for Mass Spectrometry, 2003. **14**(9): p. 931-942.

63. Walker, J.M., *The Protein Protocols Handbook*. 2 ed. 2002, Totowa, NJ Humana Press Inc.
64. Shevchenko, A., et al., *Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels*. Analytical Chemistry, 1996. **68**(5): p. 850-858.
65. Riviere, L.R. and P. Tempst, *Current Protocols in Protein Science*. June, 1995, Brooklyn: John Wiley & Sons, Inc.
66. Fenyő, D., *Identifying the proteome: software tools*. Current Opinion in Biotechnology, 2000. **11**(4): p. 391-395.
67. Gasteiger, E., et al., *ExpASY: the proteomics server for in-depth protein knowledge and analysis*. Nucl. Acids Res., 2003. **31**(13): p. 3784-3788.
68. Clauser, K.R., P. Baker, and A.L. Burlingame, *Role of Accurate Mass Measurement (± 10 ppm) in Protein Identification Strategies Employing MS or MS/MS and Database Searching*. Analytical Chemistry, 1999. **71**(14): p. 2871-2882.
69. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data*. Electrophoresis, 1999. **20**(18): p. 3551-3567.
70. Roepstorff, P. and J. Fohlman, *Proposal for a common nomenclature for sequence ions in mass spectra of peptides*. Biomedical mass spectrometry, 1984. **11**(11).
71. Hernandez, P., M. Müller, and R.D. Appel, *Automated protein identification by tandem mass spectrometry: Issues and strategies*. Mass Spectrometry Reviews, 2006. **25**(2): p. 235-254.
72. Nesvizhskii, A.I., *Protein Identification by Tandem Mass Spectrometry and Sequence Database Searching*. 2006. p. 87-119.
73. Ma, B., et al., *PEAKS: powerful software for peptide sequencing by tandem mass spectrometry*. Rapid Communications in Mass Spectrometry, 2003. **17**(20): p. 2337-2342.
74. Bern, M., Y. Cai, and D. Goldberg, *Lookup Peaks: A Hybrid of de Novo Sequencing and Database Search for Protein Identification by Tandem Mass Spectrometry*. Analytical Chemistry, 2007. **79**(4): p. 1393-1400.
75. Ghaemmaghami, S., M.C. Fitzgerald, and T.G. Oas, *A quantitative, high-throughput screen for protein stability*. Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(15): p. 8296-8301.
76. Zhu, M.M., R. Chitta, and M.L. Gross, *PLIMSTEX: a novel mass spectrometric method for the quantification of protein-ligand interactions in solution*. International Journal of Mass Spectrometry, 2005. **240**(3): p. 213-220.
77. Gau, B.C., et al., *Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding*. Analytical Chemistry, 2009. **81**(16): p. 6563-6571.
78. Schickor, P. and H. Heumann, *Hydroxyl Radical Footprinting*. 1994. p. 21-32.
79. Tullius, T.D. and J.A. Greenbaum, *Mapping nucleic acid structure by hydroxyl radical cleavage*. Current Opinion in Chemical Biology, 2005. **9**(2): p. 127-134.
80. Heyduk, T., et al., *Determinants of RNA polymerase alpha subunit for interaction with beta, beta', and sigma subunits: hydroxyl-radical protein footprinting*. Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(19): p. 10162-10166.
81. Heyduk, E. and T. Heyduk, *Mapping Protein Domains Involved in Macromolecular Interactions: A Novel Protein Footprinting Approach*. Biochemistry, 1994. **33**(32): p. 9643-9650.
82. Wong, J.W.H., S.D. Maleknia, and K.M. Downard, *Hydroxyl radical probe of the calmodulin-melittin complex interface by electrospray ionization mass spectrometry*. Journal of the American Society for Mass Spectrometry, 2005. **16**(2): p. 225-233.
83. Xu, G. and M.R. Chance, *Radiolytic Modification and Reactivity of Amino Acid Residues Serving as Structural Probes for Protein Footprinting*. Analytical Chemistry, 2005. **77**(14): p. 4549-4555.

84. Charvátová, O., et al., *Quantifying Protein Interface Footprinting by Hydroxyl Radical Oxidation and Molecular Dynamics Simulation: Application to Galectin-1*. Journal of the American Society for Mass Spectrometry, 2008. **19**(11): p. 1692-1705.
85. Borch, J., T.J.D. Jørgensen, and P. Roepstorff, *Mass spectrometric analysis of protein interactions*. Current Opinion in Chemical Biology, 2005. **9**(5): p. 509-516.
86. Suckau, D., M. Mak, and M. Przybylski, *Protein surface topology-probing by selective chemical modification and mass spectrometric peptide mapping*. Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(12): p. 5630-5634.
87. Hambly, D.M. and M.L. Gross, *Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale*. Journal of the American Society for Mass Spectrometry, 2005. **16**(12): p. 2057-2063.
88. Sclavi, B., et al., *Time-resolved synchrotron X-ray "footprinting", a new approach to the study of nucleic acid structure and function: application to protein-DNA interactions and RNA folding*. Journal of Molecular Biology, 1997. **266**(1): p. 144-159.
89. Chance, M.R., et al., *Examining the conformational dynamics of macromolecules with time-resolved synchrotron X-ray footprinting*. Structure, 1997. **5**(7): p. 865-869.
90. Halliwell, B. and J.M.C. Gutteridge, *Biologically relevant metal ion-dependent hydroxyl radical generation An update*. FEBS Letters, 1992. **307**(1): p. 108-112.
91. Dainton, F.S., *The Primary Quantum Yield in the Photolysis of Hydrogen Peroxide at 3130 Å. and the Primary Radical Yield in the X- and γ-Radiolysis of Water*. Journal of the American Chemical Society, 1956. **78**(7): p. 1278-1279.
92. Weeks, J.L. and M.S. Matheson, *The Primary Quantum Yield of Hydrogen Peroxide Decomposition*. Journal of the American Chemical Society, 1956. **78**(7): p. 1273-1278.
93. Volman, D.H. and J.C. Chen, *The Photochemical Decomposition of Hydrogen Peroxide in Aqueous Solutions of Allyl Alcohol at 2537 Å.1*. Journal of the American Chemical Society, 1959. **81**(16): p. 4141-4144.
94. Aye, T.T., T.Y. Low, and S.K. Sze, *Nanosecond Laser-Induced Photochemical Oxidation Method for Protein Surface Mapping with Mass Spectrometry*. Analytical Chemistry, 2005. **77**(18): p. 5814-5822.
95. Hambly, D.M. and M.L. Gross, *Cold Chemical Oxidation of Proteins*. Analytical Chemistry, 2009. **81**(17): p. 7235-7242.
96. Stadtman, E.R. and B.S. Berlett, *Fenton chemistry. Amino acid oxidation*. Journal of Biological Chemistry, 1991. **266**(26): p. 17201-17211.
97. Xu, G. and M.R. Chance, *Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics*. Chemical Reviews, 2007. **107**(8): p. 3514-3543.
98. Sharp, J.S., et al., *Photochemical surface mapping of C14S-Sml1p for constrained computational modeling of protein structure*. Analytical Biochemistry, 2005. **340**(2): p. 201-212.
99. Maleknia, S.D. and K.M. Downard, *Radical approaches to probe protein structure, folding, and interactions by mass spectrometry*. Mass Spectrometry Reviews, 2001. **20**(6): p. 388-401.
100. Maleknia, S.D., M. Brenowitz, and M.R. Chance, *Millisecond Radiolytic Modification of Peptides by Synchrotron X-rays Identified by Mass Spectrometry*. Analytical Chemistry, 1999. **71**(18): p. 3965-3973.
101. Elliott, W.H. and D.C. Elliott, *Biochemistry and Molecular Biology*. 4th ed. 2009, New York: Oxford University Press Inc. p. 567.
102. Xu, G. and M.R. Chance, *Radiolytic Modification of Acidic Amino Acid Residues in Peptides: Probes for Examining Protein-Protein Interactions*. Analytical Chemistry, 2004. **76**(5): p. 1213-1221.

103. Xu, G., K. Takamoto, and M.R. Chance, *Radiolytic Modification of Basic Amino Acid Residues in Peptides: Probes for Examining Protein–Protein Interactions*. Analytical Chemistry, 2003. **75**(24): p. 6995-7007.
104. Xu, G. and M.R. Chance, *Radiolytic Modification of Sulfur-Containing Amino Acid Residues in Model Peptides: Fundamental Studies for Protein Footprinting*. Analytical Chemistry, 2005. **77**(8): p. 2437-2449.
105. Guan, J.-Q. and M.R. Chance, *Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectrometry*. Trends in Biochemical Sciences, 2005. **30**(10): p. 583-592.
106. Morris, G.M., et al., *Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4*. Journal of Computer-Aided Molecular Design, 1996. **10**(4): p. 293-304.
107. Goodsell, D.S. and A.J. Olson, *Automated docking of substrates to proteins by simulated annealing*. Proteins: Structure, Function, and Genetics, 1990. **8**(3): p. 195-202.
108. Morris, G.M., et al., *Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function*. Journal of Computational Chemistry, 1998. **19**(14): p. 1639-1662.
109. Hussain, R., *Review: ELISA: Theory and Practice*. Epidemiology and Infection, 1995. **115**(3): p. 624.
110. Mirzaei, H. and F. Regnier, *Enhancing Electrospray Ionization Efficiency of Peptides by Derivatization*. Analytical Chemistry, 2006. **78**(12): p. 4175-4183.
111. Synder, A.P., *Interpreting Protein Mass Spectra: A Comprehensive Resource* 2000: An American Chemical Society Publication.
112. Maleknia, S.D., J.G. Kiselar, and K.M. Downard, *Hydroxyl radical probe of the surface of lysozyme by synchrotron radiolysis and mass spectrometry*. Rapid Communications in Mass Spectrometry, 2002. **16**(1): p. 53-61.
113. Hubbard, S.J. and J.M. Thornton, *NACCESS*. University College London, London, 1993.
114. Sharp, J.S., J.M. Becker, and R.L. Hettich, *Protein surface mapping by chemical oxidation: Structural analysis by mass spectrometry*. Analytical Biochemistry, 2003. **313**(2): p. 216-225.
115. Kiselar, J.G., et al., *Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry*. International Journal of Radiation Biology, 2002. **78**(2): p. 101-114.

