STRUCTURAL CHARACTERIZATION OF IMMUNOGLOBULIN G ANTIBODIES
WITH
LC-MS BASED APPROACHES

A dissertation presented

by

Yi Wang

to
The Department of Chemistry and Chemical Biology

In partial fulfillment of the requirements for the degree of
Doctor of Philosophy

in the field of
Chemistry

Northeastern University
Boston, Massachusetts
April, 2012
STRUCTURAL CHARACTERIZATION OF IMMUNOGLOBULIN G ANTIBODIES
WITH
LC-MS BASED APPROACHES

by

Yi Wang

ABSTRACT OF DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate School of Northeastern University
April, 2012
ABSTRACT

Therapeutic monoclonal antibodies (mAb) have been one of the fastest-growing areas in the biotechnological and pharmaceutical industry due to their high specificity as anticancer drugs with few side effects. Comprehensive characterization of such mAbs is critical to the biotech industry. Particularly, detection of functionality-sensitive posttranslational modifications (PTMs) and chemical modifications in recombinant monoclonal antibodies can have significant quality implications. Liquid chromatography-mass spectrometry (LC-MS) is currently the most sophisticated and powerful tool for protein characterization due to its superior resolution, sensitivity, and accuracy. This dissertation focuses on the structural characterization of antibodies (IgGs) using liquid chromatography - mass spectrometry (LC-MS) based methods.

Chapter 1 reviews the importance of mAb drugs, structure of mAbs, common chemical modifications and the current mass spectrometry methods used in their structural characterization.

In Chapter 2, the disulfide linkages present in three therapeutic mAbs, anti-HER2 mAb, anti-CD11a mAb and GLP-1 with IgG4-Fc fusion protein, were completely characterized by our LC-MS methods with both electron-transfer dissociation (ETD) and collision induced dissociation (CID) fragmentation. Scrambled disulfides in heat-stressed samples were also identified easily and confidently by this analytical method. We then compared the disulfide scrambling patterns of these three different mAbs and how the disulfide scrambling process may initiate.

In Chapter 3, a new chip-based LC-MS method for direct analysis of N-glycosylation present in intact mAbs was evaluated. Immobilized PNGase F and porous graphitized carbon
(PGC) column are integrated together in this mAb-Glyco chip for on-line release of the glycans, separation and mass spectrometric analysis. For a mAb with only one glycosylation site, this automation and absence of manual sample pre-treatment simplifies the procedure and reduces analysis time significantly. The glycan distribution of anti-HER2 mAb generated from mAb-Glyco chip was compared to the results from two other chip-based MS methods: PGC chip MS for offline released glycans analysis and a C18 chip MS for glycopeptide analysis. We then evaluated day-to-day and run-to-run reproducibilities of these methods and then analyzed different types of mAbs, including innovator and biosimilar products with different formulations.

In Chapter 4, a range of LC-MS methods have been used for comprehensive characterization of the primary structure, disulfides, glycan structures, and other modifications of innovator and biosimilar anti-HER2 mAbs. The LC-MS methods include identification of glycopeptides and disulfide linkages by the combination of CID and ETD fragmentation and profiling glycans by an enzyme immobilized chip-based LC-MS approach. The extent of N-terminal pyroglutamic acid cyclization, heavy chain C-terminal lysine processing, oxidation, deamidation and isomerization were compared between innovator and biosimilar anti-HER2 mAbs.

In Chapter 5, the glycan distributions of IgGs in blood samples of patients at varying stages of AIDS were studied using two glycopeptide analysis methods and two glycan analysis methods. For glycopeptide analysis, tryptic peptides were analyzed by C18 reversed phase high-performance liquid chromatography - Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) while Lys-C digested peptides were analyzed by C18 reversed phase high-performance microfluidic chip liquid chromatography quadruple time-of-flight mass spectrometry (Q-TOF MS). The PNGase F offline released glycans were identified by
microfluidic-based porous graphitized carbon (PGC) chip with Q-TOF MS or intact IgGs were directly analyzed by a mAb-Glyco chip. Glycopeptide analysis can provide site-specific information and therefore the glycan distribution of subclasses IgGs can be determined. Both glycopeptide analysis platforms provided similar IgG subclass-specific profiles and with a similar amount of measurement variation. The pros and cons of the different glycan analysis methods for the generation of total glycan profiles will be discussed in this chapter.
ACKNOWLEDGEMENTS

I must first thank my advisor, Dr. William S. Hancock, and my supervisor, Dr. Shiaw-lin Wu, for providing me the great opportunity to work on these projects. Your guidance and deep insights into mass spectrometry science are great values in my life; your support, encouragement and patience make me grateful forever. I would like also thank Professor Barry Karger for proving an excellent research environment at the Barnett Institute.

To my committee members: Dr. Penny J. Beuning, Dr. Michael P. Pollastri and Dr. Zhaohui (Sunny) Zhou, thanks for serving on my committee and for giving me valuable advice.

I gratefully thank former and present members from Dr. Hancock and Dr. Karger’s group: Dr. Marina Hincapie, Dr. Xiaoyang Zheng, Dr. Haitao Jiang, Dr. Majlinda Kullolli, Dr. Qiaozhen Lu, Zhi Zeng, Xiaomei He, Suli Liu, Yue Zhang, Fangfei Yan and Fan Zhang; Dr. Sangwon Cha, Dr. Shujia Dai, Dr. Quanzhou Luo, Dr. Dongdong Wang, Zhenke Liu, Chen Li, Siyang Li and Wenqin Ni for their support and encouragement. People come and go, but the spirits, the wisdom, and the friendships remain.

My special thanks to Jean Harris and Andrew Bean from Department of Chemistry and Chemical Biology, Jeffrey Kasilman, Jana Volf, Felicia Hopkins and Nancy Carbone from Barnett Institute for their kind help during my study in Northeastern University.

I gratefully thank Dr. Tyler Carlage and Suli Liu for reviewing my dissertation draft.

Thank you to my husband for your understanding and support throughout. To my son, Eric, your brilliant smile and funny words ease my troubles and make me happy every day.

Finally, I would like to dedicate this dissertation especially to my grandparents and my mother. Without your selfless love, tolerance, and support, I would not be where I am today.
TABLE OF CONTENTS

ABSTRACT .................................................................................................................................................. 2

ACKNOWLEDGEMENTS .......................................................................................................................... 6

TABLE OF CONTENTS .......................................................................................................................... 7

LIST OF FIGURES ................................................................................................................................... 13

LIST OF TABLES ....................................................................................................................................... 17

LIST OF ABBREVIATIONS ....................................................................................................................... 19

Chapter 1 .................................................................................................................................................. 22

OVERVIEW OF MASS SPECTROMETRIC METHODS FOR STRUCTURAL CHARACTERIZATION OF THERAPEUTIC MONOCLONAL ANTIBODIES

1.1 Abstract ............................................................................................................................................. 23

1.2 Antibodies ........................................................................................................................................... 23

  1.2.1 Human immunoglobulin G antibodies ......................................................................................... 23

  1.2.2 History of therapeutic monoclonal antibodies (mAbs) ............................................................... 27

  1.2.3 Types of therapeutic mAbs .......................................................................................................... 30

  1.2.4 Advantages and limitations of therapeutic mAbs ......................................................................... 31

  1.2.5 Common chemical modifications in therapeutic mAbs ............................................................... 32

  1.2.6 Importance of structural characterization of therapeutic mAbs.................................................... 33

  1.2.7 Anti-Her2 mAb and treatment of breast cancer ............................................................................. 34

1.3 Mass Spectrometry ............................................................................................................................ 36

  1.3.1 Ionization methods ....................................................................................................................... 36

    1.3.1.1 Electrospray ionization (ESI) ............................................................................................... 37
1.3.1.2 Matrix-assisted laser desorption ionization (MALDI) ..................................38
1.3.2 Mass analyzers ........................................................................................................38
  1.3.2.1 Quadrupole ........................................................................................................39
  1.3.2.2 Paul ion trap .......................................................................................................40
  1.3.2.3 Linear Ion Trap (LIT) .........................................................................................40
  1.3.2.4 Time-of-Flight (TOF) .........................................................................................40
  1.3.2.5 Orbitrap .............................................................................................................41
  1.3.2.6 Fourier Transform Ion Cyclotron Resonance (FTICR) ....................................41
1.3.3 Hybrid mass spectrometer instruments .................................................................42
  1.3.3.1 LTQ-Orbitrap ...................................................................................................43
  1.3.3.2 LTQ-FT .............................................................................................................43
  1.3.3.3 Q-TOF ...............................................................................................................43
1.3.4 Tandem mass spectrometry ....................................................................................44
  1.3.4.1 Tandem Mass Spectrometry in Space or in Time ...............................................45
  1.3.4.2 Fragmentation in Tandem Mass Spectrometry ..................................................45
1.3.5 Liquid Chromatography-Mass Spectrometry (LC-MS) ........................................46
1.4 Mass Spectrometry (MS)-based Strategies for Structural Characterization of Therapeutic MAbs .................................................................47
  1.4.1 Top-down approach ..............................................................................................47
  1.4.2 Middle-up approach ............................................................................................49
  1.4.3 Bottom-up approach ............................................................................................49
1.5 Structural Characterization of Therapeutic MAbs by Bottom-up Mass Spectrometry .......51
  1.5.1 Primary Sequence analysis ...................................................................................51
1.5.2 Disulfide-linkage determination .......................................................... 52
1.5.3 Glycosylation analysis ........................................................................ 55
  1.5.3.1 Glycopeptide analysis ................................................................. 55
  1.5.3.2 Glycan analysis ......................................................................... 57
1.5.4 Characterization of common chemical modifications characterization .......... 58
1.5.5 Quantitation ....................................................................................... 60
  1.5.5.1 Absolute quantitation (AQUA) ...................................................... 60
  1.5.5.2 Relative quantitation .................................................................... 61
1.6 Conclusions ........................................................................................... 62
1.7 References ............................................................................................. 63

Chapter 2 ..................................................................................................... 74

CHARACTERIZATION AND COMPARISON OF DISULFIDE LINKAGES AND SCRAMBLING PATTERNS IN THERAPEUTIC MONOCLONAL ANTIBODIES: USING LC-MS WITH ELECTRON TRANSFER DISSOCIATION

2.1 Abstract ................................................................................................. 75
2.2 Introduction ............................................................................................ 76
2.3 Experimental Section ............................................................................ 78
  2.3.1 Materials ......................................................................................... 78
  2.3.2 Heat-Stressed Conditions ............................................................... 79
  2.3.3 Enzymatic Digestion ...................................................................... 79
  2.3.4 LC-MS .......................................................................................... 80
2.4 Results and Discussion .......................................................................... 80
  2.4.1 Digestion Strategy .......................................................................... 80
  2.4.2 LC-MS with CID and ETD ............................................................. 82
2.4.3 Disulfide Assignment ................................................................. 82
2.4.4 Heat-Stressed Study ................................................................. 84
2.4.5 Scrambled Disulfides ............................................................... 85
2.4.6 Stability for Other Types of mAb ................................................. 88
2.5 Conclusions .................................................................................. 92
2.6 References ..................................................................................... 93
2.7 Supporting Information ................................................................. 97

Chapter 3 ............................................................................................... 101

ANALYSIS OF N-GLYCOSYLATION IN THERAPEUTIC MONOCLONAL
ANTIBODIES BY LC-MS WITH A NEW CHIP MS METHOD

3.1 Abstract ......................................................................................... 102
3.2 Introduction .................................................................................... 103
3.3 Experimental Section ..................................................................... 105
  3.3.1 Materials .................................................................................. 105
  3.3.2 Enzymatic Digestion .................................................................. 106
  3.3.3 LC-MS ..................................................................................... 107
  3.3.4 Data analysis ............................................................................. 109
3.4 Results and Discussion .................................................................. 109
  3.4.1 Analysis of glycans of anti-HER2 mAb ....................................... 109
  3.4.2 Comparison of mAb-Glyco-chip MS with other LC-MS methods ......... 112
  3.4.3 Reproducibility of mAb-Glyco-chip MS analysis of mAb ................. 114
  3.4.4 Analysis and comparison of innovator and biosimilar samples ........... 116
  3.4.5 Analysis of different types of monoclonal antibodies ....................... 117
3.5 Conclusions .................................................................................... 118
### Chapter 4

**STRUCTURAL CHARACTERIZATION OF INNOVATOR AND BIOSIMILAR THERAPEUTIC MONOCLONAL ANTIBODIES BY LC-MS BASED APPROACHES**

4.1 Abstract

4.2 Introduction

4.3 Experimental Section

4.3.1 Materials

4.3.2 SDS-PAGE

4.3.3 Enzymatic digestion

4.3.4 LC-MS

4.4 Results and Discussion

4.4.1 Structure of anti-HER2 mAb

4.4.2 SDS-PAGE results

4.4.3 Primary structure identification

4.4.4 Glycopeptides Identification

4.4.5 Disulfide Linkages Identification

4.4.6 Chemical Modifications

4.4.6.1 N-terminal pyroglutamic acid cyclization

4.4.6.2 Heavy chain C-terminal lysine processing

4.4.6.3 Oxidation

4.4.6.4 Deamidation

4.4.6.5 Aspartic acid Isomerization

4.5 Conclusions
Chapter 5 ............................................................................................................................... 154

N-GLYCOSYLATION ANALYSIS OF IMMUNOGLOBULIN GS FROM HIV-1 PATIENTS BY DIFFERENT ANALYSIS PLATFORMS WITH LC-MS

5.1 Abstract .................................................................................................................................. 155
5.2 Introduction ............................................................................................................................... 156
5.3 Experimental Section ............................................................................................................... 159
  5.3.1 Materials .............................................................................................................................. 159
  5.3.2 SDS-PAGE .......................................................................................................................... 159
  5.3.3 Enzymatic Digestion .......................................................................................................... 160
  5.3.4 LC-MS Analysis .................................................................................................................. 160
  5.3.5 Data analysis ....................................................................................................................... 162
5.4 Results and Discussion ............................................................................................................. 163
  5.4.1 SDS-PAGE Result of IgG isolated from HIV-1 patients ...................................................... 163
  5.4.2 LC-MS method development ............................................................................................. 165
    5.4.2.1 Glycopeptide analysis .................................................................................................... 165
    5.4.2.2 Cleaved Glican Analysis .............................................................................................. 169
    5.4.2.3 Comparison of results from 4 analytical approaches .................................................. 170
  5.4.3 Comparison of the IgG glycosylation profiles obtained for four different patients with HIV-1 infection by four LC-MS platforms .................................................................................. 172
5.5 Conclusions ............................................................................................................................. 173
5.6 References ............................................................................................................................... 186

FUTURE WORK ........................................................................................................................ 190
LIST OF FIGURES

Figure 1-1. Schematic structure of the human IgG molecule..........................................................25
Figure 1-2. Schematic structures of four human IgG subclasses.........................................................28
Figure 1-3. Structure of anti-HER2 mAb............................................................................................36
Figure 1-4. Peptide fragments generated in tandem mass spectrometry............................................46
Figure 1-5. Analysis of a disulfide-linked peptide from anti-HER2 mAb by the combination of CID and ETD fragmentation.................................................................54
Figure 1-6. CID and ETD spectra of glycopeptide (with G1) from anti-HER2 mAb .......................56
Figure 1-7. LC-MS analysis of the T2 and the oxidized T2 peptides from the tryptic digest of rhGH (Nutropin AQ)...........................................................................................................59

Figure 2-1. Disulfide structures of anti-HER2 mAb (A) and GLP-1 fusion mAb (B) ......................77
Figure 2-2. Using CID-MS2 (A), ETD-MS2 (B), and CID-MS3 (C) to analyze a disulfide-linked peptide (anti-HER2).....................................................................................................................83
Figure 2-3. Heat-stressed study for mAb (anti-HER2) using SDS-PAGE separation with nonreduced (A) and reduced (B) conditions..................................................................................84
Figure 2-4. Using CID-MS^2 (A), ETD-MS2 (B), and CID-MS3 (C) to analyze a disulfide scrambled peptide (anti-HER2). ................................................................................................................86
Figure 2-5. Heat-stressed study for anti-CD11a (A) and GLP-1 with IgG4-Fc fusion mAb (B) using SDS-PAGE separation without DTT reduction.................................................................90
Figure 2-6. CID and ETD analysis of an inter- to intrachain disulfide scrambled peptide in GLP-1 with IgG4-Fc fusion mAb.........................................................................................................91
Figure 2-S1. ETD analysis of scrambled disulfides at N-terminal and C-terminal region of the light chain of anti-HER2 (band A2 from Figure 2-3A)................................. 97

Figure 2-S2. ETD analysis of scrambled disulfides at Fab and Fc region of the heavy chain of anti-HER2 (band A1a from Figure 2-3A)......................................................... 97

Figure 2-S3. Intra-disulfide in CxxC (T19) from Anti-HER2............................................ 98

Figure 2-S4. Intra-disulfide in CxxC (T19) from Anti-CD 11a........................................ 98

Figure 3-1. Structure of IgG1 (A), structure of glycans and glycosylation site at IgG1 Fc region (B) ..................................................................................................................103

Figure 3-2. Extracted ion chromatogram (EIC) of glycans from Ab1 analyzed by mAb-Glyco chip MS. ................................................................................................................111

Figure 3-3. Targeted MS/MS analysis of (A) Man9, m/z 941.85 (2+) and (B) G1+SA, m/z 958.38 (2+). ................................................................................................................113

Figure 3-4. Comparison of glycan distribution of anti-HER2 by three different chip-based Q-TOF MS methods.........................................................................................113

Figure 3-5. Reproducibility of run to run. ........................................................................115

Figure 3-6. Comparison of glycan distribution in innovator (A) and biosimilar (B) anti-HER2 mAb by mAb-Glyco-chip-Q-TOF MS.................................................................116

Figure 3-7. Analysis of different types mAbs in different buffers by mAb-Glyco-chip-Q-TOF MS. .........................................................................................................................117

Figure 4-1. Structure of anti-HER2 mAb.................................................................................129

Figure 4-2. Gel analysis of innovator and biosimilar anti-HER2 mAbs with reduced (left) and non-reduced (right) conditions.........................................................129

Figure 4-3. LC-MS analysis of N-terminal peptide of innovator anti-HER2 mAb light chain
produced by in-gel RCM trypsin digestion. .................................................................131

Figure 4-4. LC-MS analysis of glycopeptide (Lys-C digestion) (G1, m/z=1014.71, 5+) of innovator anti-HER2 mAb. ...........................................................................................................135

Figure 4-5. Glycopeptides distribution comparison between innovator and biosimilar anti-HER2 mAbs. .............................................................................................................136

Figure 4-6. LC-MS analysis of the heavy chain N-terminal peptide and the N-terminal cyclized peptide from tryptic digest of biosimilar anti-HER2. .........................................................138

Figure 4-7. LC-MS analysis of the heavy chain C-terminal peptide and the C-terminal truncated peptide from tryptic digest of biosimilar anti-HER2. .........................................................140

Figure 4-8. LC-MS analysis of the T10 and the oxidized T10 peptide from tryptic digest of biosimilar anti-HER2 heavy chain. ................................................................................................142

Figure 4-9. LC-MS analysis of the biosimilar anti-HER2 tryptic T36 (H) and the T36 (H) peptide with succinimide intermediate formed during asparagine deamidation..........................144

Figure 4-10. LC-MS analysis of the biosimilar anti-HER2 tryptic T12 (H) and the T12 (H) peptide with succinimide intermediate formed during aspartic acid isomerization. ...............146

Figure 5-1. IgG and Glycan structure. .........................................................................................164

Figure 5-2. SDS-PAGE Analysis of IgG isolated from HIV-1 patients. .......................................164

Figure 5-3. Tryptic glycopeptide analysis by LTQ-FT MS. ..........................................................166

Figure 5-4. Lys-C glycopeptide analysis by Chip Q-TOF MS........................................................168

Figure 5-5. Targeted MS/MS of m/z 1112.41 (2+), which is corresponding to glycan G2+2SA-F. ..................................................................................................................................169

Figure 5-6. Glycosylation profile of IgG subclasses obtained by FTICR MS analysis of tryptic glycopeptides from four HIV-1 infected patients. ..............................................................175
Figure 5-7. Glycosylation profile of IgG subclasses obtained by C18 chip-Q-TOF MS analysis of tryptic glycopeptides from four HIV-1 infected patients. .................................................................176

Figure 5-8. Total glycan profile obtained by PGC-chip QTOF MS analysis of PNGase F released glycans..........................................................177

Figure 5-9. Total glycan profile obtained by mAb-Glyco-chip QTOF MS analysis of intact IgGs. ..........................................................177
LIST OF TABLES

Table 1-1. Structural and biological properties of human antibody isotypes. .......................... 26
Table 1-2. Physicochemical properties of four human IgG subclasses. ................................. 29
Table 1-3. Comparison of the common mass analyzers. ....................................................... 39
Table 1-4. Performance comparisons of the hybrid mass spectrometry instruments. ............... 42
Table 1-5. Summary of mass shift of common chemical modifications. ............................... 58

Table 2-1. Theoretical tryptic peptides of anti-HER2 mAb without reduction, including linked disulfides, cysteine position and molecular weight. ......................................................... 81
Table 2-2. Assignment of all correct disulfides in band A1 from the time point at 0 min of Figure 2-3A with the observed m/z at the charge of the highest intensity and the retention time......... 87
Table 2-3. Assignment of different chain combinations along with their corresponding disulfide linkages and major scrambled disulfides in the gel bands from the time point at 35 min of Figure 2-3A................................................................................................................................. 87
Table 2-S1A. Theoretical tryptic peptides of anti-CD11a mAb without reduction, including linked disulfides, cysteine position, and molecular weight ........................................... 99
Table 2-S1B. The assignment of different chain combinations along with their corresponding disulfide linkages and major scrambled disulfides in the gel bands from the time point at 35 min of Figure 2-7A................................................................................................................................. 100

Table 3-1. Percentage of glycan distribution generated on 6 different days. .......................... 115

Table 4-1. Summary of Primary Structure Coverage.............................................................. 132
Table 4-2. Comparison of the percentage of terminal modification for innovator and biosimilar anti-HER2 mAbs. .................................................. 138
Table 4-3. Comparison of the percentage of oxidation for innovator and biosimilar anti-HER2 mAbs. .................................................. 142
Table 4-4. Comparison of the percentage of succinimide intermediate formed during asparagine deamidation for innovator and biosimilar anti-HER2 mAbs. .................................................. 144
Table 4-5. Comparison of the percentage of succinimide intermediate formed during aspartic acid isomerization for innovator and biosimilar anti-HER2 mAbs. .................................................. 146

Table 5-1. Possible tryptic glycopeptides of IgG subclasses .................................................. 178
Table 5-2. Possible Lys-C glycopeptides of IgG subclasses .................................................. 180
Table 5-3. IgG isotype-specific Fc glycosylation profiles obtained by FTICR MS analysis of tryptic glycopeptides from four HIV-1 infected patients. .................................................. 182
Table 5-4. IgG isotype-specific Fc glycosylation profiles obtained by QTOF analysis of Lys-C glycopeptides from four HIV-1 infected patients. .................................................. 183
Table 5-5. Total glycan profile obtained by PGC-chip QTOF MS analysis of PNGase F released glycans .................................................. 184
Table 5-6. Total glycan profile obtained by mAb-Glyco-chip QTOF MS analysis of intact IgGs. ........................................................................................................ 185
Table 5-7. Patients Information .................................................................................. 185
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ab</td>
<td>antibody</td>
</tr>
<tr>
<td>ADCC</td>
<td>antibody-dependent cell-mediated cytotoxicity</td>
</tr>
<tr>
<td>Asn</td>
<td>asparagine</td>
</tr>
<tr>
<td>Asp</td>
<td>aspartic acid</td>
</tr>
<tr>
<td>CDC</td>
<td>complement-dependent cytotoxicity</td>
</tr>
<tr>
<td>CDR</td>
<td>complementarity determining region</td>
</tr>
<tr>
<td>CID</td>
<td>collision induced dissociation</td>
</tr>
<tr>
<td>CV</td>
<td>coefficient of variation</td>
</tr>
<tr>
<td>Cys</td>
<td>cysteine</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>ECD</td>
<td>electron capture dissociation</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>ESI</td>
<td>electrospray ionization</td>
</tr>
<tr>
<td>ETD</td>
<td>electron transfer dissociation</td>
</tr>
<tr>
<td>Gal</td>
<td>galactose</td>
</tr>
<tr>
<td>GleNAc</td>
<td>N-acetylglucosamine</td>
</tr>
<tr>
<td>Gln</td>
<td>glutamine</td>
</tr>
<tr>
<td>Glu</td>
<td>glutamic acid</td>
</tr>
<tr>
<td>Fab</td>
<td>fragment antigen-binding</td>
</tr>
<tr>
<td>Fc</td>
<td>fragment crystallizable</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
</tbody>
</table>
FTICR  fourier transform ion cyclotron resonance
F        fucose
HCl      hydrochloric acid
HPLC     high performance liquid chromatography
IEF      isoelectric focusing
IgG      immunoglobulin G
IsoAsp   isoaspartic acid
ITRAQ    isobaric tag for relative and absolute quantitation
kDa      kilodalton
LC-MS    liquid chromatography mass spectrometry
LC-MS/MS liquid chromatography with tandem mass spectrometry
LIT      linear Ion Trap
m/z      mass to charge ratio
mAb      monoclonal antibody
MALDI    matrix assisted laser desorption ionization
Man      mannose
Met      methionine
mg       milligram
mL       milliliter
mm       millimeter
MS       mass spectrometry
MS\(^2\)  tandem mass spectrometry
MS/MS    tandem mass spectrometry
PGC  porous graphitic carbon
PTM  posttranslational modification
PyroGlu.pyroglutamic acid
QTOF.quadrupole coupled with time-of-flight
RF  radio frequency
RCM  reduced and carboxymethylated
RP  reversed phase
RP-HPLC  reversed phase high performance liquid chromatography
RSD  relative standard deviation
SA  sialic acid
SDS  sodium dodecyl sulfate
SDS-PAGE  sodium dodecyl sulfate polyacrylamide gel electrophoresis
SILAC  stable isotopic labeling of amino acids in cell culture
TOF  time-of-flight
Tris  2-Amino-2-hydroxymethyl-propane-1,3-diol
Trp  tryptophan
μg  microgram
μL  microliter
Chapter 1
Overview of Mass Spectrometric Methods for Structural Characterization of Therapeutic Monoclonal Antibodies
1.1 Abstract

Antibodies, or immunoglobulins, play an important role in treatment of cancers or other inflammatory diseases. However all therapeutic antibodies are heterogeneous due to posttranslational modifications (PTMs) occurring physiologically, or other chemical modifications induced during all stages of production. Comprehensive characterization of therapeutic is critical to ensure drug efficacy and safety. Mass spectrometry (MS) is a central analytical tool for both global and fine structural characterization of therapeutic antibodies due to its superior resolution, sensitivity, and accuracy.

In this chapter, an overview of therapeutic antibodies structure and the importance of structural characterization of therapeutic mAbs are first provided. Then, the principles and technologies of mass spectrometry (MS) are reviewed. The strategies of MS-based methods for the structural characterization of monoclonal antibodies are discussed. Among three MS-based strategies, bottom-up mass spectrometry is the most frequently used method for structural characterization and is capable of characterizing modifications at the amino acid level. Therefore, the applications of bottom-up mass spectrometry for primary sequence analysis, disulfide-linkage determination, glycosylation analysis, common chemical modifications characterization and quantitation are further discussed in detail.

1.2 Antibodies

1.2.1 Human immunoglobulin G antibodies

Antibodies (Abs), also known as immunoglobulins (Ig), are heavy (~150 kDa) globular plasma proteins produced by B cell to recognize and eliminate foreign material, such as bacteria and viruses. Each antibody is composed of two identical heavy chains and two identical light
chains linked by disulfide bonds to form a "Y" shaped structure (also known as Ig unit). Human antibodies are divided into five classes, IgA, IgD, IgE, IgG, and IgM, based on their heavy chain structure (α, δ, ε, γ, and μ, respectively). The structural and biological properties of human antibodies are listed in Table 1-1. IgGs are the most abundant antibodies, which constitute 75% of total serum antibodies. In addition, IgGs have the longest half-life in blood compared to other classes of Igs. Furthermore, antibody-dependent cell-mediated cytotoxicity (ADCC) and antibody-mediated complement-dependent cytotoxicity (CDC), two major IgG effector functions and also the most focused aspects in cancer immunotherapeutic approach, are triggered by IgGs. Due to these distinguished properties, currently all therapeutic antibody drugs on market are based on an IgG construct.

Human IgG (Figure 1-1) is a monomer with two identical heavy chains (γ) approximately 50 kDa and two identical light chains (either κ or λ) approximately 25 kDa. Each chain is composed of domains. The domain is a characteristic immunoglobulin fold, in which two antiparallel beta sheets are held together by a disulfide bond to form a barrel shape. The light chain has one variable domain and one constant domain. The heavy chain consists of one variable domain (VH) and three constant domains (CH1, CH2, and CH3). As indicated by their names, the amino acid sequences in constant domains are the same for the IgGs from the same subclass but different in variable domains between IgG molecules even from the same subclass. Each variable domain has three short, separated, but most variable sequences, called complementarity determining regions (CDRs: CDR1, CDR2 and CDR3). CDRs form a unique surface where antibody specifically complements antigen’s shape. Thus, CDRs play an important role in antigen recognition. The region between the CH1 and CH2 domains is called the hinge region due to some flexibility in the antibody at this point. The two arms of the Y-shaped
molecule are Fab (fragment antigen-binding) regions and its base part is Fc (fragment crystallizable) region. The Fab region consists of whole light chain and VH and CH1 domains from the heavy chain. As its name indicates Fab region is the site for recognizing and binding foreign objects (antigen). The Fc region is composed of CH2 and CH3 domains from heavy chains and is the site to interact with cell surface Fc receptors and some proteins of the complement system. Thus, the Fc region plays a role in activating various immune responses such as antibody-dependent cell-mediated cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC), etc. N-linked glycans, typically on the conserved Asn297 in CH2 domain, can strongly influence the binding of IgG to Fc receptor and therefore modify Fc receptor-mediated activity.\textsuperscript{1-2,5}

\textbf{Figure 1-1. Schematic structure of the human IgG molecule.}
Green parts are light chains and orange parts are heavy chains. V indicates variable domains; C indicates constant domains; VL and CL are domains of the light (L) chain; VH, CH1, CH2, and CH3 are domains of the heavy (H) chain; the red regions in variable domains, CDR1, CDR2 and CDR3, are complementary determining regions; –S-S- indicates disulfide bond.
Table 1-1. Structural and biological properties of human antibody isotypes. This Table is modified from Refs 3 and 6.

<table>
<thead>
<tr>
<th>Properties</th>
<th>Classes and subclasses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IgG</td>
</tr>
<tr>
<td></td>
<td>IgG1</td>
</tr>
<tr>
<td>Molecular weight of secreted form (kDa)(^{(A)})</td>
<td>150</td>
</tr>
<tr>
<td>Adult level range (age 16-60) in serum (g/L)</td>
<td>5-12</td>
</tr>
<tr>
<td>Approximate % total IG in adult serum</td>
<td>45-53</td>
</tr>
<tr>
<td>Biological half-life (day)</td>
<td>21</td>
</tr>
<tr>
<td>Placental transfer</td>
<td>+++</td>
</tr>
<tr>
<td>Activation of complement classical pathway (C1q)</td>
<td>++</td>
</tr>
<tr>
<td>Binding macrophages and other phagocytic cells (FcγR or FCGR)(^{(B)})</td>
<td>+++</td>
</tr>
<tr>
<td>Binding to mast cells and basophils (FcεR or FCER)</td>
<td>-</td>
</tr>
</tbody>
</table>

(A) approximate molecular weight (m=monomer, d=dimer, p=pentamer)
(B) +/-: binding depends on the FCGR isotype and on the cell type
(C) IgG2 binds only to two known functionally expressed FcγR IIa alleles (FcγR IIa-H131)
The constant domains of the heavy chain determine IgG subclasses and they are named in order of their abundance: IgG1, IgG2, IgG3, and IgG4 (IgG1 > IgG2 > IgG3 ≈ IgG4). In addition to amino acid differences, the inter-chain disulfide bonds within the hinge region and between heavy and light chains are also different among IgG subclasses (Figure 1-2 and Table 1-2). The difference of their structure determines the difference of their biological function and stability (Table 1-1). For example, IgG3 has 11 inter-heavy chain disulfide bonds within its long hinge region. The initiation of the complement classical pathway is the binding of C1q to the sites on the CH2 domains of IgG. The longest hinge of IgG3 makes it most flexible and therefore the most effective complement activator. By contrast, IgG4 does not activate complement due to steric hindrance of the complement binding sites caused by its structure. However, the long hinge region of IgG3 is prone to proteolysis and results in a short half-life in blood (approximate 7 days compared to 21 days of IgG1, IgG2, and IgG4), which makes IgG3 seldom used as a therapeutic agent. So far, except IgG3, other subclasses of IgGs are widely used in therapeutic purpose. Among them, IgG1 is most suitable for cancer treatment due to its highest ADCC and CDC activity and longest blood half-life.

1.2.2 History of therapeutic monoclonal antibodies (mAbs)

A century ago, Paul Ehrlich, the founder of cancer chemotherapy and a Nobel laureate, first proposed the idea of a “magic bullet” that can selectively target tumors and kill the targeted cancer cells only. The concept of a "magic bullet" was fully realized in therapeutic use with the ability of production of murine monoclonal antibodies (mAbs) by hybridoma technology established by Georges Köhler and César Milstein in 1975. However, the clinical utility of these murine mAbs was limited by their immunogenicity in humans and poor ability to induce human immune effector responses. Later great effort was made to overcome the problems of murine mAbs. In 1984, chimeric monoclonal antibodies containing
66% human and 34% mouse sequences were engineered using recombinant DNA technology. In 1988, Greg Winter and his team pioneered the techniques to successfully humanize antibodies, which have only 5-10% mouse sequences. Fully human antibodies could be produced by transgenic-mouse technologies and phage-display technologies in 1990.\textsuperscript{8-9, 13-14}

Muromonab-CD3 (trade name Orthoclone OKT3, marketed by Janssen-Cilag) was the first therapeutic mAb to be approved by the U.S. Food and Drug Administration (FDA) in 1986. It was a murine IgG2 mAb targeted at T cell CD3 receptor and used to lower the body's natural immunity in patients with organ transplants.\textsuperscript{2, 8} Since then until 2007, a total of 26 therapeutic mAb drug products have entered American market and a $ 36.4 billion record had been achieved on the global market on 2009.\textsuperscript{15-16}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{IgG_subclasses}
\caption{Schematic structures of four human IgG subclasses. This figure is modified from Ref 17.}
\end{figure}
Table 1-2. Physicochemical properties of four human IgG subclasses. This Table is modified from Refs 7 and 17.

<table>
<thead>
<tr>
<th></th>
<th>IgG1</th>
<th>IgG2</th>
<th>IgG3</th>
<th>IgG4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavy chain type</td>
<td>γ1</td>
<td>γ2</td>
<td>γ3</td>
<td>γ4</td>
</tr>
<tr>
<td>Amino acids in hinge region</td>
<td>15</td>
<td>12</td>
<td>62</td>
<td>12</td>
</tr>
<tr>
<td>Core hinge sequences</td>
<td>EPKSCDKTHTCPPCP</td>
<td>ERKCCVECPPCP</td>
<td>ELKTPGDTTTHTCPPRCP (EPKSCDTPPPPCPPRCP)3</td>
<td>ESKYGPPCPPSCP</td>
</tr>
<tr>
<td>Inter-heavy chain disulfide bonds (in hinge region)</td>
<td>2</td>
<td>4</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Susceptibility to proteolytic enzymes</td>
<td>++</td>
<td>+/-</td>
<td>+++</td>
<td>+</td>
</tr>
</tbody>
</table>
1.2.3 Types of therapeutic mAbs

A mAb is a specific antibody produced by a single B lymphocyte clone and therefore these antibodies are theoretically homogeneous in structure. Because of this structural homogeneity, including the amino acid sequences in variable region, the mAbs from the same cell clone specifically target the same epitope.\textsuperscript{18,19}

Currently, all marketed therapeutic mAbs are either from a murine source or human source, or from a combination of both. According to their sequence source, therapeutic mAbs are classified into four types: murine antibodies; chimeric antibodies; humanized antibodies; and fully human antibodies.\textsuperscript{8-9,18,20} A murine antibody has 100% mouse sequence. A chimeric antibody with murine variable regions and human constant regions contains approximately 34% mouse and 66% human sequences. Humanized antibody is constructed with murine complementarity determining regions (CDRs, 5-10% of total sequence) and the remainder is from a human source. Fully human antibody contains 100% human sequence.\textsuperscript{8-9,16,18,20} The immunogenicity is decreased as the percentage of human sequence is increased.\textsuperscript{8-9,18} Due to this reason, the clinical study of fully human mAbs increased significantly while the other three types of mAbs reduced in 2002.\textsuperscript{14} The generic name of mAbs ends with –mab and the infixes reflect the type and sequence source of mAbs. Four infixes, -o-, -xi-, -zu-, and -u-, are used for indication of mouse, chimeric, humanized, and fully human mAb, respectively. For example, tositumomab (brand name: Bexxar) is murine, rituximab (brand name: Rituxan and MabThera) is chimeric, trastuzumab (brand name: Herceptin) is humanized, and Ofatumumab (brand name: Arzerra, also known as HuMax-CD20) is human.\textsuperscript{8}
1.2.4 Advantages and limitations of therapeutic mAbs

The development of monoclonal antibodies as therapeutic agents has been one of the fastest-growing areas in biotechnological and pharmaceutical industry due to several advantages over small-molecule drugs. MAbs are clones of antibodies that selectively bind to a specific epitope. Because of their high specificity, fewer side effects are usually caused by mAbs. MAbs have longer half-life compared to small molecules allowing less frequent dosing requirements for mAb drugs.\textsuperscript{8, 16, 21}

In addition, the technical advances and expanding knowledge make the mAb less immunogenic and enhance clinical potential. For example, the production of fully human mAbs by transgenic-mouse and phage-display technologies increases their efficacy by modification of their Fc glycan structures.\textsuperscript{8, 13, 22-23} The minimized immunogenicity and specific therapeutic actions greatly increase their success rate in clinical development and FDA approval. It has been reported that approval rates of mAbs compared to small-molecule drugs are 20\% vs 5\%.\textsuperscript{8, 13, 21} Furthermore, due to their unique structure and function, far-reaching patent protection for therapeutic mAbs can be obtained.\textsuperscript{21}

However, therapeutic mAbs are expensive for both company production and customer consumption.\textsuperscript{13} Compared to small molecules (for example 500 Da), the large molecular size and surface charge of mAbs (approximately 150 kDa) hinder their penetration through cell membrane, therefore limiting their clinical applications to cell surface or secreted targets.\textsuperscript{13, 24} Furthermore, the antibodies can be easily digested in stomach and gastrointestinal tract and restricted diffusion through the gastrointestinal epithelium prevent their reach to targeted antigens. With the combination of these factors, mAbs have to be always administered intravenously.\textsuperscript{13, 24-25}
1.2.5 Common chemical modifications in therapeutic mAbs

Because mAbs are produced from a single cell clone, their primary structure should be theoretically homogenous. However, mAbs may undergo a series of modifications during their manufacturing, formulation, and storage and become heterogeneous. The most common modifications in mAbs include: N-terminal pyroglutamic acid cyclization, heavy chain C-terminal lysine processing, oxidation, deamidation, and isomerization.\(^1,26\)

The N-terminal residue of most therapeutic mAbs is glutamine or glutamic acid and they are easily cyclized to pyroglutamic acid. This nonenzymatic modification is spontaneous and may not have significant impact on Ab structure and function.\(^1,26-28\) However, the release of N-terminal primary amine during cyclization of glutamine to pyroglutamate makes the cyclized forms more acidic. As a result, mAbs with N-terminal glutamine elute later than their cyclized counterparts on cation exchange chromatography but earlier on reversed phase chromatography.\(^26\) Cyclization of N-terminal glutamine and glutamic acid causes a release of ammonia and water, respectively. The release of ammonia and water results in a 17 and 18 Da mass difference in turn between non-cyclized and cyclized species.

Complete or partial removal of heavy chain C-terminal lysine by carboxypeptidases during cell culture production is one of the most common modifications of mAbs. Similar to N-terminal pyroglutamic acid cyclization, C-terminal lysine processing has no significant influence on Ab structure and functions such as antigen binding and receptor binding to Fc.\(^1,26,29-30\) The loss of one C-terminal lysine residue reduces the positive charge by 1 unit.\(^26\)

Oxidation of methionine (Met) to methionine sulfoxide is often observed during purification and storage.\(^31\) CH2 domain stability could be decreased if this modification occurs in
Oxidation of tryptophan (Trp) residue in the CDR regions has also been reported, but less commonly observed than Met oxidation.\textsuperscript{1, 26, 33-34}

Deamidation of asparagines is a common modification that can occur at any time but especially after long-term storage. The most susceptible site for deamidation is the asparagine followed by glycine or serine.\textsuperscript{26} The process goes through the loss of amine in the side chain of asparagine to form succinimide intermediate. This intermediate is usually hydrolyzed quickly to aspartic acid and isoaspartate in a molar ratio of 1:3.\textsuperscript{1, 26, 35} Compared to asparagine residue, the succinimide intermediate has a 17 Da reduced mass but the deamidated products, aspartic acid and isoaspartate, have a 1 Da increased mass charge. In the hydrophobicity aspect, succinimide intermediate is most hydrophobic followed by aspartic acid containing peptide while peptide with asparagine instead is most hydrophilic.\textsuperscript{36-37}

Aspartic acid followed by glycine allows release of a molecule of water to form a succinimide, the same intermediate as in the asparagine deamidation. The succinimide intermediate is unstable and is quickly hydrolyzed to isoaspartic acid, which introduces a methyl group to the peptide backbone.\textsuperscript{26} Asparagine deamidation and aspartic acid isomerization both induce a structural change and may be immunogenic or decrease drug potency.\textsuperscript{38-39} There is no mass difference between aspartic acid and isoaspartic acid but different chromatographic behavior is observed.\textsuperscript{26}

### 1.2.6 Importance of structural characterization of therapeutic mAbs

A mAb is a large molecule with complex 3D structure and function, its structure and modifications have significant influences on the quality, safety, and efficacy of drug product. For example, CH2 domain stability could be decreased if this oxidation of methionine occurs in the
CH2 domain.\textsuperscript{32} Asparagine deamidation and aspartic acid isomerization both induce a structural change and may potentially be immunogenic or decrease drug potency.\textsuperscript{38-39} Glycan profiles can significantly influence the efficacy, safety and pharmacokinetics. For example, it has been reported that IgG Abs with high-mannose glycans on Fc region can clear more quickly \textit{in vivo}.\textsuperscript{41} Deglycosylated IgG is unable to activate its effector function but afucosylated IgG is more efficacious.\textsuperscript{42} Incorrect glycosylation can lead to a very short half-life in bloodstream.\textsuperscript{43} Nonhuman cell lines, which are the major cell lines currently used for production of therapeutic mAbs, can express a small amount of antibodies with nonhuman glycans such as galactose α (1-3) galactose and \textit{N}-glycolyneuraminic acid glycans. These glycans can be immunogenic.\textsuperscript{44-46} Incomplete formation of disulfide bonds and cysteinylation of free cysteine could cause loss of drug efficacy.\textsuperscript{47-50} Therefore, comprehensive characterization of therapeutic mAbs structure is critical to biotech and pharmaceutical industry.

\textbf{1.2.7 Anti-Her2 mAb and treatment of breast cancer}

Trastuzumab (trade name: Herceptin) from Genentech was approved in 1998 for treatment of HER2 (Human epidermal growth factor receptor 2; also known as c-neu or ErbB-2) overexpressing metastatic breast cancer.\textsuperscript{51} HER2 is a cell membrane surface receptor and able to form a dimer complex by binding a ligand, for example growth factor. The formation of the dimer complex on the cell surface leads to activation of intrinsic protein tyrosine kinase activity, triggering complex signal transduction cascade and resulting in cell division and proliferation.\textsuperscript{52} In tumor cells, overexpressed Her2 receptors activate the cell to divide and multiply at an accelerated rate thus contributing to tumor growth. Anti-HER2 mAb (Herceptin) can specifically bind to the HER2 receptors, and therefore disrupt the formation of heterodimer receptor complex
on the breast cancer cell surface and block the corresponding signal transduction cascade. As a result, tumor cell proliferation is inhibited and breast cancer can be controlled.53

This mAb against the overexpressed HER2 receptor (anti-HER2) is a humanized, IgG1-type mAb. It is composed of two identical light chains (about 25 kDa) and heavy chains (50 kDa) stabilized by 4 inter- and 12 intra-chain disulfide bonds (Figure 1-3). There are 450 amino acids in the heavy chain and 214 amino acids in the light chain. The light chain and heavy chain are linked together by an inter-chain disulfide bond between the C-terminal end of the light chain (Cys214) and Cys223 in the heavy chain. Two parallel inter-chain disulfides are formed between two heavy chains by connecting Cys229 and Cys232 in one heavy chain and the same amino acid sites in the other heavy chain. Each light chain and heavy chain has two and four intra-chain disulfide bonds, respectively. N-linked bi-antennary glycans are attached to Asn300 on heavy chains.

In 2009, this drug had sales of $4.85 billion, among top five best-selling mAbs.15 On the other hand, patients struggle for this big payment. It has been reported that the average cost for Trastuzumab per year per patient is $37,000.54 With the loss of patent protection, a generic product, called “biosimilar” in Europe or “follow-on” in US, could be developed with a lower-price and is of great benefit to both the pharmaceutical industry (stimulate further research) and customers (cost savings).55
Figure 1-3. Structure of anti-HER2 mAb.
Green part is its variable domains and black part is its constant domains.

1.3 Mass Spectrometry

Mass spectrometry is an analytical technique to identify molecules by measuring their mass-to-charge ratio. A mass spectrometer is usually composed of ion source, mass analyzer, mass detector, and data analysis system. Among these components ion source and mass analyzer are the most important parts of a mass spectrometer.

1.3.1 Ionization methods

Ion source is a place where samples are transformed into gas phase and introduced into mass spectrometer. A variety of ionization methods are available for different types of samples and different analysis purposes. Proteins and peptides are polar, nonvolatile and large molecules. Due to these properties, electrospray ionization (ESI) and matrix-assisted laser desorption
ionization (MALDI) are the most common ionization methods for protein and peptide characterization.¹

1.3.1.1 Electrospray ionization (ESI)

Positive ion analysis mode is often used for proteins and peptides analyzed by ESI. A strong electric field is generated between the probe tip at the end of the separation pipeline and the entrance of mass spectrometer, which cause negative ions to move towards the conductive probe tip and positive ions to move towards the liquid surface to form a Taylor cone. The tip of the cone emits a stream of charged droplets and the spray appears.⁵⁷ Desolvated ions are finally obtained through evaporation of the solvent or by direct ion emission. When positive ions are transferred into mass spectrometer, the same number of negative ions, which accumulates on conductive probe tip, are neutralized by oxidation occurring in the probe tip. Through this electrochemical process, the charge balance is maintained to allow continuous production of charged ions. And this electrochemical process makes ESI sensitive to sample concentration not to the total sample amount.⁵⁷

ESI is operated under atmospheric pressure, which makes the sample introduction into mass spectrometer with high vacuum inside much easier. In addition, other separation techniques such as LC and CE can be adaptable to ESI. The further separation of samples significantly improves the dynamic range of MS analysis.⁵⁷ Most importantly, ESI is a soft ionization technique. It is able to transfer large and nonvolatile proteins or peptides from liquid phase into gas phase with no fragmentation and produce multiply charged ions. The high charge states make the analysis of large molecules possible for mass analyzers with a relatively small mass range. The disadvantages of ESI include the presence of salts and denaturants, which are often used in
sample pretreatment and could greatly suppress the protein and peptides ionization. Multiple charging can make data analysis confusing especially when a complex sample is analyzed.\textsuperscript{57}

1.3.1.2 Matrix-assisted laser desorption ionization (MALDI)

Analyte compounds are embedded in a crystalline matrix composed of small organic molecules. A short, intense laser pulse strikes on the dried mixture and induces rapid heating on matrix. Then matrices are sublimated and expanded into gas phase, which carries analyte compounds with it and then helps them transfer into gas phase.\textsuperscript{57}

MALDI is a soft ionization method and generates mostly singly charged ions, which is very useful for MW measurement of intact proteins. MALDI is a pulsed ionization technique and well suited to time-of-flight (TOF) mass analyzer since TOF requires a bundle of ions. MALDI is able to desorb and ionize molecules with MW as high as 100 kDa. With the connection to TOF, having high-resolution, high-mass accuracy and wide mass analysis range, MALDI-TOF MS is a powerful technique for top-down analysis of intact mAb.\textsuperscript{56}

However, the reproducibility of the mass spectra generated from shot-to-shot is low and matrix background is a problem in low mass range spectrum.\textsuperscript{56-57}

1.3.2 Mass analyzers

Mass analyzers can be divided into two classes: scanning mass analyzers and non-scanning mass analyzers. Scanning mass analyzers, such as quadrupole, scan and transfer ions with specific m/z at any given time. Non-scanning mass analyzers transmit all ions simultaneously, which includes ion-beam (TOF) and trapping mass analyzers (ion trap, Orbitrap, and FTICR).\textsuperscript{57} The comparison of common mass analyzers is summarized in Table 1-3.
Table 1-3. Comparison of the common mass analyzers. This Table is taken from Ref 57.

<table>
<thead>
<tr>
<th></th>
<th>Quadrupole</th>
<th>Ion Trap</th>
<th>TOF reflectron</th>
<th>FTICR</th>
<th>Orbitrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass limit</td>
<td>4000 Th</td>
<td>6000 Th</td>
<td>10000 Th</td>
<td>30000 Th</td>
<td>50000 Th</td>
</tr>
<tr>
<td>Resolution</td>
<td>2000</td>
<td>4000</td>
<td>20000</td>
<td>500000</td>
<td>100000</td>
</tr>
<tr>
<td>FWHM Accuracy (m/z 1000)</td>
<td>100 ppm</td>
<td>100 ppm</td>
<td>10 ppm</td>
<td>&lt;5 ppm</td>
<td>&lt;5 ppm</td>
</tr>
<tr>
<td>Ion Sampling</td>
<td>Continuous</td>
<td>Pulsed</td>
<td>Pulsed</td>
<td>Pulsed</td>
<td>Pulsed</td>
</tr>
<tr>
<td>Pressure</td>
<td>$10^{-3}$ Torr</td>
<td>$10^{-3}$ Torr</td>
<td>$10^{-6}$ Torr</td>
<td>$10^{-10}$ Torr</td>
<td>$10^{-10}$ Torr</td>
</tr>
<tr>
<td>Tandem mass spectrometry</td>
<td>MS/MS</td>
<td>MS\textsuperscript{n}</td>
<td>MS/MS</td>
<td>MS\textsuperscript{n}</td>
<td>MS\textsuperscript{n}</td>
</tr>
</tbody>
</table>

1.3.2.1 Quadrupole

A quadrupole is a mass filter, in which an oscillating electric field is formed and ions with stable trajectories can pass through quadrupole and be detected. It is composed of four parallel metal rods applied to rods. A radio frequency (RF) voltage and a direct current (DC) voltage on each opposing rod pair make a time-varying electric field that determine which m/z has stable trajectory and are able to go through the metal rods.\textsuperscript{57}

Quadrupole is a low-resolution mass analyzer and due to its fast scan speed, Q is well compatible with chromatographic coupling. Another important property of Q is its ability to focus the trajectory of the ions toward the center of the rock when it operates in RF-only mode. This property of Q, now extending to hexapoles and octaples, has been extensively used in ion guides and collision cells.\textsuperscript{57}
1.3.2.2 Paul ion trap

Paul ion trap is a trapping mass analyzer in which ions with stable trajectories in oscillating electric field are trapped and ions with destabilized trajectories are expelled and detected. It consists of a ring electrode and two endcap electrodes and a radio frequency voltage is applied to the ring electrode to define oscillating electric field.\textsuperscript{57}

The most important feature of the ion trap is that it can isolate and trap ions with particular \( m/z \) by ejecting all other ions. The isolated ions can be further accumulated to get a better ion signal or be fragmented to do MS\textsuperscript{n} scan. However, an ion trap is also a low-resolution mass analyser and its dynamic range is limited due to space charge effects.\textsuperscript{57}

1.3.2.3 Linear Ion Trap (LIT)

LIT is a two dimensional Paul ion trap and composed of four-rod quadrupole and ending lenses. By application of radio frequency voltage on rod and direct current voltage on the end electrodes, ions are trapped inside the rods. Linear ion trap has a greater trapping volume compared to quadrupole ion trap and ions are focused along the central line in LIT rather than around a point in Paul IT. As a result, LIT has a more than 10-fold higher ion storage capacity than Paul IT. In addition, LIT has a higher trapping efficiency. The sensitivity and the dynamic range of LIT are enhanced due to these advantages.\textsuperscript{57}

1.3.2.4 Time-of-Flight (TOF)

As its name indicates, TOF analyzer analyses ions based on the time they fly in a field-free region. In TOF, ions are first accelerated so that ions with the same charge have the same kinetic energy. Then the velocity, and therefore the time, of the ions is determined by their \( m/z \).\textsuperscript{57}
In principle, the TOF instrument has unlimited mass range, which makes it especially suitable to MALDI. All the ions in the packet, theoretically, are transmitted and detected and therefore TOF has very high sensitivity. In addition the analysis speed of TOF is very fast and the instrument is relatively low cost. However, TOF requires high vacuum and coupling to continuous ion sources, such as ESI, is not easy.\textsuperscript{57}

1.3.2.5 Orbitrap

Orbitrap is an electrostatic ion trap. It is composed of a barrel-shaped outer electrode and a spindle-shaped central electrode. The ions are injected tangentially into long axis of the orbitrap and start to turn around the central electrode because their electrostatic attraction to the central electrode is balanced by centrifugal force from initial tangential velocity. In addition, the ions also oscillate back and forth along the central electrode. The oscillation frequency is directly related to the m/z ratio of injected ions but not linked to their kinetic energy. The oscillation frequency is then measured and converted to a mass spectrum by fourier transform (FT).\textsuperscript{57} Because of the energy independence and larger trapping volume compared to Paul ion trap, Orbitrap has very high mass resolution and mass accuracy.\textsuperscript{57}

1.3.2.6 Fourier Transform Ion Cyclotron Resonance (FTICR)

FTICR analyzes ions based on the cyclotron frequency of the ions in a fixed magnetic field. It is composed of a cube with six plates and placed inside a large magnet. Ions are trapped and resonate in the cube by an oscillating electric field perpendicular to the magnetic field. The current image of the ions is measured by receiver plates and converted to a mass spectrum by fourier transform (FT).
FTICR has very high mass accuracy and the highest mass resolution of all mass analyzers, which makes it very attractive for top-down proteomics research and bottom-up characterization. MS\textsuperscript{n} in a non-destructive way can be performed on FTICR so that the fragments can be detected repetitively with an increased sensitivity and resolution. However, in order to achieve high resolution, a very high cell vacuum is necessary for FTICR and its dynamic range is limited due to its finite total trap charge capacity. In addition, FTICR has a big size and is very expensive.\textsuperscript{57}

1.3.3 Hybrid mass spectrometer instruments

To take advantages of each mass analyzer and avoid their weaknesses, several types of mass analyzers are often integrated into one hybrid instrument. The most widely used hybrid instruments include: LTQ-Orbitrap hybrid instrument (Thermo Scientific; linear ion trap with orbitrap), LTQFTICR (Thermo Scientific; linear ion traps with FTICR), and Q-TOF (Agilent; quadrupole with TOF). Table 1-4 highlights comparative features of these three instruments.

Table 1-4. Performance comparisons of the hybrid mass spectrometry instruments. This table is modified from Ref 56.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Resolution</th>
<th>Mass accuracy</th>
<th>Sensitivity</th>
<th>Dynamic range</th>
<th>Scan rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTQ-Orbitrap</td>
<td>100,000</td>
<td>2 ppm</td>
<td>Femtomole</td>
<td>1.E+04</td>
<td>Moderate</td>
</tr>
<tr>
<td>LTQ-FT</td>
<td>500,000</td>
<td>&lt; 2 ppm</td>
<td>Femtomole</td>
<td>1.E+04</td>
<td>Slow</td>
</tr>
<tr>
<td>QTOF</td>
<td>10,000</td>
<td>2-5 ppm</td>
<td>Attomole</td>
<td>1.E+06</td>
<td>Moderate</td>
</tr>
</tbody>
</table>
1.3.3.1 LTQ-Orbitrap

There are three major components in the LTQ-Orbitrap: liner ion trap, C-trap and Orbitrap. The linear ion trap is used for isolation, fragmentation and MS^n analysis; C-trap is used to transform ion beams into very short pulsed ion packets suitable for Orbitrap mass analyzer; Orbitrap is used as a mass analyzer and detector to provide high-resolution and high-mass accuracy analysis. Therefore, LTQ-Orbitrap has the advantages of high resolution and mass accuracy of the Orbitrap and high sensitivity and fast speed of linear ion trap.\(^{57}\) In addition, LTQ-Orbitrap can be operated in parallel mode: full MS scans are acquired in Orbitrap while linear ion trap performs fragmentation reactions. Compared to LTQ-FT, the LTQ-Orbitrap is able to couple to ETD and therefore analysis of some types of peptides, such as disulfide-bridged cyclic peptide, is easier and accurate on LTQ-Orbitrap.\(^{58}\) LTQ-orbitrap is quite cheaper than LTQ-FT at both purchase and maintenance but provides comparable performance.

1.3.3.2 LTQ-FT

Similar to LTQ-Orbitrap, LTQ-FT combines the advantages of both linear ion trap and FTICR. However, LTQ-FT has a broader mass-to-charge range and higher resolution compared to LTQ-Orbitrap. The better performance of LTQ-FT mainly is because ions can be trapped and detected for longer than in FT than Orbitrap.\(^{56-57}\)

1.3.3.3 Q-TOF

Q-TOF is another attractive hybrid mass spectrometry instrument. It integrates the simplicity of the quadrupole and the high sensitivity and high mass accuracy of TOF. Q-TOF
consists of a quadrupole analyser Q1 and a hexapole collision cell q2 and a reflecting TOF with orthogonal injection of ions.

Q-TOF can be operated in MS mode and MS/MS mode. In MS mode, Q1 and q2 are operated in RF-only mode acting as an ion transmission element, while TOF is the mass analyzer to record spectra. The generated results benefit from the high resolution and mass accuracy of the TOF but with a limited mass range from the transmission of the quadrupoles.\(^{59}\) In MS/MS mode, Q1 is operated in mass filter mode to select and transmit only precursor ions of interest; q2 is a collision cell where precursor ions undergo CID fragmentation after collision with argon or nitrogen contained in q2; TOF is used to analyze product ions. Since both MS scan and MS/MS scan are recorded by TOF, the accurate mass measurement of precursor and product ions can be achieved.

1.3.4 Tandem mass spectrometry

Tandem mass spectrometry is also known as two-stage mass spectrometry, MS/MS or MS\(^2\). ESI and MALDI are both suitable for protein and peptide ionization, and produce ions with little or no fragmentation and therefore provide limited structural information. As a result, tandem mass spectrometry emerges as a solution and generates fragmentation between stages.\(^{60}\) The common tandem mass spectrometry experiment involves: isolation of precursor ions by a first mass analyzer; fragmentation of the isolated precursor ions by activation to yield fragment ions; analysis of the product ions by a second mass analyzer. The number of stages could be increased, and the corresponding mass analysis is referred as multiple stage mass spectrometry or MS\(^n\).\(^{57}\) Depending on fragmentation techniques, the possible fragment ions that can be generated from peptides are shown in Figure 1-4.\(^{61}\)
1.3.4.1 Tandem Mass Spectrometry in Space or in Time

Tandem mass spectrometry can be performed in space or in time. Tandem mass spectrometry in space is accomplished by two physically distinct mass analyzers while tandem in time is implemented by performing a sequence of steps over time but in the same place. Tandem in space can be performed on triple quadrupoles, TOF with a reflectron, or Q-TOF. A FTICR, Orbitrap, and ion trap can be used for tandem in time.57

1.3.4.2 Fragmentation in Tandem Mass Spectrometry

Currently, collision-induced dissociation (CID; also named collision-activated decomposition, CAD) and electron transfer dissociation (ETD) are the most often used fragmentation techniques in tandem mass spectrometry.

In CID, the isolated precursor ions collide with the helium or nitrogen present in the collision cell. The collision process results in a vibrational excitation and converts part of the kinetic energy of precursor ions into internal energy, which induces dissociation of amide bond along the peptide backbone to produce b and y ions (Figure 1-4).62 The fragmentation efficiency of CID is best on doubly-charged, short and middle-sized protonated peptide ions.63 However, many posttranslational modifications (PTM) of proteins are labile under CID. This is because the bonds to link modifications are usually weaker than peptide bond and then are cleaved first. Therefore, CID is limited in PTM characterization.56

ETD (electron transfer dissociation) is a newly developed fragmentation technique. As indicated by its name, by transferring electrons to positively charged peptides, peptides are converted into radicals, become unstable and undergo fragmentation.62 Since it is difficult to produce free electron under atmospheric conditions, radical anions such as flouranthene are often
used as an electron carrier. ETD cleavages at Cα – N bond along the peptide backbone and generate c and z ions. Since fragmentation in ETD is induced by electron transferring and no vibrational excitation occurs, ETD cleaves peptide backbone extensively and keeps PTM attached and intact.\textsuperscript{57} This feature makes ETD extremely useful for PTM characterization.\textsuperscript{64} Compared to CID, ETD is more efficient at fragmentation of peptides with higher charge states.\textsuperscript{63}

1.3.5 Liquid Chromatography-Mass Spectrometry (LC-MS)

With the combination of the separation capability of LC and mass analysis capability of MS, LC-MS is a powerful technique for protein and mAb analysis with high sensitivity and specificity. Separation of samples by LC could increase the dynamic range and make identification of low-abundance species possible. In addition to separation advantage, LC can be easily coupled with ESI.\textsuperscript{65}

![Peptide fragments generated in tandem mass spectrometry.](image)

If the charges remain on the N terminal fragments, a, b, or c is used; if the charges stay on C terminal fragments, x, y or z is used. The subscript indicates the number of amino acid residues in the fragment. This figure is modified from Ref 61.
1.4 Mass Spectrometry (MS)-based Strategies for Structural Characterization of
Therapeutic MAbs.

Protein characterization, identification and quantification by mass spectrometry-based
methods is based on mass analysis. Due to its superior sensitivity, selectivity and accuracy, MS is
currently the most sophisticated and powerful tool for protein and mAb characterization.\(^{66-68}\)
Currently MS-based strategies for structural characterization of therapeutic mAbs include: top-
down, middle-up and bottom-up approaches.

1.4.1 Top-down approach

In top-down approaches, intact mAbs are directly analyzed without sample digestion. The
intact mAbs are ionized and their masses are first analyzed (the “top” part). Then, the mAb of
interest is isolated (“a gas-phase purification”) and fragmented in the mass spectrometer by
tandem MS (MS/MS). The masses of product ions are measured. With the mass of intact mAbs
and informative fragment ions, mapping of the primary structure and the modification sites (the
“down” part) could be achieved by peptide sequence tag searching method or the ‘\textit{de novo}’
determination method.\(^{56, 69-74}\)

Both ESI and MALDI can be used to introduce intact mAb ions into gas phase in the top-
down approach.\(^{69}\) However, multiply charged precursor ions produced by ESI are fragmented
more efficiently than singly charged precursor ions usually generated by MALDI in tandem MS.
In addition, more fragmentation methods are available in ESI than MALDI. Therefore, ESI is
preferable to MALDI in top-down analysis of mAbs.\(^{72}\)

Since the mass of intact mAbs and the corresponding fragmented product ions need to be
measured and resolved confidently in the top-down approach, high resolution and mass accuracy
mass spectrometer, such as TOF, FTICR or Orbitrap, is usually required to resolve isotopes of coeluting proteins.\textsuperscript{56,72-74} ECD and ETD instead of CID, the most common fragmentation method in mass spectrometry, are used in top-down approach. CID is an energetic dissociation method and preferably breaks bonds of the lowest activation energy. As a result, the labile PTMs (such as phosphorylation bonds and glycosidic bonds) are usually lost in CID since they are weaker than backbone peptide bonds and usually cleaved first. By contrast, the labile PTMs keep intact and produce more complete backbone cleavages in nonergodic fragmentation method, ECD and ETD, making the PTMs localization and sequence variants identification by top-down approach possible. However, ECD is only available at high-end FTICR instruments and ETD has been implemented in both FTMS and LTQ-Orbitrap.\textsuperscript{56,72-74}

In the top-down approach, tandem MS is performed on an intact protein ion, so that the one of the major advantages of top-down approach is that the entire sequence is theoretically available for examination, which allows better complete characterization of the protein and any associated PTMs or unexpected modifications.\textsuperscript{72,74} Since the quantitation in the top-down approach is based on protein level instead of peptide level, another advantage of top-down is the better reliability of protein quantitation.\textsuperscript{56} In addition, the analysis is straightforward and redundant protein digestion and identifications of peptides in bottom-up approach can be avoided.\textsuperscript{74}

However, there are several disadvantages of top-down approach, which limit its application. The disadvantages of the top-down approach include a large sample requirement with high purity and high concentration (0.5–1 μg/μL) of mAb material since the sensitivity and detection limit of MS instrument for intact mAb are much poorer than for digested peptides.\textsuperscript{72} In addition, the required MS instrument, such as an FTICR or Orbitrap mass spectrometer, are very
expensive (>\$750,000). Also, ECD and ETD fragmentation are not efficient enough to cleave large proteins such as mAbs. Due to these advantages, currently top-down approach analysis of mAbs is limited to determination of mAb mass and analysis protein modification on the protein level such as, glycan isoforms, using ESI-TOF or ESI-Q-TOF.  

1.4.2 Middle-up approach

In middle-up approach, mAbs are cleaved into several large fragments and then the fragments are separated by RPHPLC and analyzed by MS. The common middle-up approaches include reduction of disulfide bonds to generate heavy and light chains and limited digestion using papain, pepsin, or Lys-C under native conditions to generate Fab, Fab, F(ab)_2, and Fc. Due to the smaller size of these fragments and lack of heterogeneous glycosylation on most fragments, the requirement for the mass spectrometer is low and can be done on most MS instruments. In addition, since only a few fragments are generated, the separation and analysis of these few fragments are much simpler. Compared to top-down analysis of intact mAb, mass analysis of these smaller fragments are much easier and provide more structural information such as, PTMs, chemical modification, and antibody conjugates characterization. However, location of a single amino acid change, such as deamidation and oxidation, cannot be achieved by a middle-up approach.

1.4.3 Bottom-up approach

In the bottom-up approach, mAbs are digested into peptides using in-gel or in-solution proteolytic digestion. Then the digested peptides (the “bottom” part) are separated by RPHPLC followed by ESI MS (mass of peptides) and MS tandem analysis (peptide sequence) or, the
digested peptides are directly analyzed by MALDI MS. The mAb (the “up” part) identification is based on the MS/MS spectra of peptides through searching against the spectra of theoretical fragmentations of the mAb using software such as SEQUEST in Thermo Bioworks.¹

The size of digested peptides is important for the bottom-up approach. Very small peptides (2-3 amino acid residues) elute in the solvent front due to poor retention on a typical reversed phase column and cannot be analyzed by MS. Too large peptides (>4 kDa) cannot be ionized and fragmented efficiently. The proper enzyme or multiple enzymes should be used to get appropriately sized peptides.¹ The common used enzymes in bottom-up approach includes: trypsin and Lys-C. Trypsin cleaves C-terminus of arginine (R) or lysine (K) residue if they are not followed by proline (P). Lys-C cleaves C-terminus of lysine (K) residue and generates a smaller number of larger peptides compared to trypsin. Trypsin is preferable to Lys-C when CID is used in tandem MS since the peptide size and charge state generated by trypsin digestion is suited for CID fragmentation.⁷⁶ However, Lys-C is used instead of trypsin when ETD is used to fragment peptides because longer peptides generated by Lys-C can retain more charges which are more compatible with ETD.⁷⁷ In addition, Lys-C is more robust and is able to tolerate denaturing condition up to 8 M urea, which is used to improve solubilization of proteins.⁵⁸ Asp-N and Glu-C, which cleave at the N-terminus of aspartic acid and C-terminus of glutamic and aspartic acids, respectively, are occasionally used. To increase the sequence coverage of a mAb, multiple enzyme digestions, either individually or in combination, is reportedly useful.⁷⁸-⁷⁹

CID is the most common fragmentation method used in bottom-up approach. ETD is more useful to characterize disulfide bonds and other posttranslational modifications.

The bottom-up approach is often used to confirm mAb peptide sequence and identify and locate the modifications at amino acid residue level. The common chemical modifications in
mAbs analyzed by the bottom-up approach include: N-terminal pyroglutamic acid cyclization, heavy chain C-terminal lysine processing, oxidation, deamidation, isomerization, non-enzymatic glycation, disulfide linkage, and glycosylation.\(^1\) The application of the bottom-up approach on the characterization of these common modifications on mAbs is described in detailed in Section 1.5. Another advantage of the bottom-up approach is that it is able to detect modifications at relatively low abundance.\(^67\)

However, the bottom-up approach typically cannot achieve 100% mAb peptide sequence coverage if only one run is analyzed due to intrinsic limitations such as poor recovery of large or hydrophobic peptides and loss of peptides during digestion, difficulty of retention of small peptides on reverse phase column, and poor ionization of large peptides.\(^65\) In addition, artifacts could be introduced during digestion. After enzymatic digestion, many small peptides are generated from one mAb molecule, which significantly increases the sample complexity. These two facts make MS analysis more complicated. Furthermore, the bottom-up approach is more time consuming and labor-intensive compared to the other two approaches.\(^1,67\)

1.5 Structural Characterization of Therapeutic MAbs by Bottom-up Mass Spectrometry

1.5.1 Primary Sequence analysis

The bottom-up approach used to identify primary sequence of proteins is also referred to as peptide mapping. As discussed in the previous section, small peptides instead of intact mAbs are analyzed in the bottom-up MS approach. A MAb is a large protein with a complex and defined three dimensional structure. MAbs should be denatured first followed by reduction of disulfide bonds and alkylation of free thiol groups to promote efficient enzymatic digestion. Based on the mAb primary sequence, enzymes are chosen to get appropriate peptide size.
Denaturant (e.g. guanidine hydrochloride, SDS, urea), which is added in the first step, should be removed or diluted before enzymatic (typically trypsin) digestion starts since denaturant will interfere with trypsin digestion and subsequent MS analysis.

After enzymatic digestion, a large number of peptides will be generated from a mAb molecule. Usually RP-HPLC coupled to ESI-MS is used for peptide mapping since RP-HPLC can separate peptides and improve detection limit, therefore increasing peptide sequence coverage. Usually both MS and MS/MS are acquired for peptide identification. CID is the most common fragmentation method in peptide mapping. As discussed in Section 1.3.4, CID breaks amide bonds on the peptide backbone and generates b and y product ions, which are further analyzed by the mass analyzer and produce MS/MS spectrum. Software such as SEQUEST in Thermo Bioworks and MassAnalyzer can be used for automated peptide identification in which experimental MS/MS spectra are searched against theoretical MS/MS spectra. However, manual inspection of the mass spectrometry results are often required to get higher protein sequence coverage and to minimize artifact peaks.

1.5.2 Disulfide-linkage determination

Digestion of mAbs under non-reducing condition with common enzymes such as trypsin or Lys-C generally produces disulfide-linked peptides that can be used to structurally characterize the disulfide linkages. The single disulfide-containing peptide is much easier to identify compared to peptides containing multiple disulfide-linkages. In the case of a mAb, the conventional strategy used for disulfide-linkage identification by MS method includes: (a) digestion of the mAb by an appropriate enzyme under non-reducing and reducing conditions; (b) mapping of non-reducing and reducing digested peptides by LC-MS and LC-MS/MS; (c)
assignment of peptides in both digestion conditions; and (d) assignment of disulfide bonds by comparison of certain peptide ions that disappear in non-reducing digestion but appear in reducing digestion.\textsuperscript{80-82} In this conventional strategy, CID is used for fragmentation in LC-MS/MS. As discussed in Section 1.3.4, the fragmentation of CID is induced by transferring of kinetic energy to internal energy through collision of precursor ions with neutral gas molecules. The internal energy breaks the weakest bond first. It has been reported that the energy required to cleave disulfide bonds (\(\sim 40-70\) kcal/mol) is higher than the one to break amide bonds (\(\sim 25-40\) kcal/mol).\textsuperscript{83} Consequently, CID fragment amide bonds on the peptide backbone and disulfide bonds keep intact. Usually two digestion conditions, reducing and non-reducing, are necessary for assignment of disulfide linkage. This strategy is tedious, time and labor consuming.

Another fragmentation technique, ETD is a powerful method for disulfide bond characterization. During electron transfer process, the electron seems more easily captured by disulfides than by other bonds on peptide backbones.\textsuperscript{77,84-85} Accordingly, ETD preferentially breaks disulfide linkages to produce two polypeptide product ions with the highest abundances in ETD-MS\(^2\) spectrum. By using the online LC-MS with ETD approach, all 16 disulfide bonds in mAbs have been successfully mapped and scrambled disulfides in degraded samples have also been identified easily.\textsuperscript{85} In this approach, the first survey MS scan is followed by CID-MS\(^2\), ETD-MS\(^2\), and CID-MS\(^3\) in a consecutive manner. CID-MS\(^2\) and ETD-MS\(^2\) scan steps were performed on the same precursor ion from first survey MS scan, while CID-MS\(^3\) targets the most abundant ions from ETD-MS\(^2\). To assign disulfide, the precursor ions with the anticipated linked-peptides could be used for initial screening. The likely candidates (from the matched masses) can then be further confirmed by their fragmentation ions produced under either CID or ETD or even CID-MS3 fragmentation.\textsuperscript{85} As an example, the identification of a heavy chain
disulfide, linkage between cysteine in T35 and cysteine T40 of anti-HER2, is shown in Figure 1-5. The disulfide was still linked together after fragmentation in CID (Figure 1-5A), while the disulfide was broken into two separate polypeptides by ETD (P1 and P2 in Figure 1-5B). The dissociated peptide, such as P2 (or T40) could then be further identified by CID-MS3 (Figure 1-5C). Through this LC-MS with ETD strategy, only non-reducing digestion condition is required and it provides more unambiguous identification of disulfide bonds. The advantage of ETD of identification of disulfide bond is more obvious in characterization of multiple intertwined disulfides which have been successfully achieved by Wu et al using this approach.\textsuperscript{77, 84}

**Figure 1-5.** Analysis of a disulfide-linked peptide from anti-HER2 mAb by the combination of CID and ETD fragmentation. 
A. CID-MS\textsuperscript{2} of the precursor ion m/z 770.60 (5+) ion; B. ETD-MS\textsuperscript{2} of the precursor ion m/z 770.60 (5+) ion, the same precursor ion as CID-MS\textsuperscript{2}; C. CID-MS\textsuperscript{3} of the precursor ion m/z 1372.81 (2+) ion, the highest-intensity product ion from ETD-MS2. This figure is taken from Ref 85.
1.5.3 Glycosylation analysis

All mAbs are N-glycosylated at least at the conserved Fc position. Glycosylation analysis can be achieved at two levels: glycopeptide and free glycan levels.

1.5.3.1 Glycopeptide analysis

Glycopeptide analysis by MS allows site-specific analysis of glycosylation. If there are more than one glycosylation sites, glycopeptide analysis could provide specific glycopeptide information on different sites. During CID analysis of glycopeptides, the cleavages occur primarily on glycans since glycosidic bonds are weaker than peptide bonds (Figure 1-6A). As a consequence, little amino acid sequence information is obtained and the glycosylation site is not known. To resolve this problem, removal of N-linked glycans by PNGase F along with stable isotope $^{18}$O labeling at the N-glycosylation site has been developed. PNGase F digestion cleaves the N-glycosidic bond between glycans and Asn residue, converting Asn to Asp and releasing free glycans. When this digestion is performed in an $^{18}$O-labeled water condition, Asn will be converted to $^{18}$O-labeled Asp, which means there is an addition of 3 Da to the corresponding aglycosylated peptide. The presence of Asp $^{18}$O labeling strongly indicates that the peptide was formerly N-glycosylated and the site of glycosylation can also be localized. By MS/MS analysis of the peptides containing Asp $^{18}$O labeling, the amino acid sequence can be confirmed. However, this strategy needs another digestion step besides trypsin digestion and labeling, which is time and labor intensive.

ETD is a powerful technique to characterize glycopeptides. The glycans on glycopeptides remain intact when ETD is applied on the glycopeptide. The cleavages predominantly take place on the peptide backbone and generate c and z ions. As shown in Figure 1-6B , compared to few
peptide backbone cleavages, ETD provides almost complete fragmentation on the peptide backbone and the glycan remains intact.\textsuperscript{63, 87}

Figure 1-6. CID (A) and ETD (B) spectra of glycopeptide (with G1) from anti-HER2 mAb. This figure is taken from Ref 87.
1.5.3.2 Glycan analysis

Glycan analysis provides a glycan profile of the mAb but the glycosylation site cannot be localized. N-linked glycans can be released chemically or enzymatically. PNGase F is the common enzyme for cleavage of N-linked glycans. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) is a widely used approach for profiling released glycans directly or in a derivatized form. However, glycan identification by MALDI-TOF MS is based on the detected mass and no separation is used. Therefore, glycans isomers with identical mass, such as G1 isomers, cannot be differentiated by MALDI-TOF. In addition, the neutral (e.g. G1 glycan) and acidic glycans (e.g. G2+SA glycan) have to be analyzed in different modes.88-89

Released glycans cannot be retained on a regular C18 stationary phase, however a porous graphitized carbon (PGC) stationary phase can strongly bind glycans.90 The mechanism of PGC interaction with glycans is not clear understood but it is generally believed that hydrophobic, polar and ionic interactions play roles in retention of glycans on PGC.90 The most important feature of PGC is that the isoform glycans (e.g. G1 and G1’) can be separated by PGC.91

Another new technique, Agilent mAb-Glyco-Chip, integrates sample preparation and sample analysis into one LC chip. In this approach, intact mAbs are directly injected and followed by online PNGase F digestion. Afterwards, the released glycans are separated by PGC column and analyzed by Q-TOF MS. No sample pretreatment is required and the once tedious and time consuming glycan analysis workflow can be achieved as short as 12 minutes with good results.91
1.5.4 Characterization of common chemical modifications characterization

The common chemical modifications occurring on mAbs are listed in Table 1-5 and the mechanisms of these modifications are discussed in Section 1.2.5.

An example to identify the oxidized peptides from rhGH is shown in Figure 1-7. First, non-oxidized peptide (T12) was identified by accurate mass and the characteristic b and y ions in MS2 spectrum (Figure 1-7A). Due to the addition of one oxygen atom, the mass of oxidized T12 increases by 16 Da compared to nonoxidized T12. The oxidized T12 can be identified by database searching with 16 Da modification of methionine, or manual identification can be achieved by extracting the anticipated m/z of oxidized T12. The matched m/z and charge state (in the insert of Figure 1-7B) indicates the possible candidate and peptide sequence was further confirmed by CID-MS2 spectrum. The position of oxidation, which was localized at the methionine (M) residue, was also identified by the observation of addition of 16 Da on product ions y3, y5, y6, y7 and b7 but not on y1, y2 and b2-5.

Table 1-5. Summary of mass shift of common chemical modifications.

<table>
<thead>
<tr>
<th>Modification</th>
<th>Original AA</th>
<th>Modified AA</th>
<th>Mass shift (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxidation</td>
<td>Met/Trp</td>
<td>Met-O/Trp-O</td>
<td>+16</td>
</tr>
<tr>
<td>Deamidation</td>
<td>Asn</td>
<td>Asp/IsoAsp</td>
<td>+1</td>
</tr>
<tr>
<td>Isomerization</td>
<td>Asp</td>
<td>IsoAsp</td>
<td>0</td>
</tr>
<tr>
<td>Succinimide</td>
<td>Asn/Asp</td>
<td>Suc</td>
<td>-17/-16</td>
</tr>
<tr>
<td>N-terminal pyroglutamic acid cyclization</td>
<td>Gln/Glu</td>
<td>PyroGlu</td>
<td>-17/-18</td>
</tr>
<tr>
<td>C-terminal lysine processing</td>
<td></td>
<td>-Lys</td>
<td>-128</td>
</tr>
</tbody>
</table>
Figure 1-7. LC-MS analysis of the T2 and the oxidized T2 peptides from the tryptic digest of rhGH (Nutropin AQ). A. CID-MS2 of the T2 peptide, m/z 490.2543 (2+) ion; B. MS2 scan of the oxidized T2 peptide, m/z 498.2516 (2+) ion. The FTICR MS measurements of both precursor ions were in the insert. This figure is taken from Ref 92.

Other chemical modifications can be easily identified in the same method as the oxidation but with different mass shifts. However, the characterization of aspartic acid (Asp) containing peptide and corresponding isoaspartic acid (isoAsp) peptide is a challenge since there is no mass difference between Asp and isoAsp, and CID fragmentation of peptides containing these isomers generates similar product ions. One strategy for identification of Asp and isoAsp is the combination of Asp-N digestion and ETD-MS. Asp-N selectively cleaves peptide bonds on the N-terminus of Asp and Cys but keep isoAsp peptides intact. Therefore, isoAsp containing
peptides can be enriched due to the degradation of Asp. Fragmentations of Asp and isoAsp by ETD produce different product ion patterns, in which a pair of reporter ions (c+57 and z-57) is unique to isoAsp. Through this strategy, the Asp and isoAsp containing peptides can be differentiated confidently.35

1.5.5 Quantitation

MS-based quantitation can be divided into absolute and relative quantitation based on exact or relative amount of protein or peptide is determined.93

1.5.5.1 Absolute quantitation (AQUA)

In the absolute quantitation method, a peptide of interest is synthesized with isotopically labeled amino acids, which increases the mass of the synthetic peptide by 6-10 Da compared to the targeted peptide. A known amount of synthetic and isotopically labeled peptide is added to the sample. The sample is then digested and analyzed by LC-MS. Since the synthetic and native peptides have the same amino acid sequence, they should have identical retention time and ionization characteristics but different masses. The absolute amount of native peptide can be determined from the responses of the spiked-in synthetic peptides.93-94

This absolute quantitation (AQUA) method is a highly accurate and selective measurement of the peptide of interest. However, because of the requirement of labeled peptide and depending on the distribution of the peptides of interest and the duty cycle of the mass spectrometer, only one or a few proteins may be analyzed in a single run.95-96
1.5.5.2 Relative quantitation

Relative quantitation can be further divided into isotope–labeling method or label-free method.

In the isotope-labeling method, isotopic labels can be incorporated during protein biosynthesis (SILAC), enzyme digestion in the presence of H$_2^{18}$O, through chemical labeling some active groups (e.g. free amino groups; lysine-specific labeling; phosphopeptide labeling, N- or C-terminus of peptides), or through introduction of isotopic tags (ICAT, a thiol-specific group; mTRAQ, an amine-reactive group; ALICE, acid-labile isotope-coded). After labeling, isotope-labeled or non-labeled samples are pooled together and analyzed in a single run. The isotope-labeled peptide and non-labeled counterpart should behave in the same way under both LC and MS analysis but they can be distinguished by their mass difference. The relative abundances of labeled and non-labeled pairs can be accurately determined by their peak intensities or peak area.

In label-free quantitation method, each sample is prepared and analyzed separately. After classic LC-MS or MS/MS is done, the relative protein abundance is determined by comparing the spectra counting of the same protein, peak intensity or peak area of the same peptide.

Since no labeling is required, label free quantitation is less expensive and much easier to perform. In addition, an unlimited number of samples can be compared and the relative amount of any kind of protein can be determined. However, labeling quantitation is less accurate compared to other quantitation methods and each sample has to be analyzed separately.
1.6 Conclusions

Comprehensive characterization of therapeutic monoclonal antibodies (mAbs) is critical for drug integrity and regulatory filing. The development of high mass accuracy and mass resolution MS instruments with versatile fragmentation methods makes the global and fine structural characterization of mAbs more easily and confidently. A range of LC-MS methods have been developed for comprehensive characterization of primary structure, disulfides, glycostructures, and other modifications. However, to better support mAb characterization, multiple and complementary LC-MS methods should be used.
1.7 References


31. Weissbach, H.; Resnick, L.; Brot, N. Methionine sulfoxide reductases: history and cellular role in protecting against oxidative damage. *Biochim Biophys Acta* 2005, 1703 (2), 203-
12.


Cytotechnology 2007, 53 (1-3), 121-5.


57. Hoffmann, E. d.; Stroobant, V., Mass Spectrometry: Principles and Applications. 3 ed.;


78. Choudhary, G.; Wu, S. L.; Shieh, P.; Hancock, W. S. Multiple enzymatic digestion for


Chapter 2
Characterization and Comparison of Disulfide Linkages and Scrambling Patterns in Therapeutic Monoclonal Antibodies: Using LC-MS with Electron Transfer Dissociation

Publication:
2.1 Abstract

The disulfides in three monoclonal antibodies (mAb), the anti-HER2, anti-CD11a, and GLP-1 with IgG4-Fc fusion protein, were completely mapped by LC-MS with the combination of electron-transfer dissociation (ETD) and collision induced dissociation (CID) fragmentation. In addition to mapping the 4 inter- and 12 intrachain disulfides (total 16), the identification of scrambled disulfides in degraded samples (heat-stress) was achieved. The scrambling was likely attributed to an initial breakage between the light (Cys 214) and heavy (Cys 223) chains in anti-HER2, with the same observation found in a similar therapeutic mAb, anti-CD11a. On the other hand, the fusion antibody, with no light chain but containing only two heavy chains, generated much less scrambling under the same heat-stressed conditions. The preferred sites of scrambling were identified, such as the intrachain disulfide for CxxC in the heavy chain, and the C194 of the heavy chain pairing with the terminal Cys residue (C214) in the light chain. The interchain disulfides between the light and heavy chains were weaker than the interchain disulfides between the two heavy chains. The relative high abundance ions observed in ETD provided strong evidence for the linked peptide information, which was particularly useful for the identification of the scrambled disulfides. The use of sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) helped the separation of these misfolded proteins for the determination of scrambled disulfide linkages. This methodology is useful for comparison of disulfide stability generated from different structural designs and providing a new way to determine the scrambling patterns, which could be applied for those seeking to determine unknown disulfide linkages.
2.2 Introduction

Therapeutic monoclonal antibodies (mAb) can bind to specific epitopes on surface receptors and disrupt the interaction between signaling molecules and surface receptors, thereby inhibiting cancers or other inflammatory diseases.\(^1\)\(^-\)\(^2\) For example, anti-HER2 mAb (Herceptin) binds to the HER2 receptor, disrupting the receptor complex on the breast cancer cell surface and thus blocking the corresponding signal transduction cascade. The binding domain (Fab) and IgG1 construct (Fc) of anti-HER2 mAb is connected with disulfides, which link light and heavy chains, and also the two heavy chains (see Figure 2-1A). There are a total of sixteen disulfides in one anti-HER2 mAb molecule, 12 intra- and 4 interchain disulfides. Among the four interchain disulfides, two link together the heavy chains and the other two connect the light and heavy chains. The C-terminal end of the light chain (Cys 214) forms a disulfide bond with Cys 223 in the heavy chain, which thus connects the light and heavy chains. Cys 229 and Cys 232 in one heavy chain link with the other heavy chain at the same Cys 229 and Cys 232 sites to form two parallel interchain disulfides between the two heavy chains. These four polypeptide chains (two heavy and two light chains) are connected by these four interchain disulfide bonds to form a tetramer, which is required for mAb to function effectively,\(^3\)\(^-\)\(^8\) and incomplete formation of disulfide bonds could cause loss of drug efficacy.\(^3\),\(^9\) Thus, it is important to characterize all these disulfides and establish the presence of correct connectivity to ensure drug function and quality.
The disulfides in mAb, similar to other classes of protein, are conventionally characterized by Edman sequencing and tryptic mapping with reduced and nonreduced conditions.\(^4\),\(^7\),\(^10\)-\(^11\) Due to its superior selectivity and sensitivity, mass spectrometry (MS) coupled with liquid chromatography has recently become the choice for characterization of disulfides.\(^12\)-\(^15\) The MS approach often involves the characterization of disulfide-dissociated peptides with reduction and then the determination of disulfide bonds without reduction through collision induced dissociation (CID).\(^7\),\(^10\)-\(^11\),\(^16\) However, these approaches can be tedious and may not be feasible for multiple intertwined disulfides in one peptide. Recently, Wu et al. used online LC-MS with electron-transfer dissociation (ETD) approach to successfully identify disulfide linkages and multiple intertwined disulfides in several recombinant therapeutic proteins.\(^17\)-\(^18\)

It has been reported that disulfide scrambling can occur, particularly at alkaline pH or in the presence of free cysteine residues.\(^7\),\(^19\)-\(^21\) However, the determination of the scrambling sites
in mAb has not been studied extensively but is of importance in determining product stability as well as in selection of proper formulation. In this report, the disulfides of three different mAb, anti-HER2 and anti-CD11a (both with IgG1 construct) (Figure 2-1A), and a fusion protein (GLP-1 with IgG4 Fc domain) (Figure 2-1B) are characterized with the LC-MS with ETD in combination with efficient separation and multienzyme approach. This detailed characterization pinpoints the scrambled sites, which could then shed light on the relative strength of specific disulfides, the preferred sites of scrambling, and the scrambling process during a heat-stress study.

2.3 Experimental Section

2.3.1 Materials

Three therapeutic monoclonal antibodies were used in this study: (1) anti-HER2 (Herceptin, Genentech, South San Francisco, CA), a liquid formulation product (22 μg/μL); (2) anti-CD11a (Raptiva, Xoma, Berkeley, CA), a lyophilized powder; (3) GLP-1 with IgG4 Fc fusion protein (CNTO 736, Centocor R&D, Radnor, Pennsylvania), a liquid formulation product (10 μg/μL). Herceptin and Raptiva were the gift from Genentech, and CNTO 736 was the gift from Centocor. Trypsin (sequencing grade) was purchased from Promega (Madison, WI). Guanidine hydrochloride, ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide (IAA), and formic acid (FA) were from Sigma-Aldrich (St. Louis, MO). LC-MS grade water was from J.T. Baker (Phillipsburg, NJ), and HPLC grade acetonitrile from ThermoFisher Scientific (Fairlawn, NJ). Precast NuPAGE 4–12% Bis-Tris sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels, NuPAGE MES SDS running buffer (20X), NuPAGE LDS sample buffer (4X), and SimplyBlue SafeStain were all from Invitrogen (Carlsbad, CA).
2.3.2 Heat-Stressed Conditions

For the heat-stressed stability study, an aliquot of 20 μg monoclonal antibody (i.e., 2 μL) was mixed with 5 μL NuPAGE LDS Sample Buffer (4X) (pH 7.8) and 13 μL NuPAGE MES SDS Running Buffer (1X) to achieve the final volume of 20 μL. The mixtures were incubated at 70 °C for 0, 10, 15, 20, 25, 30, and 35 min, separately. For comparison purposes (under reduced conditions), an aliquot of 20 μg monoclonal antibody (i.e., 2 μL) was mixed with 2 μL DTT, 5 μL NuPAGE LDS sample buffer (4X), and 11 μL NuPAGE MES SDS running buffer (1X) to achieve the final volume of 20 μL. The mixtures were incubated at 90 °C for 10, 15, 20, 25, 30, and 35 min, separately. After incubation, the mixtures (20 μL) were loaded onto an SDS-PAGE gel (4–12% Bis-Tris polyacrylamide) at 200 V for 40 min and then stained with Coomassie blue.

2.3.3 Enzymatic Digestion

In-gel digestion: bands of interest were excised and destained by alternating use of acetonitrile and 0.1 M ammonium bicarbonate (pH 7.8) until no visible color was present. The destained gel pieces were completely dehydrated with acetonitrile and then dried in a SpeedVac (Labconco, Centrivap Cold Trap). Trypsin or Lys-C (1:50 w/w) was added to just cover the dried gel pieces. After 45–50 min incubation at 4 °C, the remaining supernatant was removed and buffer without trypsin was added to cover gel pieces for overnight (~12 h) at 37 °C. The digested peptides were then extracted with 5% formic acid and ACN (1:2, v/v) and then dried down until only 10 μL were left (if not, the samples were reconstituted to 10 μL for subsequent analysis).
2.3.4 LC-MS

An aliquot of 5 μL of the digest was injected onto a self-packed C18 column (Magic C18, 200 Å pore and 5 μm particle size, 75 μm i.d. × 15 cm) (Michrom Bioresources, Auburn, CA). Mobile phase A was 0.1% formic acid in water, and mobile B was 0.1% formic acid in acetonitrile. The peptides were eluted at 200 nL/min using 2% B to 65% B in 65 min, then from 65% B to 80% B for 10 min. The MS experiments were performed on an LTQXL with ETD mass spectrometer (Thermo Fisher Scientific, San Jose, CA). The instrument was operated in a data-dependent mode: the first survey MS (scan 1) from \( m/z \) 400 to 1900 followed by three consecutive ion activation steps: CID-MS\(^2\) (scan 2), ETD-MS\(^2\) (scan 3), and CID-MS\(^3\) (scan 4). The CID-MS\(^2\) and ETD-MS\(^2\) scan steps were performed on the same precursor ion (see the details in our previous paper\(^17\)). The precursor ions of these disulfides were also measured by an LTQ-FT MS instrument (ThermoFisher Scientific).

2.4 Results and Discussion

In this section, the digestion strategy, LC-MS method, and principles for assignment of scrambled disulfides will be first described, followed by the analysis of disulfides for three different mAb in a heat-stressed study.

2.4.1 Digestion Strategy

Identification of a single disulfide linkage is straightforward because there is usually only one possibility for connection. So, enzymes, which can cut proteins to the peptides containing only single disulfide, are desired. In addition, peptide sizes are preferred at 1–5 kDa since the recovery and electrospray ionization efficiency can be a problem for larger peptides, while
smaller peptides may not retain well on a typical reversed phase column. Thus, the selection of proper enzymes or the use of multiple enzymes should be considered for the size adjustment. In addition, the presence of a disulfide bond may prevent cleavage of neighboring residues (steric hindrance). Therefore, during data analysis, additional miscleavages need to be considered. In this work, trypsin was found to be sufficient for in-gel digestion except for the need to use Lys-C for the analysis of a small disulfide-linked peptide (see Table 2-1). However, Lys-C plus trypsin digestion is preferred than trypsin only (produced less miscleavages) for in-solution digestion (e.g., protein digestion without SDS-PAGE). The reason could be that the cleavage of the disulfide-linked proteins may need Lys-C to trim the size for reducing steric hindrance first.

**Table 2-1.** Theoretical tryptic peptides of anti-HER2 mAb without reduction, including linked disulfides (lines), cysteine position (parentheses), and molecular weight.

<table>
<thead>
<tr>
<th>Tryptic peptides in heavy chain</th>
<th>MW= 2384.0934</th>
<th>MW= 7916.9351</th>
<th>MW= 5454.8148</th>
<th>MW= 2328.1135</th>
<th>MW= 3844.8393</th>
<th>MW= 4819.2579</th>
<th>MW= 3555.7647</th>
<th>MW= 912.3587</th>
<th>MW= 1260.5021</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T1</strong></td>
<td>20-30</td>
<td>LSCAASGFNIK (Cys 22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T11</strong></td>
<td>88-98</td>
<td>AEDTAVYYCSR (Cys 96)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T14</strong></td>
<td>137-150</td>
<td>STSGTAALGCLVK (Cys 147)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T15</strong></td>
<td>151-213</td>
<td>DYLPEPVTSNWGALTSGVHTFPALQSSGLYSLSVTVVSSSLGQTQYICNVHKPSNTK (Cys 203)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T19</strong></td>
<td>226-251</td>
<td>THITCPPCPAPELLGGSFLFPPKPK (Cys 229; Cys 232)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T19</strong></td>
<td>226-251</td>
<td>THITCPPCPAPELLGGSFLFPPKPK (Cys 229; Cys 232)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T21</strong></td>
<td>259-277</td>
<td>TPEVTCVVVDVSHEDEPVK (Cys 264)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T27</strong></td>
<td>324-325</td>
<td>CK (Cys 324)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T35</strong></td>
<td>364-373</td>
<td>NQVSLTCLVK (Cys 370)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T40</strong></td>
<td>420-442</td>
<td>WQQGQVFCG7SVMHELHNYTQK (Cys 428)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T</strong></td>
<td>19-24</td>
<td>VTTTCR (Cys 23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T</strong></td>
<td>67-103</td>
<td>SGDFITLTTSSLQEDAPFYCQQHYTTPFGQG (Cys 88)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T7</strong></td>
<td>127-142</td>
<td>SGTSVXCLNNFYP (Cys 134)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T7</strong></td>
<td>191-207</td>
<td>VYACEVTHQGLSSVPVT (Cys 194)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T18(H)</strong></td>
<td>222-225</td>
<td>SCDK (Cys 223 from H chain)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T19(L)</strong></td>
<td>212-214</td>
<td>GEI (Cys 214 from L chain)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T18(H)</strong></td>
<td>222-225</td>
<td>SCDK (Cys 223 from H chain)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T18T19(L)</strong></td>
<td>208-214</td>
<td>SFNRGEI (Cys 214 from L chain)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lys-C digested peptide between heavy and light chain

<table>
<thead>
<tr>
<th>MW= 912.3587</th>
<th>MW= 1260.5021</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T18(H)</strong></td>
<td>222-225</td>
</tr>
<tr>
<td><strong>T18T19(L)</strong></td>
<td>208-214</td>
</tr>
</tbody>
</table>
2.4.2 LC-MS with CID and ETD

The disulfide-linked peptides are characterized by the combination of CID and ETD; specifically, the disulfides are preferentially dissociated into two polypeptides by ETD while CID generates mainly peptide backbone cleavage (with the disulfides intact). This strategy can be further enhanced by CID-MS\textsuperscript{3} (after ETD) on the dissociated peptides or the charge-reduced species if necessary, as described in our previous papers.\textsuperscript{17-18}

2.4.3 Disulfide Assignment

When two polypeptides are linked by a disulfide bond, the assignment of such peptides could not be achieved by typical approaches such as using either Sequest or Mascot software. Often, the precursor ions with the anticipated linked-peptides could be used for initial screening. The likely candidates (from the matched masses) can then be further confirmed by their fragmentation ions produced under either CID or ETD or even CID-MS\textsuperscript{3} fragmentation.

As an example, the identification of a heavy chain disulfide, linkage between T35 and T40 of anti-HER2, is shown in Figure 2-2. The disulfide was still linked together after fragmentation in CID (Figure 2-2A), while the disulfide was broken into two separate polypeptides by ETD (P1 and P2 in Figure 2-2B). The dissociated peptide, such as P2 (or T40) could then be further identified by CID-MS3 (Figure 2-2C). The high intensity of P1 and P2 in ETD (preferred cleavage) was likely due to the free electron being more easily captured by sulfurs (i.e., disulfides) than by peptide backbones (i.e., N–Cα bonds) during the electron capture or transfer process (see our previous references for more details).\textsuperscript{17-18}
Figure 2-2. Using CID-MS2 (A), ETD-MS2 (B), and CID-MS3 (C) to analyze a disulfide-linked peptide (anti-HER2).
2.4.4 Heat-Stressed Study

Anti-HER2 samples were heated (without reduction with DTT) at 70 °C for 0, 10, 15, 20, 25, 30, and 35 min and then separated by SDS-PAGE (Figure 2-3 left panel). As shown, the intensities of additional bands were increased with longer heating times. However, these same samples under reduction (DTT treatment) were collapsed into two major bands (heavy and light chains) regardless of the heating time. Thus, the additional bands, while disappearing after reduction, could be related to mAb structures with different combinations of disulfide linkages.

Figure 2-3. Heat-stressed study for mAb (anti-HER2) using SDS-PAGE separation with nonreduced (A) and reduced (B) conditions.
2.4.5 Scrambled Disulfides

The disulfides with the known and expected linkages were identified in band A1 at the 0 time point (see Table 2-2). However, the additional bands generated at different heating times (i.e., A1a, A1b, A1c, and etc. in Figure 2-3A) were found to contain different disulfide connectivity. As an example, a scrambled disulfide was identified in band A1d. In this band, in addition to the anticipated disulfide linkage, T35 with T40, the scrambled linkage of T35 with T11 was identified. The disulfide-linked peptide was dissociated into two polypeptides by ETD (T11 or P1 and T35 or P2 in Figure 2-4B). The observation of two relative high abundance polypeptides (T11 and T35) provided strong evidence of these two linked peptides, which was further confirmed by CID-MS3 (Figure2-4C) or CID-MS2 (Figure 2-4A). However, for the robust identification of scrambled disulfides, a combination of all possible cysteine-containing tryptic peptides for the mAb sequence need to be considered. After making a mass list of all possible combinations with different charge states, the likely precursors were then further examined by their corresponding ETD and CID fragmentation ions. The high abundant ETD ions were typically examined first since the disulfide-dissociated peptides were often produced with relative high intensities. If these disulfide-dissociated peptides in ETD matched with any of the cysteine-containing tryptic peptides, their corresponding CID-MS2 and CID-MS3 (after ETD) were then examined for further confirmation. Overall, the ETD spectrum provided an initial match for the presence of scrambled disulfides from the list of potential precursor ions. The examples of the light and heavy chain scrambling were displayed in the Supporting Information. Two examples of the light chain scrambling in band A2 were displayed in the Supporting Information; see Figure 2-S1 (A and B) for the light chain scrambling in band A2 and Figure 2-S2 (A and B) for the heavy chain scrambling in band A1a. For simplicity, only the ETD spectra
were displayed for the identification. The identification of different combinations of light and heavy chains (along with their corresponding disulfide linkages and major scrambled disulfides) in these additional bands was listed in Table 2-3. These assignments were based on the disulfides and peptide sequences found at their corresponding gel bands.

From these heat-stressed studies, the initiation of the scrambling could be proposed to occur at the linkage between the light and heavy chain. Under storage at mild temperatures (i.e., 4 °C for 3 yrs in a liquid formulation), a minor band at the light chain position was observed (band A2 at 0 time point in Figure 2-3A). This minor band was identified by peptide sequence as a breakage of light chain from the intact mAb. There is only one linkage between the light (Cys 214) and heavy chains (Cys 223) (see Figure 2-1A). Thus, this breakage should produce the nuclei (free cysteine) for initiation of scrambling, which was manifested with different combinations of scrambled disulfides and increased more with longer heating times.

Figure 2-4. Using CID-MS2 (A), ETD-MS2 (B), and CID-MS3 (C) to analyze a disulfide scrambled peptide (anti-HER2).
Table 2-2. Assignment of all correct disulfides in band A1 from the time point at 0 min of Figure 2-3A with the observed m/z at the charge of the highest intensity and the retention time.

<table>
<thead>
<tr>
<th>correct disulfide in heavy chain</th>
<th>observed m/z (charge)</th>
<th>highest intensity</th>
<th>retention Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2 with T11</td>
<td>796.0530 (3+)</td>
<td>$1.38 \times 10^7$</td>
<td>26.50</td>
</tr>
<tr>
<td>T14 with T15</td>
<td>1132.7753 (7+)</td>
<td>$2.73 \times 10^6$</td>
<td>37.47</td>
</tr>
<tr>
<td>T19 with T19</td>
<td>780.7271 (7+)</td>
<td>$1.11 \times 10^5$</td>
<td>35.76</td>
</tr>
<tr>
<td>T21 with T27</td>
<td>777.3900 (3+)</td>
<td>$2.25 \times 10^7$</td>
<td>25.72</td>
</tr>
<tr>
<td>T35 with T40</td>
<td>770.3930 (5+)</td>
<td>$1.07 \times 10^7$</td>
<td>29.05</td>
</tr>
</tbody>
</table>

To correct disulfide between heavy and light chains

| T18 (H) with T18T19 (L)\(^a\) | 631.3666 (2+) | $6.13 \times 10^5$ | 23.83               |

To correct disulfide in light chain

| T2 with T10                     | 965.4877 (5+)  | $9.76 \times 10^6$ | 34.11               |
| T10 with T17                    | 712.5760 (5+)  | $2.38 \times 10^7$ | 31.97               |

\(^a\) Obtained from Lys-C digestion.

Table 2-3. Assignment of different chain combinations along with their corresponding disulfide linkages and major scrambled disulfides in the gel bands from the time point at 35 min of Figure 2-3A.

<table>
<thead>
<tr>
<th>band no.</th>
<th>chain combination(^a)</th>
<th>disulfides</th>
<th>observed m/z (charge)</th>
<th>highest intensity</th>
<th>retention time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>2 heavy chains + 2 light chains (144 kDa)</td>
<td>T17 (L) with T18T19 (L)</td>
<td>658.0743 (4+)</td>
<td>$4.76 \times 10^6$</td>
<td>23.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T18T19 (H)(^c) with T18T19 (H)(^c)</td>
<td>791.3063 (8+)</td>
<td>$1.41 \times 10^6$</td>
<td>36.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T2 (H) with T21 (H)</td>
<td>798.4250 (4+)</td>
<td>$8.35 \times 10^5$</td>
<td>29.74</td>
</tr>
<tr>
<td>A1a</td>
<td>2 heavy chains + 1 light chain (121 kDa)</td>
<td>T18T19 (L)(^b) with T18T19 (H)(^c)</td>
<td>795.6248 (5+)</td>
<td>$2.76 \times 10^4$</td>
<td>32.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T11 (H) with T18T19 (H)(^c)</td>
<td>740.5518 (6+)</td>
<td>$1.71 \times 10^5$</td>
<td>32.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T11 (H) with T18T19 (H)(^c)</td>
<td>740.5517 (6+)</td>
<td>$4.10 \times 10^5$</td>
<td>32.60</td>
</tr>
<tr>
<td>A1b</td>
<td>2 heavy chains (98 kDa)</td>
<td>T18T19 (H)(^c) with T18T19 (H)(^c)</td>
<td>791.4333 (8+)</td>
<td>$1.41E \times 10^6$</td>
<td>36.24</td>
</tr>
<tr>
<td>A1c</td>
<td>1 heavy chain + 1 light chain (72 kDa)</td>
<td>T18T19 (L)(^b) with T18T19 (H)(^c)</td>
<td>795.4141 (5+)</td>
<td>$5.71 \times 10^5$</td>
<td>32.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T11 (H) with T19 (H)</td>
<td>683.1274 (4+)</td>
<td>$8.37 \times 10^5$</td>
<td>34.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T11 (H) with T35 (H)</td>
<td>595.8024 (4+)</td>
<td>$2.38 \times 10^6$</td>
<td>27.58</td>
</tr>
<tr>
<td>A1d</td>
<td>1 heavy chain (49 kDa)</td>
<td>T11 (H) with T18T19 (H)(^c)</td>
<td>740.5486 (6+)</td>
<td>$2.27 \times 10^6$</td>
<td>32.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intra disulfide in T19 (H)</td>
<td>683.1265 (4+)</td>
<td>$5.55 \times 10^6$</td>
<td>34.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T17 (L) with T18T19 (L)</td>
<td>658.0746 (4+)</td>
<td>$4.76 \times 10^6$</td>
<td>23.86</td>
</tr>
<tr>
<td>A2</td>
<td>1 light chain (23 kDa)</td>
<td>T2 (L) with T17 (L)</td>
<td>628.0838 (4+)</td>
<td>$3.18 \times 10^6$</td>
<td>25.85</td>
</tr>
</tbody>
</table>

\(^a\) The assignments were based on the disulfides and peptide sequences found at their corresponding gel bands. \(^b\) Not cleaved (miscleavage) between T18 and T19 of the light chain. \(^c\) Not cleaved (miscleavage) between T18 and T19 of the heavy chain; in which T19 formed an intrachain disulfide.
2.4.6 Stability for Other Types of mAb

As examples of other therapeutic mAb, anti-CD11a in a lyophilized powder form was investigated with the same analytical approach. Except for the presence of different amino acid sequences in the Fab region, anti-CD11a has the same numbers and types of disulfides as anti-HER2 (same IgG1 type). After heating, anti-CD11a generated a very similar band pattern as anti-HER2 under nonreduced conditions (see Figure 2-5A) and similar scrambling was also found in the additional bands (see Tables S1A and S1B in the Supporting Information). Another fusion antibody was also investigated under the same conditions. In this fusion antibody, the light chain was removed and the heavy chain in the Fab region was replaced with a GLP-1 sequence (CNTO 736). The heavy chains in the IgG4 Fc domain have the same types and numbers of two parallel interchain disulfides to connect the two heavy chains and four intrachain disulfides within the heavy chains (see Figure 2-1B). After a similar heating stability study, the additional bands observed in SDS-PAGE were significantly reduced for the fusion antibody (compare Figure 2-5B to 2-5A). This reduction was attributed again to the absence of light chains, which likely initiated the scrambling in anti-HER2 and anti-CD11a. The results also suggest that the interchain disulfides between the light and heavy chains are weaker than the interchain disulfides between the two heavy chains. In this heat-stress study, the dissociated heavy chains migrated at the molecular weight corresponding to one-half size of the intact molecule (band 2 vs band 1 in Figure 2-5B). This dissociated heavy chain was found to contain an intrachain disulfide bond connected between the Cys 229 and Cys 232 (see Figure 2-6). The assignment was based on the accurate precursor mass measurement (the loss of 2H from the peptide containing 2 cysteines), as well as the corresponding fragmentation patterns of the precursor ion in CID-MS2 (Figure 2-6A) and ETD-MS2 (Figure 2-6B). As shown, the fragment ions in this peptide were derived
from outside of the disulfide ring. Particularly, the fragmentation at the proline residues by CID (y ions) was observed with high abundance near but not within the intrachain disulfide. The disulfide exchange from inter- to intrachain disulfide in the CxxC region of the heavy chain could also be found in anti-HER2 and anti-CD11a mAb, the latter two being IgG1 type Fc as compared to the IgG4 type for the fusion protein (see Figures 2-S3 and 2-S4 in the Supporting Information). For the fusion protein, this disulfide exchange in the CxxC seemed to be the preferred scrambled site as compared to other scrambled sites (judged by the 1−2 order magnitude higher intensity; see Supporting Information in Tables 2-S2A and 2-S2B). Since different peptides could have different ionization efficiency, we also used relative comparison to compare samples at different heating times (0−35 min as shown in Figure 2-5B). Indeed, the relative ratio of the intrachain disulfide of CxxC (ratio against an internal peptide without disulfides) was increased much faster and higher than any other scrambling disulfides with longer heating time. For mAb with the light and heavy chains (anti-HER2 and anti-CD11a), the scrambling at CxxC was then the second most preferred site relative to the disulfide linkage between the light and heavy chains. The formation of the intrachain disulfide (CxxC) caused the two heavy chains breaking into a single heavy chain (molecular weight shift as shown in the band 2 of Figure 2-5B or band A1d of Figure 2-3A). For this scrambled site, neither the two free cysteines nor the cross-link (i.e., between C229 of the heavy chain 1 with C232 of the heavy chain 2 and C232 of the heavy chain 1 with C229 of the heavy chain 2) could be observed.

From characterization and comparison of these scrambled disulfides, a focus should be placed on the disulfide connection between the light and heavy chains and also the disulfide connection between the two heavy chains in the manufacturing processes or during formulation development. Any inconsistent recovery (intensity) or retention time for these two disulfide-
linked peptides would provide a warning signal for scrambling. A time course study with relative comparison, as shown here, should provide a good indication of disulfide stability for mAb development. In addition, any artifact (e.g., scrambling caused by digestion conditions) can also be determined and eliminated from the time course study.

Figure 2-5. Heat-stressed study for anti-CD11a (A) and GLP-1 with IgG4-Fc fusion mAb (B) using SDS-PAGE separation without DTT reduction.
Figure 2-6. CID and ETD analysis of an inter- to intrachain disulfide scrambled peptide in GLP-1 with IgG4-Fc fusion mAb.

The fragment ions containing the intra-chain disulfide were labeled with red; otherwise, they were labeled with black.
2.5 Conclusions

Our method successfully identified the scrambled disulfides with linkage sites in mAb. The scrambling that we observed in our heat-stressed study was likely due to the initial breakage between the light (Cys 214) and heavy (Cys 223) chains. For proteins containing only Fc region, the breakage of interchain disulfides between the two heavy chains (CxxC) was facilitated by the formation of the intrachain disulfide linkage.

In this method, the combination of ETD with CID fragmentation successfully mapped the correct and scrambled disulfides of three different mAb, with the relative high abundant ions observed in ETD providing strong evidence of the linked peptide. The use of SDS-PAGE to separate the misfolded proteins helped to decide the scrambled disulfide linkages. The relative strength of disulfides with their correct linkages and preferred scrambling sites could be examined from a time course study.

The higher order structures such as disulfide mapping in proteins are usually difficult task to characterize comprehensively because of the technique challenge. The new ETD mass spectrometry technology, developed recently, allowed us to investigate this important phenomenon. The method in this study provided a new way to determine the scrambling patterns, which could be applied for those seeking to determine unknown disulfide linkages. This method also set the stage to compare the higher order structure such as disulfide linkages for biotechnology industry at different development stages or biosimilar comparison.
2.6 References


17. Wu, S. L.; Jiang, H.; Lu, Q.; Dai, S.; Hancock, W. S.; Karger, B. L. Mass spectrometric determination of disulfide linkages in recombinant therapeutic proteins using online LC-MS with


2.7 Supporting Information

Figure 2-S1. ETD analysis of scrambled disulfides at N-terminal (A) and C-terminal (B) region of the light chain of anti-HER2 (band A2 from Figure 2-3A).

Figure 2-S2. ETD analysis of scrambled disulfides at Fab (A) and Fc (B) region of the heavy chain of anti-HER2 (band A1a from Figure 2-3A).
Figure 2-S3. Intra-disulfide in CxxC (T19) from Anti-HER2.

Figure 2-S4. Intra-disulfide in CxxC (T19) from Anti-CD11a.
Table 2-S1A. Theoretical tryptic peptides of anti-CD11a mAb without reduction, including linked disulfides (lines), cysteine position (parentheses), and molecular weight.

<table>
<thead>
<tr>
<th>Tryptic peptides in heavy chain</th>
<th>MW=</th>
<th>Tryptic peptides in light chain</th>
<th>MW=</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2 20-38 LSCAAASGYSFTGHWMNWVR (Cys 22)</td>
<td>3430.5115</td>
<td>T2 19-24 VTTITGR (Cys 23)</td>
<td>5311.4548</td>
</tr>
<tr>
<td>T11 88-98 AEDTAVYVYCAR (Cys 96)</td>
<td></td>
<td>T7 62-103 GTK (Cys 88)</td>
<td>5555.7647</td>
</tr>
<tr>
<td>T14 138-151 STSGGTAALGQLVK (Cys 148)</td>
<td></td>
<td>T10 127-142 SGTASVVCLLLNNFYPR (Cys 134)</td>
<td></td>
</tr>
<tr>
<td>T15 152-214 DVFPEPVTSENSGALTSGVHTFPAVLQSSGLYSLSSVV TVPSSSLGTQTYICNVNHKSNTK (Cys 204)</td>
<td>7916.9351</td>
<td>T17 191-207 VYACEVTHQGLSSPVTK (Cys 194)</td>
<td>3844.8393</td>
</tr>
<tr>
<td>T19 227-252 THTCPPCPAPELLGGPSVFLFPKPK (Cys 230; Cys 233)</td>
<td>5454.8148</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T21 260-278 TPEVTVVVDVSHEDPEVK (Cys 265)</td>
<td></td>
<td>T27 325-326 CK (Cys 324)</td>
<td></td>
</tr>
<tr>
<td>T27 325-326 CK (Cys 324)</td>
<td></td>
<td>T35 365-374 NQVSLTCVLKV (Cys 371)</td>
<td></td>
</tr>
<tr>
<td>T40 421-443 WQOGNVSDFCSVMHEALHNHYTQK (Cys 428)</td>
<td>3844.8393</td>
<td>T40 421-443 WQOGNVSDFCSVMHEALHNHYTQK (Cys 428)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T40 421-443 WQOGNVSDFCSVMHEALHNHYTQK (Cys 428)</td>
<td></td>
</tr>
<tr>
<td>T19(L) 212-214 GEC (Cys 214 from L chain)</td>
<td>912.3587</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T18(L) 208-214 SFNREGEC (Cys 214 from L chain)</td>
<td>1260.5021</td>
</tr>
</tbody>
</table>

Lys-C digested peptide between heavy and light chain

T18(H) 223-226 SCDK (Cys 224 from H chain) | MW= 912.3587 |

T18T19(L) 208-214 SFNRGERC (Cys 214 from L chain) | MW= 1260.5021 |
Table 2-S1B. The assignment of different chain combinations along with their corresponding disulfide linkages and major scrambled disulfides in the gel bands from the time point at 35 min of Figure 2-7A.

<table>
<thead>
<tr>
<th>Band #</th>
<th>chain combination&lt;sup&gt;a&lt;/sup&gt;</th>
<th>disulfides</th>
<th>observed m/z (charge)</th>
<th>highest intensity</th>
<th>retention time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>1L</td>
<td>T17 (L) with T18T19 (L)&lt;sup&gt;b&lt;/sup&gt; T2 (L) with T17 (L)</td>
<td>658.33 (4+) 628.42 (4+)</td>
<td>6.63 x 10&lt;sup&gt;5&lt;/sup&gt; 1.84 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td>29.55 32.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A1d</td>
<td>T11 (H) with T14 (H) Intra disulfide in T19 (H)</td>
<td>842.25 (3+) 683.33 (4+)</td>
<td>2.06 x 10&lt;sup&gt;5&lt;/sup&gt; 2.95 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td>33.62 43.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A1c</td>
<td>T18T19 (L)&lt;sup&gt;b&lt;/sup&gt; with T18T19 (H)&lt;sup&gt;c&lt;/sup&gt; Intra disulfide in T19 (H)</td>
<td>795.81 (5+) 683.58 (4+)</td>
<td>1.01 x 10&lt;sup&gt;5&lt;/sup&gt; 2.75 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td>40.69 44.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A1b</td>
<td>T18T19 (H)&lt;sup&gt;c&lt;/sup&gt; with T18T19 (H)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>791.97 (8+)</td>
<td>1.41 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td>41.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A1a</td>
<td>T11 (H) with T14 (H) T18T19 (L)&lt;sup&gt;b&lt;/sup&gt; with T18T19 (H)&lt;sup&gt;c&lt;/sup&gt; Intra disulfide in T19 (H)</td>
<td>842.25 (3+) 795.81 (5+) 683.58 (4+)</td>
<td>3.43 x 10&lt;sup&gt;5&lt;/sup&gt; 2.76 x 10&lt;sup&gt;4&lt;/sup&gt; 9.04 x 10&lt;sup&gt;4&lt;/sup&gt;</td>
<td>33.37 32.46 43.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A1</td>
<td>T18 (H) with T18T19 (L)&lt;sup&gt;b&lt;/sup&gt; T19 (H) with T19 (H)</td>
<td>631.87 (2+) 781.29 (7+)</td>
<td>4.43 x 10&lt;sup&gt;5&lt;/sup&gt; 8.58 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td>30.27 43.34</td>
</tr>
</tbody>
</table>

<sup>a, b, and c</sup>: the same footnotes as in Table 2-3.
Chapter 3
Analysis of N-Glycosylation in Therapeutic Monoclonal Antibodies by LC-MS with A New Chip MS Method

Publication:
3.1 Abstract

A robust and high throughput method is needed to measure the consistency of the N-glycan distributions in therapeutic monoclonal antibodies (mAbs). A new chip-based LC-MS approach has been developed for direct glycan analysis of intact mAbs. Immobilized PNGase F and graphitized carbon are integrated together in the chip to on-line release the glycans followed by mass spectrometric analysis. The pros and cons of the glycan analysis using the new chip LC-MS approach are compared with other chip based LC-MS methods such as C18 chip for the analysis of glycopeptides after enzymatic digestion and PGC chip to analyze glycans released from offline PNGase F treatment. The reproducibility of the new chip LC-MS method was evaluated and the applications of the new approach to analyzed different types of mAbs were also explored.
3.2 Introduction

N-Glycosylation is an important post-translational modification of IgG and almost all therapeutic mAbs are glycosylated at a conserved Fc region (Figure 3-1). For example, the glycosylation site for IgG1-type mAb is at Asn297 in the conserved sequence as shown in Figure 3-1B. The Fc glycans have a fucosylated core structure with terminal 0, 1, or 2 galactose residues (G0, G1, or G2, respectively). A very low level of terminal sialic acid or highly branched structures may be also present in some mAbs.¹

N-glycosylation significantly affects the quality, stability, and pharmacokinetics of therapeutic monoclonal antibody (mAb) drugs and the change of the glycan profiles can cause serious safety and efficacy issues.²⁻⁷ MAb glycans, however, can be highly heterogeneous in glycoforms and changed easily under different production conditions. Thus, comprehensive characterization of glycans in therapeutic monoclonal antibodies is critical to pharmaceutical and biotechnology companies.

Figure 3-1. Structure of IgG1 (A), structure of glycans and glycosylation site at IgG1 Fc region (B).
Currently, three major strategies are used for the analysis of N-glycosylation: capillary electrophoresis, HPLC and mass spectrometry. In CE, glycans are released from IgGs and labeled with fluorophore. The labeled glycans are then analyzed by CE combined with laser-induced fluorescence (LIF) detection. The glycan structure is identified by comparison of the retention time of released unknown glycan with the glycan standards. The released and fluorophore labeled glycans can also be analyzed by HPLC coupled with fluorescence detector or electric chemical detector using the same principle as CE. However, these methods can only match the released glycans with known standard. Thus, the glycans that do not exist in the glycan standards will not be identified.

Mass spectrometry is a powerful approach to analyze N-glycosylation. Several MS analysis platforms have been developed for different purposes. Glycopeptide analysis by MS allows characterization of glycans with site-specific information. Overall glycan profiling can be achieved by MS analysis of released glycans. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) is a widely used approach for profiling released glycans directly or in a derivatized form. However, glycan identification by MALDI-TOF MS is based on the detected mass without any separation. Therefore, glycans isomers with identical mass, such as G1 isomers, cannot be differentiated by MALDI-TOF. In addition, the neutral (e.g. G1 glycan) and acidic glycans (e.g. G2+SA glycan) have to be analyzed in different modes, which increase the complexity of analysis process. All these methods above do require sample pre-treatments such as enzymatic digestion for glycopeptide analysis, chemical or enzymatic release of glycans, and labeling released glycans. A new LC-MS approach, Agilent mAb-Glyco-Chip couple to QTOF, is developed for analysis of mAb glycans without any pre-treatment. In this approach, intact mAbs are directly
injected to the LC-MS system. Following the online enzyme digestion, the released glycans are separated by PGC column and analyzed by Q-TOF MS. No sample preparation is required and the total analysis of mAb can be as short as 12 min.\textsuperscript{10}

In this report, we have analyzed the glycan distributions of mAbs using this new chip MS approach. Glycan distributions of anti-her2 obtained from two different chip-based MS methods were compared with mAb-Glyco-chip MS. In addition, run-to-run and day-to-day reproducibility of this approach were tested. Finally, the applications of this approach to analyze mAbs from different manufacturers (IgG1, IgG4, and Fc-fusion type of protein) and to compare the glycan profiles from innovator and biosimilar mAbs were also reported.

3.3 Experimental Section

3.3.1 Materials

Therapeutic monoclonal antibodies were used in this study: (1) anti-HER2 (Herceptin, Genentech, South San Francisco, CA), a liquid formulation product (22 μg/μL); (2) anti-CD11a (Raptiva, Xoma, Berkeley, CA), a lyophilized powder; (3) GLP-1 with IgG4 Fc fusion protein (CNTO 736, Centocor R&D, Radnor, Pennsylvania), a liquid formulation product (10 μg/μL). Herceptin and Raptiva were the gift from Genentech and CNTO 736 was the gift from Centocor.(4) IgG1 type mAbs (mAb 1: 100 mg/mL) was obtained from Biogen Idec (Cambridge, Massachusetts); (5) anti-HER2 biosimilar mAb was a gift from Agilent, China, a lyophilized powder (10μg/μL). Achromobacter protease I (Lys-C) was obtained from Wako Co. (Richmond, VA). Trypsin (sequencing grade) was purchased from Promega (Madison, WI). Guanidine hydrochloride, ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide (IAA), and formic acid (FA) were from Sigma-Aldrich (St. Louis, MO). LC-MS grade water was from J.T. Baker.
(Phillipsburg, NJ), and HPLC grade acetonitrile from ThermoFisher Scientific (Fairlawn, NJ).

Precast NuPAGE 4–12% Bis-Tris sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels, NuPAGE MES SDS running buffer (20X), NuPAGE LDS sample buffer (4X), and SimplyBlueSafeStain were all from Invitrogen (Carlsbad, CA). Microcon YM-3 Centrifugal Filter Unit was obtained from Millipore (Bedford, MA). Deglycosylation buffer was from Agilent Technologies (Santa Clara, CA).

3.3.2 Enzymatic Digestion

PNGase F digestion: an aliquot of 20 μg anti-HER2 was mixed with 5 μL NuPAGE LDS Sample Buffer (4X) (pH 7.8) and 13 μL NuPAGE MES SDS Running Buffer (1X) to achieve the final volume of 20 μL. The mixture was incubated at 70 °C for 10 min. After incubation, the mixture (20 μL) was loaded onto an SDS-PAGE gel (4–12% Bis-Tris polyacrylamide) at 200 V for 40 min and then stained with Coomassie blue. The bands of heavy chain were excised and destained by alternating use of acetonitrile and 0.1 M ammonium bicarbonate (NH₄HCO₃) until no visible color was present. The destained gel pieces were completely dehydrated with acetonitrile and then dried in a SpeedVac (Labconco, Centrivap Cold Trap). PNGase F (1 Unit PNGase F for 100 ug protein) in 0.1 M NH₄HCO₃ was added to just cover the dried gel pieces. After 45–50 min incubation at 4 °C, the remaining supernatant was removed and incubated overnight (16 hr) at 37 °C. The digested peptides were then extracted with 5% formic acid and ACN (1:2, v/v) and then dried down until only 20 μL were left (if not, the samples were reconstituted to 10 μL for subsequent analysis).

Lys C digestion: an aliquot of 250 μg anti-her2 was denatured by 6 M guanidine hydrochloride in 0.1 M NH₄HCO₃ to get 2.5 μg/μL diluted sample. To reduce disulfide bond, 5
mM dithiothreitol (DTT) in NH₄HCO₃ was added to diluted sample and incubated at 37 °C for 30 minutes. Then the sample was subsequently alkylated at room temperature and in the dark for 60 min with 20 mM IAA in NH₄HCO₃. Then sample was buffer exchanged with 0.1M ammonium bicarbonate over an YM-3 Centrifugal Filter Unit, and then endoproteinase Lys-C (1:100 w/w) was added to digest the protein for 16 hr at 37 °C. Digestion was stopped by addition of 10% formic acid.

3.3.3 LC-MS

The Agilent 1200 HPLC-Chip system coupled with the Agilent 6520 Q-TOF through an Agilent 1100 Chip Cube (G4240A) was used for analysis of glycans distributions in mAbs. An Agilent 1100 autosampler was used for automatic sample injection. Mass spectra were acquired in positive ion mode. Three chips from Agilent Technologies (Santa Clara, CA) were used in this paper: a PGC chip (G4240-64010), a large capacity C18 chip (G4240-62010), and a mAb-Glyco chip (G4240-64020).

For PGC chip MS and C18 chip MS, mobile phase A with 0.1% formic acid in water and mobile phase B with 0.1% formic acid in acetonitrile were used on both CapPump and NanoPump. The electrospray voltage for mass spectrometer was set to 1880 V, with a drying gas of 6 L/min nitrogen at 325 °C. The fragmentor voltage was 175 V, and the skimmer voltage was 65 V. Data was acquired at 2 GHz, in extended dynamic range mode. The instrument was operated in a data-dependent mode: the first survey MS (scan 1) from m/z 400 to 2000 followed by one MS² scan within range m/z 150 to 2000. For PGC chip (G4240-64010) MS analysis of glycans, released glycans by off-line PNGase F digestion were enriched and separated by a 43 mm porous graphitized carbon chip with 40 nL trap column. For sample loading and enrichment,
5 μL of released and concentrated glycan was loaded by CapPump at 3μL/min with 2% of mobile phase B. For sample analysis, NanoPump at 0.3 μL/min was used with a gradient from 2% to 60% of mobile phase B between 0-20 minute. Mobile phase washed at 60% of mobile phase B for 2 minutes and followed by 2% B within 1 minute. Finally, the chip was balanced at 2% of mobile phase B for 15 minutes.

For C18 chip MS, 5 μL of Lys-C digested peptides were separated by a large capacity C18 chip (G4240-62010, 150 mm 300 A C18 chip w/ 160 nL trap column). The CapPump flushed at 3 μL/min of 2% of mobile phase B. The gradient (0.3 μL/min) used on NanoPump for elution peptides was: increase mobile phase B from 2% to 65% in 60 minutes followed by increasing mobile phase B from 65% to 80% in 10 minutes. Held mobile phase B at 80% for 10 minutes and then the mobile phase B was dropped back to 2% in 5 minutes and balanced the column at 2% of mobile phase B for 15 minutes.

For mAb-Glyco chip MS, mAb-Glyco chip (G4240-64020) is designed as an inner/outer rotor device and composed of three columns. One column is 310 nL enzyme reactor packed with immobilized PNGase F beads, which can be switched into or out of the capillary pump flow path for filling sample, deglycosylation, and transferring digested samples to enrichment column. Manufacture optimized 4-minute incubation of antibody in the enzyme reactor was used. Another two columns are a sample enrichment column (160 nL) and a separation column (43 mm). Both columns are packed with porous graphitized carbon. The NanoPump was used to generate gradient nanoflow at 500 nL/min with mobile phase A (0.1% formic acid in water) and mobile phase B (0.1% formic acid in acetonitrile). The gradient for deglycosylation and glycan analysis is as follows: 2% B in the first 6 min, 2-32% B in 1.5 min, 32-85% B in 30 s. After a hold at 85% for 1 min, gradient quickly drop back to 2% B in 0.01 s and followed by 3 min
balance in 2% B. A longer LC gradient was used to separate G1 isomers: 2% B in the first 6 min, 2-22% B in 10 min, 22-85% B in 30 s. Held at 85% B for 1 min. Then dropped back to 2% B in 0.01 s and followed by 3 min balance in 2% B. The CapPump was used for both loading and deglycosylation with deglycosylation buffer at 3 μL/min (from Agilent Technologies, Santa Clara, CA). The electrospray voltage for mass spectrometer was set to 1950 V, with a drying gas of 3.5 L/min nitrogen at 325 °C. The fragmentor voltage was 160 V, and the skimmer voltage was 65 V. Data was acquired at 2 GHz, in extended dynamic range mode. The instrument was operated in a mass scan mode from m/z 100 to 1700. Each sample was diluted to 0.5 μg/μL by using deglycosylation buffer and 4 μL diluted sample was analyzed each time.

3.3.4 Data analysis

First, extracted ion chromatograms (EIC) of major glycans and glycopeptides were performed using a 50 ppm mass range to obtain the peak area of the EIC. Then, glycosylation distribution was estimated by the peak area of each glycan or glycopeptide divided by the sum of peak areas of all glycans or glycopeptides, expressed as a percentage.

3.4 Results and Discussion

3.4.1 Analysis of glycans of anti-HER2 mAb

Anti-HER2 mAb is a IgG1-type mAb and used to treat breast cancer. The Sample was diluted with deglycosylation buffer to get a final concentration of 0.5 μg/μL. Then 4 μL of diluted sample was injected directly by autosampler. Following the online PNGase F digestion, the released glycans were separated by PGC column and identified by Q-TOF. When a large mass range, ± 0.5 Da (m/z), was used for extracted ion chromatogram (EIC), three peaks were
observed for each glycan (Figure 3-2A). The major peak is amine form glycan (eluted between 8.7 min to 9 min). The two small peaks are the cis and trans-hydroxyl species (eluted between 9.1 min to 9.4 min). PNGase F specifically removes N-linked glycan from asparagine residue in glycoprotein. Initially, glycans are released as β-glycosylamines, the amine form of glycan, with -NH2 linked to the reducing end. These amine form glycans however are not stable. They are slowly hydrolyzed to hydroxyl glycan forms under neutral conditions. This conversion is catalyzed under acidic conditions. Due to an anomeric equilibrium, each amine form glycan yields two hydroxyl forms10. In mAb-Glyco chip, the online deglycosylation time is only 4 minutes and glycan separation time is 3 minutes. The predominant glycans are still in amine forms. But exposure of cleaved amine form glycans to acidic mobile phase in the NanoPump flow during the separation process hydrolyzes a small amount of the amine glycan to hydroxyl glycans. All these species can be resolved on the PGC column. There is a 0.984 Da difference between the amine and hydroxyl forms. The cis and trans of hydroxyl forms have no mass difference but slightly different retention times when using a PGC column for separation10. When a smaller (± 50ppm) mass range were used for EIC, one peak (amine form) for each glycan was observed (Figure 3-2B). When a shallower LC gradient was used, the amine G1 isomers can be further separated as two peaks (Figure 3-2C). In addition to glycan isomers, mannose-type glycans and glycans with sialic acid can also be identified by this approach.10 Since the glycan was identified by accurate mass not MS2 fragmentation, for the structure specificity of glycans, targeted MS2 can be performed. As an example, Figure 3-3 showed the targeted MS2 analysis of Man9 (Figure 3-3A) and G1+SA (Figure 3-3B). The characteristic product ions confirmed the glycan structures.

Because amine-form glycan is the dominant species detected during online PNGase F
digestion and the EIC results from a narrow window (± 50 ppm) are more consistent, the peak area of EIC for amine form glycan with a narrow window (± 50 ppm) was used for glycan distribution calculations. The glycan distribution was estimated by the peak area of EIC for each glycan divided by the sum of peak areas of all glycans, expressed as a percentage.

Figure 3-2. Extracted ion chromatogram (EIC) of glycans from Ab1 analyzed by mAb-Glyco chip MS.
A. EIC performed with ± 0.5 m/z; B. EIC performed with ± 50 ppm; C. A longer LC gradient to separate G1 amine isomers.
### 3.4.2 Comparison of mAb-Glyco-chip MS with other LC-MS methods

Glycan analysis provides a glycan profile of mAb. Glycopeptide analysis by LC-MS is another common approach to characterize glycosylation, which allows site-specific analysis. Since anti-HER2 mAb has only one glycosylation site, the pros and cons of the glycan analysis using the new LC-MS approach are compared to Lys-C glycopeptides analysis by C18-chip-Q-TOF MS and offline-released glycans analysis by PGC-chip-Q-TOF MS. The glycopeptides or glycans distributions were calculated as mentioned before. Figure 3-4 showed that glycan distributions from these three different LC-MS methods are very consistent for both high and low abundant glycans.

In addition to site-specific analysis of glycosylation, glycopeptide analysis enables simultaneous analysis of peptide sequences and other post-translational modifications. Since peptide sequences are identified at the same time, the time for analysis of each sample is long; usually 1-1.5 hour. Besides, the off-line digestion process is time-consuming and laborious. By contrast, glycan analysis analyzes only glycans, so that the overall analysis time can be very short. In this report, a good separation was achieved by a 3 min gradient. In addition, glycan isoforms can be separated by PGC, which are unable to be discriminated on C18 column. However, glycosylation sites are not determined by glycan profiling.
Figure 3-3. Targeted MS/MS analysis of (A) Man9, m/z 941.85 (2+) and (B) G1+SA, m/z 958.38 (2+). □ N-acetylglucosamine □ Mannose ○ Galactose ▲ Fucose ● Sialic acid

Figure 3-4. Comparison of glycan distribution of anti-HER2 by three different chip-based Q-TOF MS methods.
3.4.3 Reproducibility of mAb-Glyco-chip MS analysis of mAb

Anti-HER2 mAb was analyzed on 6 different days within 5 months by 3 different chips with three runs each sample to evaluate the run-to-run and day-to-day reproducibility. The glycan profiles in Figure 3-5 showed an overall good run-to-run reproducibility. In Table 3-1, both high abundant glycans and low abundant glycans have low variation (CV<5%) of run-to-run analysis. Then the day-to-day reproducibility was calculated from these 6 results. The high abundant glycans (G0 and G1) have a low day-to-day variation (2.1% for G0 and 1.8% for G1). However, for G2 and G0-F, due to low abundance the day-to-day variation measurement could be as high as 20%, which should be tolerable considering that three different chips were used.
Figure 3-5. Reproducibility of run to run.

Table 3-1. Percentage of glycan distribution generated on 6 different days (average of three analysis per day).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>G0</td>
<td>45.4 (1.7)</td>
<td>45.7 (0.3)</td>
<td>43.9 (1.8)</td>
<td>44.2 (1.4)</td>
<td>43.1 (2.0)</td>
<td>44.2 (1.4)</td>
<td>44.4 (2.1)</td>
</tr>
<tr>
<td>G1</td>
<td>42.2 (1.7)</td>
<td>41.5 (0.7)</td>
<td>43.5 (1.2)</td>
<td>42.1 (0.9)</td>
<td>43.2 (1.7)</td>
<td>42.1 (0.9)</td>
<td>42.4 (1.8)</td>
</tr>
<tr>
<td>G2</td>
<td>6.5 (4.5)</td>
<td>7.6 (2.8)</td>
<td>8.7 (6.6)</td>
<td>6.8 (3.7)</td>
<td>7.7 (3.2)</td>
<td>6.8 (3.7)</td>
<td>7.3 (11.2)</td>
</tr>
<tr>
<td>G0 - Fu</td>
<td>5.9 (2.1)</td>
<td>5.1 (3.8)</td>
<td>4.0 (2.7)</td>
<td>7.0 (2.8)</td>
<td>6.0 (1.6)</td>
<td>7.0 (2.8)</td>
<td>5.8 (19.7)</td>
</tr>
</tbody>
</table>

- CV% in parentheses.
- three different chips were used for the time-course study
3.4.4 Analysis and comparison of innovator and biosimilar samples

One application of mAb-Glyco chip MS approach is to quickly compare innovator and biosimilar mAbs. As shown in Figure 3-6, compared with innovator anti-HER2 mAb (Figure 3-6A), the biosimilar mAb (Figure 3-6B) has higher G0 to G1 and G0 to G2 ratios but lower G0 to G0-F ratios.

![Figure 3-6. Comparison of glycan distribution in innovator (A) and biosimilar (B) anti-HER2 mAb by mAb-Glyco-chip-Q-TOF MS.](image)
3.4.5 Analysis of different types of monoclonal antibodies

Different types of mAbs and mAb in different formulations can also be analyzed by this mAb-Glyco-chip MS approach. For example, the glycan profiles of a lyophilized powder sample, anti-CD11a mAb (Figure 3-7B) and CNTO736, a GLP-1 with IgG4 Fc fusion protein (Figure 3-7C) were characterized. Compared to IgG 1 type mAb (Figure 3-7A), the same glycans eluting at similar retention times but different glycol profiles indicated that they were from different samples.

Figure 3-7. Analysis of different types mAbs in different buffers by mAb-Glyco-chip-Q-TOF MS.
A. anti-HER2 mAb: IgG1-type mAb in liquid formulation; B. anti-CD11a mAb: IgG1-type mAb in lyophilized powder; C. CNTO736 mAb: fusion IgG4 in liquid formulation.
3.5 Conclusions

The mAb-Glyco-Chip successfully released and separated glycans from intact monoclonal antibodies for MS detection with good run-to-run and day-to-day reproducibility. The other commonly used methods for glycosylation analysis, such as glycopeptide analysis by LC-MS and off-line released glycan analysis by PGC separation and MS, produced similar results as this new approach. Nevertheless, the automation of the on-line approach (mAb-Glyco-Chip) simplified the analysis procedure and reduced time (increased the throughput) significantly. Glycan isoforms, fucosylated and non-fucosylated glycans, and sialylated glycan can be easily separated by PGC column in the chip. In addition to IgG1 type mAbs, other types of mAbs and mAbs in different formulation can also be analyzed by this chip MS method. Furthermore, innovator and biosimilar mAbs can be quickly compared by this approach. Overall, this new mAb-Glycochip based assay offers a more facile and rapid method for glycan characterization, analysis at different development stages and biosimilar comparison.
3.6 References


Chapter 4
Structural Characterization of Innovator and Biosimilar Therapeutic Monoclonal Antibodies by LC-MS Based Approaches

Publication:
4.1 Abstract

Therapeutic monoclonal antibody (mAb) is a prospective approach to prevent or treat cancer and other inflammatory diseases. As the patents of some therapeutic monoclonal antibodies are going to expire soon, the production of biosimilar products becomes attractive. The complicated recombinant mAb production process, however, will introduce various modifications to monoclonal antibodies. Therefore, thorough characterization of therapeutic monoclonal antibodies drug, particularly, detections of functionality-sensitive modifications in mAbs are necessary and critical to ensure its correct structure that is strongly associated with the quality, safety and efficacy of the drug. In this report, a range of LC-MS based methods have been developed for comprehensive characterization and comparison of the structure of innovator and biosimilar therapeutic anti-HER2 mAbs. The complete amino acid sequences and common modifications were identified in both mAbs by LC-MS using the accurate peptide mass (FTICR MS) and sequence assignment (MS/MS measurement). The extents of N-terminal pyroglutamic acid cyclization, heavy chain C-terminal lysine processing, oxidation, deamidation and isomerization were measured by relative comparison of the peak areas of modified peptide and nonmodified peptides in the same LC-MS analysis. The identification of glycopeptides and disulfide linkages were further confirmed by using the combination of CID and ETD fragmentation. Through these LC-MS analyses, no significant difference was observed between innovator and biosimilar anti-HER2 mAbs.
4.2 Introduction

Therapeutic monoclonal antibody (mAb) has been one of the fastest-growing areas in biotechnological and pharmaceutical industry due to their high specificity, few side effects, and easy manipulation.1-5 By 2009, more than 20 monoclonal antibody drugs are available on the American market for treatment of many human diseases and a $36.4 billion sales record had been achieved on the global market.6-7

Trastuzumab (trade name: Herceptin) from Genentech was approved in 1998 for treatment of HER2 overexpressing metastatic breast cancer.8 This anti-HER2 mAb targets the HER2 receptors on cancer cell surface and inhibits cancer cell proliferation by interrupting its growth signal.9 In 2009, this drug had sales of $4.85 billion, among the top five best-selling mAbs.6 On the other hand, patients struggle for affording this drug. It has been reported that the average cost for Trastuzumab per year per patient is $37,000.10 However, the patent of Trastuzumab is going to expire in 2014 in EU and 2015 in US.6 With the loss of patent protection, a copied product, called “biosimilar” Europe or “follow-on” in US, could be developed with lower-price and is of great benefit to both pharmaceutical industry (stimulate further research) and customers (cost saving).11

Since mAb is a large molecule with complex 3D structure and function, its structure and modifications have significant influences on the quality, safety, and efficacy of drug product.3, 5, 7, 9 Therefore, comprehensive characterization of therapeutic innovator mAbs and its biosimilar is critical to biotech and pharmaceutical industry.

Typically, mAb is characterized by SDS-PAGE separation, tryptic peptide mapping, direct analysis of intact mAb by mass spectrometry, or limited digestion followed by RP-HPLC-MS analysis.5, 12-19 However, due to its superior resolution, sensitivity, and accuracy, mass
spectrometry (MS) is currently the most sophisticated and powerful tool for protein and mAb characterization.\textsuperscript{17, 20-22} Several mass spectrometry-based platforms have been developed for different levels of mAb characterization.\textsuperscript{5} Typically mAbs are cleaved into small peptides and then the peptides are separated by liquid chromatography coupled online with tandem mass spectrometry (LC-MS/MS). LC will separate sample mixture first to reduce the complexity and enable more in-depth and confident analysis in later MS and MS/MS.\textsuperscript{17, 23-24} In this manner, the structure information of product samples can be generated in details.

The purpose of this paper is to achieve a detailed characterization of anti-HER2 mAb and its biosimilar through the combination of several LC-MS analysis platforms.

4.3 Experimental Section

4.3.1 Materials

The innovator therapeutic monoclonal antibody, anti-HER2 (Genentech, Co. San Francisco, CA), which was generated against the overexpressed Her2 receptor, was used in this study. The freshly made liquid formulation product (22 μg/uL) is stored in the refrigerator (4 °C) for 3 years. The anti-HER2 biosimilar was a gift from Agilent Technologies in China. Achromobacter protease I (Lys-C) was from Wako Co. (Richmond, VA), and trypsin (sequencing grade) was from Promega (Madison, WI). Guanidine hydrochloride, ammonium bicarbonate, N-Ethylmaleimide, tris(hydroxymethyl)aminomethane (Tris) base, Dithiothreitol (DTT), Iodoacetamide (IAA), and formic acid (FA) were purchased from Sigma-Aldrich (St. Louis, MO). LC-MS grade water was obtained from JTBaker (Phillipsburg, NJ), HPLC grade acetonitrile was from ThermoFisher Scientific (Fairlawn, NJ). Microcon spin column (10 kDa MWCO) was purchased from Millipore (Bedford, MA). Pre-cast NuPAGE 4-12% Bis-Tris SDS-
PAGE gels, NuPAGE MES SDS Running Buffer, NuPAGE LDS Sample Buffer, and SimplyBlue SafeStain were obtained from Invitrogen (Carlsbad, CA).

### 4.3.2 SDS-PAGE

Around 7 ug reduced (with DTT, heat at 70 ºC for 10 min) and non-reduced (no DTT and no heat) innovator and biosimilar mAbs were separated by 4–12% Bis-Tris polyacrylamide gel and stained with Coomassie blue.

### 4.3.3 Enzymatic digestion

In-gel RCM (reduced and S-carboxymethylated) digestion: bands of interest were excised and destained by alternating use of acetonitrile and 0.1 M ammonium bicarbonate until no visible color present. The destained gel pieces were completely dehydrated with acetonitrile and then dried in a SpeedVac (Labconco, Centrivap Cold Trap). DTT (10mM in 0.1 M NH₄HCO₃) was added and incubated for 30 min at 60 ºC. After alternating 0.1 M NH₄HCO₃ and acetonitrile washing, the gel pieces were dried by SpeedVac. The reduced samples were alkylated by 55 mM Iodoacetamide in 0.1 M NH₄HCO₃. After 60 min incubation at room temperature in the dark, the gel pieces were washed alternatively with 0.1 M NH₄HCO₃ and acetonitrile and dried by SpeedVac. Trypsin (12.5 ng/µL in 50 mM NH₄HCO₃ and 5 mM CaCl₂) was added to just cover the dried gel pieces. After 60 min incubation at 4 ºC, remaining supernatant was removed and buffer without trypsin was added to cover gel pieces. Overnight (14.5 hours) incubation at 37 ºC was performed and the tryptic peptides were extracted with 5% formic acid and acetonitrile (1:2, v/v). The tryptic peptides were dried in a SpeedVac until 10-20 µL were left.
In-gel native digestion: similar to the in-gel RCM digestion except that neither DTT nor IAA was added. Trypsin in 0.1 M Tris, 5 mM CaCl$_2$, pH 6.8 was used for mAb digestion.

4.3.4 LC-MS

LC-MS experiments were performed on an LTQ FT MS or LTQ Orbitrap XL hybrid FTMS instrument (ThermoFisher Scientific, San Jose, CA) equipped with a nanospray ion source (New Objective, Woburn, MA) connected to an Ultimate 3000 nanoLC pump (Dionex, Mountain View, CA). The digest was injected onto a self-packed C18 column (Magic C18, 200 Å pore and 5-μm particle size, 75 μm i.d. × 15 cm) (Michrom Bioresources, Auburn, CA) manually or by autosampler. Mobile phase A was 0.1% formic acid in water and mobile B was 0.1% formic acid in acetonitrile. The peptides were eluted at 200 nL/min using 2% B to 65% B in 60 min, then from 65% B to 80% B for 10 min. Mass spectra were acquired from 380 to 1800 m/z in the FTICR cell or orbitrap cell. One survey full-scan MS was followed by ten sequential data-dependent MS/MS scans. An LTQXL with ETD mass spectrometer (Thermo Fisher Scientific, San Jose, CA), which was composed of a newly developed linear ion trap with an additional chemical ionization source to generate fluoranthene anions, was used for disulfide bond analysis. It was operated in the data-dependent mode with dynamic exclusion and the first survey MS (scan 1) from m/z 400 to 1900 in the linear ion trap (at a target value of 30,000 ions) was followed by three consecutive ion activation steps: CID-MS2 (scan 2), ETD-MS2 (scan 3), and CID-MS3 (scan 4). The CID-MS2 (target value of 30,000 ions and 28% normalized collision energy) and ETD-MS2 (target value of 30,000 ions and 35% normalized collision energy) activation scan steps were performed on the precursor ion (starting with the most intense ion).
from the first survey MS scan. Scans 2, 3, and 4 were repeated twice in sequence. Therefore, the first three highest precursor ions generated from the first MS scan were fragmented.

4.4 Results and Discussion

4.4.1 Structure of anti-HER2 mAb

This mAb against the overexpressed HER2 receptor (anti-HER2) is an IgG1-type mAb about 150 kDa. It is composed of two identical light chains (about 25 kDa) and heavy chains (50 kDa) stabilized by 4 inter- and 12 intra-chain disulfide bonds (Figure 4-1). There are 450 amino acids in single heavy chain and 214 amino acids in single light chain. Light chain and heavy chain are linked together by an inter-chain disulfide bond between the C-terminal end of the light chain (Cys214) and Cys223 in the heavy chain. Two parallel inter-chain disulfides are formed between two heavy chains by connecting Cys229 and Cys232 in one heavy chain and the same amino acid sites in the other heavy chain. Each light chain and heavy chain has two and four intra-chain disulfide bonds, respectively. N-linked bi-antennary glycans are attached to Asn300 on heavy chains.

4.4.2 SDS-PAGE results

Approximately 7 μg nonreduced (without reduction with DTT and no heat added) and reduced (with DTT treatment) innovator and biosimilar anti-HER2 mAbs were separated by SDS-PAGE and stained with Coomassie blue. As shown in Figure 4-2 left panel, when samples were reduced by DTT, disulfides were broken and caused heavy chains and light chains to dissociate from each other. Therefore, two major bands at around 50 kDa and 25 kDa corresponding to heavy and light chains were observed. Two bands at 120 kDa and 75 kDa were
also shown for both samples. These two bands could result from disulfide scrambling. Although DTT was used to reduce disulfide bond, disulfide scrambling still occurs easily at alkaline pH condition. 12, 14, 25-27

When intact samples were separated by SDS-PAGE, because no DTT was used to break disulfide bonds in mAb, a fairly strong band was present at the 150 kDa position corresponding to the intact mAb (Figure 4-2 right panel). In addition to this strong-intensity band, two bands with less abundance at the position of 260 kDa and 110 kDa were observed for both innovator and biosimilar samples and could be induced during mAb purification or caused by mAb aggregations or disulfide scrambling. 9 The higher intensities of these two bands (at 260 kDa and 110 kDa positions) were observed in biosimilar sample, indicating that biosimilar mAb is not as pure or stable as innovator sample.
Figure 4-1. Structure of anti-HER2 mAb.

Figure 4-2. Gel analysis of innovator and biosimilar anti-HER2 mAbs with reduced (left) and non-reduced (right) conditions.
4.4.3 Primary structure identification

The correct primary structure is the basis of efficacy and safety of mAb drug. To confirm the amino acid sequence, the antibody is reduced, alkylated, and digested with enzyme in-gel or in-solution. The digested peptides are then analyzed by mass spectrometer. The right size of cleaved peptides is critical for mass spectrometer analysis since it is very difficult to ionize large peptides (>4 kDa) or to retain too small peptides on C18 column. To get appropriately sized peptides, Lys-C and trypsin digestions were performed in our study. After digestion, peptides were separated and analyzed using LC-MS and LC-MS/MS. For example, as shown in Figure 4-3, the N-terminal peptide (T1, tryptic peptides of antibody starting from N-terminus) of innovator anti-HER2 mAb light chain produced by in-gel RCM trypsin digestion was identified at 26.56 min (Figure 4-3A) and assigned the accurate mass for the peptide, m/z 939.9579 with 2+ charge (Figure 4-3B). The precursor ion (m/z 939.9579) was automatically isolated and further fragmented in the linear ion trap (CID-MS2). The b and y ions in CID-MS2 (Figure 4-3C) provided the sequencing information to identify this peptide. All other peptides of innovator and biosimilar were identified in this approach. With the combination of the results from RCM trypsin or Lys-C in-gel digestion or intact trypsin in-gel digestion, all peptides were identified for both innovator and biosimilar anti-HER2 mAbs (Table 4-1), showing that no obvious amino acid difference was observed between innovator and biosimilar samples.
Figure 4-3. LC-MS analysis of N-terminal peptide of innovator anti-HER2 mAb light chain produced by in-gel RCM trypsin digestion.
A. Extracted ion chromatogram (XIC); B. Precursor ion scan at 26.56 min using FTMS. For illustration purpose, only m/z 938-944 region is shown; C. CID-MS$^2$ of the precursor ion in B. The peptide sequences with the observed fragment ions are shown in the insert.
Table 4-1. Summary of Primary Structure Coverage.

<table>
<thead>
<tr>
<th>Heavy chain</th>
<th>Start</th>
<th>End</th>
<th>Observed in different methods*</th>
<th>Heavy chain</th>
<th>Start</th>
<th>End</th>
<th>Observed in different methods*</th>
<th>Light chain</th>
<th>Start</th>
<th>End</th>
<th>Observed in different methods*</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>1</td>
<td>19</td>
<td>A/B/C</td>
<td>T21</td>
<td>259</td>
<td>277</td>
<td>A/B/C</td>
<td>T1</td>
<td>1</td>
<td>18</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T2</td>
<td>20</td>
<td>30</td>
<td>A/B/C</td>
<td>T22</td>
<td>278</td>
<td>291</td>
<td>A/B/C</td>
<td>T2</td>
<td>19</td>
<td>24</td>
<td>A/B</td>
</tr>
<tr>
<td>T3</td>
<td>31</td>
<td>38</td>
<td>A/B/C</td>
<td>T23</td>
<td>292</td>
<td>295</td>
<td>A/B/C</td>
<td>T3</td>
<td>25</td>
<td>42</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T4</td>
<td>39</td>
<td>43</td>
<td>A/B</td>
<td>T24*</td>
<td>296</td>
<td>304</td>
<td>A/B/C</td>
<td>T4</td>
<td>43</td>
<td>45</td>
<td>A/C</td>
</tr>
<tr>
<td>T5</td>
<td>44</td>
<td>50</td>
<td>A/B/C</td>
<td>T25</td>
<td>305</td>
<td>320</td>
<td>A/B/C</td>
<td>T5</td>
<td>46</td>
<td>61</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T6</td>
<td>51</td>
<td>59</td>
<td>A/B/C</td>
<td>T26</td>
<td>321</td>
<td>323</td>
<td>A/C</td>
<td>T6</td>
<td>62</td>
<td>66</td>
<td>B</td>
</tr>
<tr>
<td>T7</td>
<td>60</td>
<td>65</td>
<td>A/B/C</td>
<td>T27</td>
<td>324</td>
<td>325</td>
<td>C</td>
<td>T7</td>
<td>67</td>
<td>103</td>
<td>A/B</td>
</tr>
<tr>
<td>T8</td>
<td>66</td>
<td>67</td>
<td>A/B/C</td>
<td>T28</td>
<td>326</td>
<td>329</td>
<td>C</td>
<td>T8</td>
<td>104</td>
<td>107</td>
<td>A/B</td>
</tr>
<tr>
<td>T9</td>
<td>68</td>
<td>76</td>
<td>A/B/C</td>
<td>T29</td>
<td>330</td>
<td>337</td>
<td>A/B/C</td>
<td>T9</td>
<td>109</td>
<td>126</td>
<td>A/B</td>
</tr>
<tr>
<td>T10</td>
<td>77</td>
<td>87</td>
<td>A/B/C</td>
<td>T30</td>
<td>338</td>
<td>341</td>
<td>A/B/C</td>
<td>T10</td>
<td>127</td>
<td>142</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T11</td>
<td>88</td>
<td>98</td>
<td>A/B/C</td>
<td>T31</td>
<td>342</td>
<td>343</td>
<td>A/B</td>
<td>T11</td>
<td>143</td>
<td>145</td>
<td>B</td>
</tr>
<tr>
<td>T12</td>
<td>99</td>
<td>124</td>
<td>A/B/C</td>
<td>T32</td>
<td>344</td>
<td>347</td>
<td>A/B/C</td>
<td>T12</td>
<td>146</td>
<td>149</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T13</td>
<td>125</td>
<td>136</td>
<td>A/B/C</td>
<td>T33</td>
<td>348</td>
<td>358</td>
<td>A/B/C</td>
<td>T13</td>
<td>150</td>
<td>169</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T14</td>
<td>137</td>
<td>150</td>
<td>A/B/C</td>
<td>T34</td>
<td>359</td>
<td>363</td>
<td>A/B/C</td>
<td>T14</td>
<td>170</td>
<td>183</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T15</td>
<td>151</td>
<td>213</td>
<td>A/B</td>
<td>T35</td>
<td>364</td>
<td>373</td>
<td>A/B/C</td>
<td>T15</td>
<td>184</td>
<td>188</td>
<td>A/B</td>
</tr>
<tr>
<td>T16</td>
<td>214</td>
<td>216</td>
<td>A/B</td>
<td>T36</td>
<td>374</td>
<td>395</td>
<td>A/B/C</td>
<td>T16</td>
<td>189</td>
<td>190</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T17</td>
<td>218</td>
<td>221</td>
<td>A/B/C</td>
<td>T37</td>
<td>396</td>
<td>412</td>
<td>A/B/C</td>
<td>T17</td>
<td>191</td>
<td>207</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T18</td>
<td>222</td>
<td>225</td>
<td>B</td>
<td>T38</td>
<td>413</td>
<td>417</td>
<td>A/B/C</td>
<td>T18</td>
<td>208</td>
<td>211</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T19 or T19*</td>
<td>226</td>
<td>251</td>
<td>A/B/C</td>
<td>T39</td>
<td>418</td>
<td>419</td>
<td>A/B/C</td>
<td>T19</td>
<td>212</td>
<td>214</td>
<td>A/B/C</td>
</tr>
<tr>
<td>T20</td>
<td>252</td>
<td>258</td>
<td>A/B/C</td>
<td>T40</td>
<td>420</td>
<td>442</td>
<td>A/B</td>
<td>T41</td>
<td>443</td>
<td>450</td>
<td>A/B/C</td>
</tr>
</tbody>
</table>

Tn: tryptic peptides of antibody starting from N-terminal end (T1).
a: A: RCM, trypsin in-gel digestion; B: RCM, Lys-C in gel digestion; C: intact, trypsin in-gel digestion
T19*: T19 peptide (heavy chain) without C-terminal PK residues, which was further cleaved by Lys-C (K-PK).
T24*: glycosylated peptide.
4.4.4 Glycopeptides Identification

The correct glycosylation is directly related to the correct functions of antibodies.\textsuperscript{5,28-29} It is known that therapeutic antibody drugs are glycosylated at least at the conserved Fc position. In the anti-HER2 mAb, glycosylation site is Asn300 in the heavy chain. The antibody (in solution) was digested by Lys-C and glycopeptides (T23-24-25) with four types of glycans eluted at 33.34 min (Figure 4-4A) and identified by accurate mass assignment of the peptides shown in Figure 4-4B. The precursor ion, G1 with a 5+ charge, was further fragmented by CID and produced abundant product ions derived from glycosidic bond cleavages (Figure 4-4C), which is a characteristic pattern of glycopeptide fragmented by low-energy CID.\textsuperscript{30} The ion present at m/z 366.25 is Man-GlcNAc. The fragment ion at m/z 1136.51 indicates that there is a galactose attached to Man-GlcNAc. The ion at m/z 1177.18 is due to the loss of Man-GlcNAc. The loss of Man (m/z 1217.63) with the ion at m/z 1136.51 indicated that this is a G1 type glycan with only one galactose attached. Because little peptide backbone information was obtained in CID, an online LC-MS platform with ETD was further used for the peptide backbone identification. Different to CID, ETD cleavage mostly occurs on peptide backbone with glycan moiety intact and therefore peptide information could be obtained\textsuperscript{30}. A different fragmentation pattern was observed when ETD-MS\textsuperscript{2} was applied, in which all the cleavages happened on the peptide backbone (Figure 4-4D). In the Figure 4-4D, the mass shift of 1606.6 Da in the z28, not in the z13, z18, and z20, indicated that the the G1 type glycan was attached on the first asparagine (Asn300). Using the ETD spectrum, the amino acid sequence of the glycopeptide could be assigned. Therefore using ETD along with CID, peptide sequences, glycosylation sites, and glycan structures could be assigned simultaneously and confidently. Then the glycopeptide distribution was calculated and compared between innovator and biosimilar anti-HER2 mAbs.
As shown in Figure 4-5, the glycopeptide distributions are different between innovator and biosimilar anti-HER2 mAbs. Higher G0 and G0-Fu levels were found in biosimilar mAb. The different glycopeptide distribution could affect the drug efficacy.\textsuperscript{31-35} It has been reported that antibodies without core fucose have enhanced drug efficacy.\textsuperscript{36} However, the galactosylation level has no influence on one of the immune responses, antibody-dependent cell-mediated cytotoxicity (ADCC), which could be induced by anti-HER2 mAb to fight HER2+ breast cancer.

\textbf{4.4.5 Disulfide Linkages Identification}

To map disulfide bonds, the traditional approach is to obtain disulfide-dissociated peptide sequence with reduction combined with the determination of disulfide bond without reduction. Since there is no disulfide linkage information in database, the assignment of disulfide-linked peptides cannot be achieved using typical database searching. Manual assignment and comparison of the two spectra are required. Consequently, this approach is tedious, labor and time-consuming and the assignment may not be accurate.\textsuperscript{37} Wu et al. used online LC-MS with electron-transfer dissociation (ETD) approach to successfully confirm disulfide linkages of several proteins with high confidence and in just one step. ETD, similar to electron-capture dissociation (ECD), has been shown to preferentially break disulfide bonds.\textsuperscript{37} In their strategy, CID-MS\textsuperscript{2} and ETD-MS\textsuperscript{2} were used to identify disulfide bonds. Then, disulfide-dissociated peptide, generated directly or indirectly from ETD-MS\textsuperscript{2}, was characterized by following MS\textsuperscript{3} step.\textsuperscript{37} The correct disulfide linkages in innovator anti-HER2 mAb have been fully characterized by LC-MS with ETD.\textsuperscript{9} The same approach was also used on anti-HER2 biosimilar and confirmed the disulfide linkages of anti-HER2 biosimilar were correct (data not shown).
major disulfide scramblings shown under heat-stressed condition\(^9\) were not observed in either innovator or biosimilar anti-HER2 mAbs without heat.

Figure 4-4. LC-MS analysis of glycopeptide (Lys-C digestion) (G1, m/z=1014.71, 5+) of innovator anti-HER2 mAb.

A. Extracted ion chromatogram (XIC); B. Precursor ion scan at 33.34 min using FTMS. For illustration purpose, only m/z 952-1047 region is shown; C. CID-MS\(^2\) of the precursor ion m/z 1014.9089 (5+) ion; D. ETD-MS\(^2\) of the precursor ion m/z 1014.9089 (5+) ion. The peptide sequence with the observed fragment ions are shown in the insert. ▲ N-acetylglicosamine  □ Mannose  ○ Galactose  △ Fucose
4.4.6 Chemical Modifications

MAbs are produced from a single cell clone. Theoretically, the primary sequence of mAbs should be the same. However, mAbs may undergo a series of modifications during the manufacturing, formulation, and storage and become heterogeneous. The most common chemical modifications in mAbs include: N-terminal pyroglutamic acid cyclization, heavy chain C-terminal lysine processing, oxidation, deamidation, and isomerization.\textsuperscript{5,38} In this section, all these common chemical modifications in innovator and biosimilar anti-HER2 mAbs were analyzed and compared.

4.4.6.1 N-terminal pyroglutamic acid cyclization

The N-terminal residue of most therapeutic mAbs is glutamine or glutamic acid and they are easily cyclized to pyroglutamic acid. This nonenzymatic modification is spontaneous and may not have significant impact on Ab structure and function.\textsuperscript{5,38-40} Cyclization of N-terminal...
glutamine and glutamic acid cause a release of ammonia and water, respectively, which are 17 and 18 Da mass differences between non-cyclized and cyclized species. These mass differences can be readily distinguished by mass spectrometry. As shown in the insert of Figure 4-6A and B, the mass difference (9.0067) of the monoisotopic ion (2+) between the non-cyclized and cyclized peptide matched the mass difference of a water molecule for a 2+ charge ion. Furthermore, similar y ions series but different b ions pattern (b11, b12, and b18, with 18 Da difference) were observed between non-modified and modified peptides in CID fragmentation, which confirmed that the modification is on N-terminal glutamic acid with a loss of water.

Similar to glycosylation distribution estimation, the relative quantitation for the extent of modification is performed by the equation:

\[
\% \text{ of N-terminal cyclization} = \frac{\text{Peak area of } \left( \text{cyclized peptide}/(\text{cyclized peptide} + \text{non-cyclized peptide}) \right)}{100\%}.
\]  

As shown in Table 4-2, the innovator and biosimilar anti-HER2 mAbs have a similar N-terminal cyclization content.
Figure 4-6. LC-MS analysis of the heavy chain N-terminal peptide and the N-terminal cyclized peptide from tryptic digest of biosimilar anti-HER2.

A. CID-MS\(^2\) of the heavy chain N-terminal peptide, m/z 941.5068 (2+) ion, with the FTICR MS measurement of the precursor ion in the insert; B. CID-MS\(^2\) of the cyclized heavy chain N-terminal peptide, m/z 932.5001 (2+) ion, with the FTICR MS measurement of the precursor ion in the insert.

Table 4-2. Comparison of the percentage of terminal modification for innovator and biosimilar anti-HER2 mAbs. The values are average of three runs ± CV%.

<table>
<thead>
<tr>
<th>Modification type</th>
<th>Innovator</th>
<th>Biosimilar</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-terminal cyclization</td>
<td>1.32 ± 0.47</td>
<td>1.46 ± 0.22</td>
</tr>
<tr>
<td>C-terminal lysine processing</td>
<td>99.55 ± 0.09</td>
<td>98.76 ± 0.06</td>
</tr>
</tbody>
</table>
### 4.4.6.2 Heavy chain C-terminal lysine processing

Removal of heavy chain C-terminal lysine by carboxypeptidases during cell culture production is one of the most common modifications of mAb. Similar to N-terminal pyroglutamic acid cyclization, C-terminal lysine processing has no significant influence on Ab structure and functions such as antigen binding and receptor binding to Fc.\(^5\,\text{[38,41-42]}\) The loss of one C-terminal lysine residue reduces the molecular weight by 128 Da, which can be identified by mass spectrometry as shown in the insert of Figure 4-7 A and B. The characteristic CID fragmentation patterns in Figure 4-7 confirmed the sequence of the C-terminal tryptic peptides and the loss of one lysine residue. The percentage of heavy chain C-terminal lysine processing was calculated in the same method as N-terminal cyclization and listed in Table 4-2. We observed that 99.55% of innovator and 98.76% of biosimilar anti-HER2 mAbs have lost their heavy chain C-terminal lysine. The percentage of removal of terminal lysine in anti-HER2 mAb is consistent with other report.\(^43\)
Figure 4-7. LC-MS analysis of the heavy chain C-terminal peptide and the C-terminal truncated peptide from tryptic digest of biosimilar anti-HER2.

A. CID-MS² of the heavy chain C-terminal peptide, m/z 788.4521 (1+) ion, with the FTICR MS measurement of the precursor ion in the insert; B. CID-MS² of the heavy chain truncated C-terminal peptide, m/z 660.3535 (1+) ion, with the FTICR MS measurement of the precursor ion in the insert.
4.4.6.3 Oxidation

Oxidation of methionine to methionine sulfoxide is often observed during mAb drug purification and storage. CH2 domain stability could be decreased if this modification occurs in CH2 domain. By addition of one oxygen atom, oxidized product, methionine sulfoxide has a 16 Da higher mass and is more hydrophilic. The observed mass difference between oxidized and nonoxidized monoisotopic ion in the insert of Figure 4-8 matches the mass difference of oxygen atom for a 2+ charge ion. The product ions in CID-MS\textsuperscript{2} of oxidized and non-oxidized peptides confirmed the peptide sequence and identified the oxidation was localized at the methionine (Met) residue. By using the same method, we also identified other oxidized peptides of both innovator and biosimilar anti-HER2 mAbs (data not shown). The oxidation content was estimated by the peak area of the oxidized peptide divided by the sum of peak area of the oxidized and nonoxidized peptides, the same approach as for the modifications described above (Equation 1). The summary of the percentage of oxidation was listed in Table 4-3. Among the five identified oxidized sites, the innovator and biosimilar anti-HER2 mAbs have the similar extent of oxidation. M255 is most susceptible to oxidation in both innovator and biosimilar anti-HER2 samples, which is consistent with another report. It should be noted that when the electrospray ion source is operated under atmospheric pressure, LC-MS analysis might produce some amount of oxidation artificially. However, the relative comparison between the innovator and biosimilar mAbs is valid due to the same or very similar amount of artificial oxidation.
Figure 4-8. LC-MS analysis of the T10 and the oxidized T10 peptide from tryptic digest of biosimilar anti-HER2 heavy chain.

A. CID-MS² of T10(H) peptide, m/z 655.8279 (2+) ion, with the FTICR MS measurement of the precursor ion in the insert; B. CID-MS² of oxidized T10(H) peptide, m/z 663.8245 (2+) ion, with the FTICR MS measurement of the precursor ion in the insert.

Table 4-3. Comparison of the percentage of oxidation for innovator and biosimilar anti-HER2 mAbs. The values are average of three runs ± CV%.

<table>
<thead>
<tr>
<th>Tryptic Peptide Sequence</th>
<th>Innovator</th>
<th>Biosimilar</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTAYLQM₈₈ NSLR (HC)</td>
<td>6.38 ± 9.18</td>
<td>4.84 ± 7.77</td>
</tr>
<tr>
<td>WGGDGFYAM₁₀₇ DYWGQGT LVTVSSASTK (HC)</td>
<td>3.52 ± 13.46</td>
<td>4.47 ± 7.31</td>
</tr>
<tr>
<td>DTLM₂₅₅ ISR (HC)</td>
<td>11.05 ± 6.1</td>
<td>10.58 ± 9.84</td>
</tr>
<tr>
<td>WQQGNVFSCSVM₄₃₁ HEALHNHYTQK (HC)</td>
<td>3.11 ± 10.03</td>
<td>3.68 ± 6.27</td>
</tr>
<tr>
<td>DIQM₄ TQSPSSLSASVGDR (LC)</td>
<td>6.39 ± 4.32</td>
<td>5.45 ± 9.2</td>
</tr>
</tbody>
</table>

a: HC means heavy chain; LC means light chain.
4.4.6.4 Deamidation

Nonenzymatic deamidation of asparagines is a common modification that can occur especially after long-term storage. The most susceptible site for deamidation is the asparagine followed by glycine or serine.\textsuperscript{38} The process goes through a loss of amine in the side chain of asparagine to form a succinimide intermediate. This intermediate is usually hydrolyzed quickly to aspartic acid and isoaspartate in a molar ratio of 1:3.\textsuperscript{5,38} Compared to asparagine residues, the succinimide intermediate has a 17 Da reduced mass but the deamidated products, aspartic acid and isoaspartic acid, have a 1 Da increased mass charge. By using the same approach as for previously discussed modifications, deamidated products were not detected. However, we found the presence of succinimide intermediate. As shown in Figure 4-9, the 8.5 Da mass difference for 2+ charged ions in inserted figures matches the mass change between peptides with asparagine residue and succinimide intermediate. The characteristic CID-MS\textsuperscript{2} fragmentation identified the sequence of peptides. In both fragmentation, the same product ions such as y6, y8, b9, b11, b16 and the 17 Da mass difference for 1+ and 8.5 Da mass difference for 2+ product ions such as y10-13, y 15 and y19 confirmed the 17 Da mass difference between these two peptides and the modification occurred on Asn387 not on Asn392 and Asn393. The percentage of succinimide intermediate was calculated in the same method as percentage of oxidation (Equation 1) by using the peak area of XIC of nonmodified and modified peptides. The innovator and biosimilar anti-HER2 mAbs showed the similar percentage of succinimide intermediate in four identified peptides (Table 4-4). The succinimide intermediate has been found in several mAbs in spite of its instability.\textsuperscript{38}
Figure 4-9. LC-MS analysis of the biosimilar anti-HER2 tryptic T36 (H) and the T36 (H) peptide with succinimide intermediate formed during asparagine deamidation.

A. CID-MS$^2$ of T36(H) peptide, m/z 1273.0720 (2+) ion, with the FTICR MS measurement of the precursor ion in the insert; B. CID-MS$^2$ of T36 peptide with succinimide intermediate, m/z 1264.5610 (2+) ion, with the FTICR MS measurement of the precursor ion in the insert.

Table 4-4. Comparison of the percentage of succinimide intermediate formed during asparagine deamidation for innovator and biosimilar anti-HER2 mAbs. The values are average of three runs ± CV%.

<table>
<thead>
<tr>
<th>Tryptic Peptide Sequence</th>
<th>Innovator</th>
<th>Biosimilar</th>
</tr>
</thead>
<tbody>
<tr>
<td>IYPTN$_{55}$GYTR (HC)</td>
<td>2.93 ± 4.02</td>
<td>2.85 ± 1.03</td>
</tr>
<tr>
<td>NTAYLQMN$_{84}$SLR (HC)</td>
<td>0.19 ± 3.92</td>
<td>0.19 ± 1.46</td>
</tr>
<tr>
<td>VVSVLTVLHQDWLN$_{318}$GK (HC)</td>
<td>1.51 ± 3.92</td>
<td>1.71 ± 2.69</td>
</tr>
<tr>
<td>GFYPSDIAVEWESN$_{387}$GQPENNYK (HC)</td>
<td>2.22 ± 2.92</td>
<td>2.10 ± 5.90</td>
</tr>
<tr>
<td>ASQDVN$_{30}$TAVAWYQQKPGK (LC)</td>
<td>0.53 ± 16.33</td>
<td>0.35 ± 6.96</td>
</tr>
</tbody>
</table>

a: HC means heavy chain; LC means light chain.
**4.4.6.5 Aspartic acid Isomerization**

Aspartic acid followed by glycine can easily release a molecule of water to form a succinimide, the same intermediate as in the asparagine deamidation. The succinimide intermediate is unstable and is quickly hydrolyzed to isoaspartic acid, which introduces a methyl group to the peptide backbone.\(^{38}\) So asparagine deamidation and aspartic acid isomerization both induce a structural change and may be immunogenic or decrease drug potency.\(^{43,46}\) There is no mass difference between aspartic acid and isoaspartic acid but different chromatographic behavior is observed between them.\(^{38}\) Here, instead of isoaspartic acid, the succinimide intermediate was observed in two peptides. Similar to asparagine deamidation, the insert in Figure 4-10 matches the mass difference (18 Da for 1+ charge ion) and the CID-MS2 fragmentation confirmed the peptide sequence and located the modification site. Similar to the calculation of percentage of asparagine deamidation (Equation 1), the percentage of succinimide intermediate was calculated and listed in Table 4-5. Biosimilar anti-HER2 mAb has a higher percentage of Asp102 succinimide but a similar content of Asp414 succinimide compared to innovator. It has been reported that drug potency is significantly reduced when the Asp 102 succinimide is formed in one heavy chain of anti-HER2 mAb.\(^{43}\) The higher percentage of Asp102 succinimide in biosimilar sample could lead to less drug efficacy.
Figure 4-10. LC-MS analysis of the biosimilar anti-HER2 tryptic T12 (H) and the T12 (H) peptide with succinimide intermediate formed during aspartic acid isomerization. A. CID-MS$^2$ of T12(H) peptide, m/z 929.0947 (3+) ion, with the FTICR MS measurement of the precursor ion in the insert; B. CID-MS$^2$ of T37 peptide with succinimide intermediate, m/z 923.0855 (3+) ion, with the FTICR MS measurement of the precursor ion in the insert.

Table 4-5. Comparison of the percentage of succinimide intermediate formed during aspartic acid isomerization for innovator and biosimilar anti-HER2 mAbs. The values are average of three runs ± CV%.

<table>
<thead>
<tr>
<th>Tryptic Peptide Sequence</th>
<th>Innovator</th>
<th>Biosimilar</th>
</tr>
</thead>
<tbody>
<tr>
<td>WGGD$_{102}$GFYAMDYGQGTVLTVSSASTK(HC)</td>
<td>0.61 ± 10.11</td>
<td>1.09 ± 7.5</td>
</tr>
<tr>
<td>TTPPVLDSD$_{414}$GSFFLYSK (HC)</td>
<td>0.92 ± 6.38</td>
<td>1.05 ± 7.52</td>
</tr>
</tbody>
</table>

a: HC means heavy chain
4.5 Conclusions

All peptide sequences of innovator and biosimilar anti-HER2 mAb including glycosylation and correct disulfide bonds were studied and identified by the LC-MS based analysis. The use of ETD combined with CID fragmentation greatly enhanced the identification of glycopeptide and disulfide-linked peptides.

The common modifications such as N-terminal pyroglutamic acid cyclization, heavy chain C-terminal lysine processing, oxidation, deamidation, and isomerization were identified in both innovator and biosimilar anti-HER2 mAbs and similar levels of modifications were observed by using relative quantitation approach.

Through our LC-MS based approaches, the structures of innovator and biosimilar anti-HER2 mAbs were comprehensively characterized and compared. No significant structural difference was observed between innovator and biosimilar anti-HER2 mAbs.
4.6 References


Chapter 5
N-Glycosylation Analysis of Immunoglobulin Gs from HIV-1 Patients by Different Analysis Platforms with LC-MS

The purification of serum samples was performed by Dr. Galit Alter’s group

Publication
5.1 Abstract

Preliminary data suggest that the effector function of IgGs may play an important role in the control of viral infections. The N-glycans on IgGs can strongly influence the affinity of IgGs binding to different Fc receptors and therefore modify IgGs effector function. However, little is known about how glycosylation pattern affects IgG Fc-binding properties and HIV infection. Four LC-MS analysis platforms are developed to characterize N-glycosylation on IgGs from patients with HIV-1 infection. After a simple albumin depletion and SDS-PAGE separation step, we either prepared the glycopeptides by trypsin or Lys-C digestion or the glycan fraction by PNGase F reaction on the IgGs. In this study, we compared two different LC-MS platforms for analysis of glycopeptides (C18 reversed phase high-performance liquid chromatography fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) and C18 reversed phase high-performance microfluidic chip liquid chromatography quadruple time-of-flight mass spectrometry (Q-TOF MS) and two platforms for the corresponding glycans (microfluidic-based porous graphitized carbon chip and integrated mAb-Glyco chip with quadrupole time-of-flight mass spectrometry). There are four IgG subclasses and more than one glycosylation sites on each IgG subclasses. Glycopeptide analysis allowed quantitation of a given glycan motif on individual IgG subclasses but glycan analysis enables profiling the glycans from all IgG subclasses and from all sites. In a comparative study, similar results were generated between two glycopeptide analysis platforms and two glycan analysis platforms. However, mAb-chip-Q-TOF MS offers a more facile and high-throughput approach for glycan quantitation and thus is more suitable for large clinical studies.
5.2 Introduction

There are four IgG subclasses in human blood and they are named in order of their abundance: IgG1, IgG2, IgG3, and IgG4. IgG antibody (Figure 5-1) is a protein around 150 kDa and consists of two identical heavy chains (around 50 kDa each chain) and two identical light chains (around 25 kDa each chain). The heavy chains and light chains are connected together by inter- and intra-chain disulfide bonds to form a Y-like shape.

N-Glycosylation is one of the most common posttranslational modifications of IgGs. All IgG molecules contain a conserved N-glycosylation site (e.g. Asparagine 297 at IgG1) in their Fc regions. In N-linked glycans, the glycan is linked to asparagine within a consensus tripeptide sequence: asparagine–X–serine/threonine, where X is any amino acid except proline. In the case of IgGs, the sequences are all N-S-T, which is a favorable sequence for glycosylation. All N-linked glycans have a common core structure containing two N-acetylglucosamine residues (GlcNAc) and three mannose residues in a biantennary structure (black square in Figure 5-1).

During in vivo biosynthesis variable terminal-sugar residues are added to the core structure and make the N-glycan structure heterogenous. The most common N-linked glycans on IgG molecule are fucosylated core structure with terminal sugars containing no (G0), one (G1), or two (G2) galatose residues.

In addition to their neutralizing properties, antibodies can mediate other effector function, such as antibody-dependent cell-mediated cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC). In ADCC, the Fc region of IgG binds to Fc receptors (FcγRs) on the surface of an effector cell, such as natural killer (NK), causing lysis of the pathogen-infected target cells that are bound by IgGs. Several reports have shown that the binding properties of antibodies to Fc receptors are modulated by the Fc glycan structure. For example, defucosylated IgG1
exhibited higher affinity to ADCC mediating Fc receptors.\textsuperscript{5} Lower binding to FcγRIIIa receptor and therefore decreased ADCC activity has been observed in IgGs with terminal sialic acids (N-acetylglucosamine). In addition, lower galactosylation has been reported in several diseases and results in decreased binding of antibody to C1q and reduced CDC activity.\textsuperscript{6-9} Preliminary data suggest that ADCC-mediating antibodies have an important role in slowing HIV progress.\textsuperscript{10-11} However, little is known about how glycosylation patterns modify IgG Fc-binding properties and finally control HIV infection. The detailed characterization of these glycans on IgGs may facilitate to understand the mechanism of IgG effector function via Fc receptors and develop better HIV vaccine or therapeutic antibodies.

Mass spectrometry is an efficient and powerful approach to detect and identify N-glycosylation due to its high sensitivity and ability to \textit{rapidly} identify \textit{complex} mixtures by MS/MS sequencing. It is not possible to do detailed glycan analysis on an intact glycoprotein largely due to extensive heterogeneity of the sample and thus the analysis will either be performed on an enzymatic digestion or cleavage of the glycans. Glycopeptides can be analyzed by LC-MS while the released glycans can be analyzed by MALDI-TOF or LC-MS directly or as a derivatized form. MALDI-TOF has been widely used for detection glycans and provides a relatively fast measurement. However, MALDI-TOF measures the mass only, and is difficult to quantify the neutral (e.g. G1 glycan) and acidic glycans (e.g. G2+SA glycan) in the same mode. In addition, glycans isomers, such as G1 isomers, cannot be separated by MALDI-TOF.\textsuperscript{12-14}

In this report, integrated analysis of glycosylation patterns on IgGs from four different patients with HIV-1 infection (two still healthy and two with progressive weakness) was achieved by LC-MS based methods.
As is shown in Figure 5-1, the analysis of glycans is simplified in a typical IgG by the occurrence of a single glycosylation site at Asn 297. There can be, however, significant glycan heterogeneity (the major forms as well as the core glycan structure are shown in the Figure 5-1). One of our goals in this study is to develop an analytical method which is specific for a given glycan in the four IgG subclasses. As shown in Tables 5-1 and 5-2, a variety of glycan structures attached to Asn 297 can be uniquely identified for a given subclass i.e. IgG1, IgG2, IgG3 or IgG4 due to specific amino acid residues which are class specific (residues 296 and 300). While IgG3 and IgG4 have different amino sequences, the corresponding trypsin or Lys-C peptides have the same monoisotopic masses and therefore these IgG subclasses cannot be differentiated through this analysis platform. Glycopeptides were analyzed on two different platforms by RPLC-MS (C18 RP nanoHPLC/nanoESI FTICR MS and nano-LC-chip-MS). As an alternative approach, we also explored the glycan analysis by LC-MS methods. Several HPLC separation methods can be used for glycan profiling. Among them, porous graphitized carbon liquid chromatography has high separation power for glycans and is able to separate glycan isomers, neutral, and sialylated glycans in a single run with no derivatization required. In this report, two chip-based LC-MS methods were evaluated. One is the offline PNGase F released glycans were studied by a porous graphitized carbon (PGC) chip-MS. Another method is Agilent mAb-Glyco-Chip-MS for rapid detection and measurement of glycans. In this new approach, samples are directly injected and followed by online PNGase F digestion. Afterwards, the released glycans are separated by PGC column and analyzed by QTOF MS. This new chip based LC-MS approach has been initially studied. Results generated from these approaches were compared. The advantages and disadvantages of each approach were discussed.
5.3 Experimental Section

5.3.1 Materials

IgGs of four different patients with HIV-1 infection (two still healthy and two progressive weakness) samples were obtained from Ragon Institute, Harvard University. The serum samples were treated with a Melon Gel IgG Purification Kit (Thermo Scientific, Waltham, MA). The non-IgG proteins were bound to the gel while the IgGs as eluent were recovered in the flow-through (non-bound) fraction. Achromobacter protease I (Lys-C) was obtained from Wako Co. (Richmond, VA). Trypsin (sequencing grade) was purchased from Promega (Madison, WI). PNGase F, guanidine hydrochloride, ammonium bicarbonate, dithiothreitol (DTT), iodoacetamide (IAA), and formic acid (FA) were from Sigma-Aldrich (St. Louis, MO). LC-MS grade water was from J.T. Baker (Phillipsburg, NJ), and HPLC grade acetonitrile from ThermoFisher Scientific (Fairlawn, NJ). Precast NuPAGE 4–12% Bis-Tris sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels, NuPAGE MES SDS running buffer (20X), NuPAGE LDS sample buffer (4X), and SimplyBlue SafeStain were all from Invitrogen (Carlsbad, CA). Microcon YM-3 Centrifugal Filter Unit was obtained from Millipore (Bedford, MA). Deglycosylation buffer was from Agilent Technologies (Santa Clara, CA).

5.3.2 SDS-PAGE

An aliquot of 30 μg eluent from a Melon Gel IgG Purification Kit was mixed with 5 μL NuPAGE LDS Sample Buffer (4X), 2 μL 1M DTT. Appropriate volume of NuPAGE MES SDS Running Buffer (1X) was added to the mixture to achieve the final volume of 20 μL. After incubation at 70 °C for 10 min, the mixture (20 μL) was separated by SDS-PAGE gel (4–12% Bis-Tris polyacrylamide) at 200 V for 40 min and then stained with Coomassie blue.
5.3.3 Enzymatic Digestion

In-gel trypsin and Lys-C digestion: gel bands of expected IgG heavy chain were excised and cut into small pieces. The gel pieces were then washed by alternating use of acetonitrile and 0.1 M ammonium bicarbonate (NH₄HCO₃) until no visible color was present. DTT (10 mM in 0.1 M NH₄HCO₃) was added into SpeedVac dried gel pieces for 30 min at 60 ºC. The reduced samples were alkylated by 55 mM Iodoacetamide in 0.1 M NH₄HCO₃. After 60 min incubation at room temperature in the dark, the gel pieces were shrunk by acetonitrile and dried by Speed-Vac. Trypsin or Lys-C (12.5 ng/µl in 50 mM NH₄HCO₃) or PNGase F (1 unit in 0.1 M NH₄HCO₃) was added to just cover the dried gel pieces. After 45–50 min incubation at 4 ºC, overnight (14.5 hours) incubation at 37 ºC was performed. Trypsin or Lys-C digested peptides were extracted with 5% formic acid/acetonitrile (1:2, v/v) while PNGase F digest was extracted with water. The digested peptides were completely dried in a SpeedVac. Then 20 µL of 0.1% formic acid in water was added to redissolve digested peptides.

5.3.4 LC-MS Analysis

Platform 1: C18 RP nanoHPLC-FTICR MS. An aliquot of 5 µL of tryptic digest was injected onto a self-packed C18 column (Magic C18, 300 Å pore and 5 µm particle size, 75 µm i.d.X15 cm) (Michrom Bioresources, Auburn, CA). Mobile phase A was 0.1% formic acid in water, and mobile B was 0.1% formic acid in acetonitrile. The peptides were eluted at 200 nL/min using 2% B to 45% B in 60 min, then from 45% B to 85% B in 5 min. The MS experiments were performed on an LTQ FT mass spectrometer (Thermo Fisher Scientific, San Jose, CA). The instrument was operated in a data-dependent mode: the first survey MS (scan 1) from m/z 400 to 2000 followed by eleven sequential MS² scans.
Platform 2: C18 RP HPLC-chip-Q-TOF MS. The Agilent 1200 HPLC-Chip system was coupled with the Agilent 6520 Q-TOF through an Agilent 1100 Chip Cube (G4240A). An Agilent 1100 autosampler was used for automatic sample injection. Mass spectra were acquired in positive ion mode. Lys-C digested peptides (5 μL) were separated by a large capacity C18 chip (G4240-62010, 150mm 300 Å C18 chip with 160 nL trap column). Mobile phase A with 0.1% formic acid in water and mobile phase B with 0.1% formic acid in acetonitrile were used on both capillary pump (CapPump) and NanoPump. The electrospray voltage for mass spectrometer was set to 1880 V, with a drying gas of 6 L/min nitrogen at 325 °C. The fragmentor voltage was 175 V, and the skimmer voltage was 65 V. The instrument was operated in a data-dependent mode: the first survey MS (scan 1) from m/z 400 to 2000 followed by one MS² scan within range m/z 100 to 2000. The CapPump flushed at 3 μL/min of 2% mobile phase B. The gradient (0.5 μL/min) used on NanoPump for elution peptides is shorter than the gradient used on LTQ-FT for a high-throughput analysis: from 2% B to 50% B in 30 min; 50% B to 90% B in 2 min.

Platform 3: PGC-chip-Q-TOF MS. PNGase F off-line released glycans (5 μL) was separated by a PGC chip (G4240-62003, 43 mm porous graphitized carbon chip with 40 nL trap column). The same instruments and mobile phases as C18 HPLC chip-Q-TOF MS analysis were used. Glycans were eluted using gradient: 2-32% B in 1.5 min, 32-85% B in 30 s, then holds 85% for 1 min. The electrospray voltage for mass spectrometer was set to 1950 V, with a drying gas of 3.5 L/min nitrogen at 325 °C. The fragmentor voltage was 160 V, and the skimmer voltage was 65 V. The instrument was operated in a mass scan mode from m/z 100 to 1700.

Platform 4: Intact IgGs analyzed by mAb-Glyco-chip-Q-TOF MS. The same instruments as C18 chip-Q-TOF MS were used. MAb-Glyco chip (G4240-64020) used in this platform is designed as an inner/outer rotor device and composed of three columns. One column is 310 nL.
enzyme reactor packed with immobilized PNGase F beads, which can be switched into or out of the CapPump flow path for filling sample, deglycosylation, and transferring digested samples to enrichment column. Manufacture optimized 4-min minute incubation of antibody in the enzyme reactor was used. Another two columns are a sample enrichment column (160 nL) and a separation column (43 mm). Both columns are packed with porous graphitized carbon. The gradient for deglycosylation and glycan analysis is as follows: 2% B in the first 6 min, 2-32% B in 1.5 min, 32-85% B in 30 s. After a hold at 85% for 1 min, gradient quickly drop back to 2% B in 0.01s and followed by 3 min balance in 2% B. Deglycosylation buffer (3 μL/min) was used on CapPump. Other mass spectrometer parameters are set as the same as PGC-chip-Q-TOF analysis of off-line released Glycans. Each sample was diluted into 0.5 μg/μL by using deglycosylation buffer and 4 μL diluted sample was analyzed each time.

5.3.5 Data analysis

Firstly, theoretical masses of IgGs glycopeptides (Tables 5-1 and 5-2, including 31 usual glycans on human IgGs) were developed. Then the extracted ion chromatograms (EIC) were plotted using a ± 50 ppm mass window centered on the theoretical m/z of each IgG glycopeptide. A final assignment of the most probable glycan within each EIC was obtained by manual inspection of the observed m/z and corresponding charge state of each IgG glycopeptide and preferred fragmentation patterns obtained in the corresponding MS/MS spectra. Once the glycopeptides were assigned in one run, the mass, charge state and retention time of precursor ion of each glycopeptide were used to assign glycopeptides in other runs. Since triply charged glycopeptides were most abundant, the masses (m/z) of triply charged glycopeptides within 50 ppm mass window were extracted and the peak area of the EIC with a S/N > 2 was considered for
manual quantitation.

Glycan data was performed using the Agilent MassHunter software version B.02. The structures of glycans were identified by searching against glycan database coming along with the mAb-Glyco chip kit. The extracted ion chromatograms (EIC) of identified glycans were plotted using 50 ppm mass window. Peak area with a S/N>2 is considered for final quantitation.

The relative abundances of glycopeptides were normalized to each subclass IgG with G1 glycoform. The relative abundances of glycans were normalized to G1. Each digest or sample was analyzed three times.

5.4 Results and Discussion

5.4.1 SDS-PAGE Result of IgG isolated from HIV-1 patients

Serum IgGs from 4 different patients with HIV-1 infection (two still healthy and two with progressive weakness) were purified by the Melon Gel IgG Purification Kit. The isolated IgGs were separated by 1D SDS-PAGE and stained by Coomassie blue. As shown in the Figure 5-2, the amounts of IgGs (both heavy and light chains) were consistent for four different samples together with a significant amount of albumin, which could be related to either overloading effects or the binding of albumin to IgG either due to specific or non-specific interactions.\textsuperscript{17} The gels also showed small amounts of high molecular weight material which could be related to IgG aggregates. The gel bands containing the expected IgG heavy chain (in red square) were cut out and destained for in-gel enzymatic digestion (trypsin, Lys-C, or PNGase F).
Figure 5-1. IgG and Glycan structure.

Figure 5-2. SDS-PAGE Analysis of IgG isolated from HIV-1 patients.
5.4.2 LC-MS method development

In this study we have reported on the evaluation of four analytical platforms for the analysis of IgG N-glycosylation for possible use in the study of serum samples collected from AIDS patients with the view of identifying patients at risk. We first explored the following aspects in development of the methods by analysis of four patient samples in triplicate: performance for accuracy and sensitivity, reproducibility, possible throughput and type of information. We will discuss the clinical parameters of the patient samples later (section 5.4.3) and in this section we will concentrate on the performance of these four analysis platforms.

5.4.2.1 Glycopeptide analysis

Tryptic glycopeptide analysis (platform 1)

Trypsin digestion of IgGs usually generates a hydrophilic glycopeptide containing 9 amino acids (E<sub>293</sub>EQX<sub>296</sub>NSTX<sub>300</sub>R<sub>301</sub>. X<sub>296,300</sub> = tyrosine (Y) in IgG1; X<sub>296,300</sub> = phenylalanine (F) in IgG2; X<sub>296</sub> = tyrosine (Y), X<sub>300</sub> = phenylalanine (F) in IgG3, X<sub>296</sub> = phenylalanine (F), X<sub>300</sub> = tyrosine (Y) in IgG4) which elutes early in the HPLC separation. We examined the analysis of tryptic digests on platform 1: the FTCIR MS with a Dionex nano-flow HPLC platform. Figure 5-3 shows an example of the identification of the family of glycopeptides IgG1 present at Asn<sub>297</sub>. The overlap of XIC of IgGs with G1 (Figure 5-3A) showed that IgG1 glycopeptides (…Y<sub>296</sub>…Y<sub>300</sub>…) eluted first, around 19.6 min, followed by IgG3/4 subclasses (22.4 min; …F<sub>296</sub>…Y<sub>300</sub>…for IgG3 and …Y<sub>296</sub>…F<sub>300</sub>…for IgG4) and IgG2 glycopeptides were eluted last (25.9 min; …F<sub>296</sub>…F<sub>300</sub>…). The separation achieved by the Magic C18 column was excellent with the elution sequence corresponding to the following amino acid substitutions. The peak areas of these glycopeptides are in accordance with the relative amount of each IgG in
normal blood with IgG1 > IgG2 > IgG3/IgG4. However, glycan variants within one glycopeptide family were rarely separated by reversed phase HPLC because their retention times are largely determined by the peptide sequence. For example, IgG1 with G0, G1, G2, and G2-F eluted around 19.6 min (Figure 5-3B). Despite this lack of separation, the accurate mass and charge state in the FT MS scan confirmed the parent ion mass of all glycopeptides (Figure 5-3B). The glycan composition was determined via the presence of abundant product ions derived from glycosidic bond cleavages in MS2 by CID, as an example (Figure 5-3C) shows the CID analysis of the G1 (see structure) form with the loss of N-GlcNAc, mannose and galactose.

Figure 5-3. Tryptic glycopeptide analysis by LTQ-FT MS. A. Extracted ion chromatograms (XIC) of IgG1, IgG2, and IgG3/4; B. Precursor ion scan of IgG1 using FTMS; C. CID-MS/MS of the precursor ion m/z 933.0413 (3+) in B.
**Lys-C glycopeptide analysis (platform 2)**

Although nanoHPLC-FTICR MS is a robust technique for glycopeptide analysis, chip-Q-TOF MS has been reported to be more sensitive (attomole for Q-TOF compared to femtomole for LTQ-FT) and potentially high-throughput. In addition, reproducible results are difficult to obtain by the nanoflow HPLC used with the LTQ-FTICR due to intricate fittings and connections. In the case of Q-TOF, we use an HPLC-Chip that integrates the nanoflow-LC components and spray tips onto a single chip and thus minimizes dead volumes and the possibility of leaks and therefore may increase chromatography efficiency. Furthermore, the integration simplifies the workflow and makes high-throughput glycoproteomic analysis possible. We, therefore, explored the use of the C18-chip-Q-TOF MS (platform 2) for analysis of the Lys-C IgG glycopeptides (see Figure 5-4A for the corresponding IgG-G1). The glycopeptide produced by Lys-C digestion is longer than the tryptic glycopeptide \( (T_{289}KPREEQXNSTXRVVSLTVXHQDWLNGK_{317}) \) and includes more hydrophobic amino acids (underlined), which resulted in stronger binding to the C18 column used in the microfluidic chip separation where we used a higher loading flow (3 \( \mu \)L/min vs. 200 nL/min in nano RPHPLC-FTICR MS). A disadvantage of this strategy is that the Lys-C but not the shorter tryptic peptide contains Asn\(_{315}\) which is readily deamidated during the enzymatic digestion. The split peak around 17.5 min was due to deamidation that occurred on Asn\(_{315}\) \( (T_{289}KPREEQXNSTXRVVSLTVXHQDWLNGK_{317}) \). Another disadvantage of the fast separation gradient used for Lys-C glycopeptide analysis is that it resulted in four-subclass IgG glycopeptides being eluted around a similar time.

For estimation of the glycan amounts we integrated the peak area for both the Asn and Asp forms of an individual Lys-C glycopeptide (the isoAsp variant was not resolved in this
separation). Similar to tryptic glycopeptides analysis by HPLC/ESI-FTICR MS, the accurate mass and the MS/MS fragments identified the glycopeptide, the glycan composition, and the glycosylation site (Figure 5-4B and 5-4C).

Figure 5-4. Lys-C glycopeptide analysis by Chip Q-TOF MS.
A. Extracted ion chromatograms (XIC) of IgG1, IgG2, and IgG3/4; B. Precursor ion scan of IgG1, IgG2 and IgG3/4 using TOF; C. CID-MS/MS of the precursor ion m/z 1014.8954 (5+) in B.
5.4.2.2 Cleaved Glycan Analysis

In this report, two platforms (platform 3 and 4) were used to profile cleaved glycans present in IgGs (IgG1-4). Platform 3 is based on offline release of glycans from IgGs by PNGase F. The free glycans were analyzed by a PGC chip coupled to Q-TOF MS. In platform 4, mAb-Glyco-chip-Q-TOF MS uses an online PNGase F release of glycans followed by PGC separation and Q-TOF MS analysis. In both analyses, we used MassHunter Qualitative Analysis Software to provide automated glycan identification and quantitation. Since the goal was to develop a high-throughput approach for glycan analysis we used only MS full scan for a better quantitation due to an improved peak shape in the MS mode. The glycan identification was based on measured accurate mass (within a 50 ppm window around the theoretical m/z of each glycan).

We then used targeted MS/MS analysis to confirm some glycan structures that are not found in glycopeptide analysis. For example, G2+2SA-F which was identified only by the glycan and not the glycopeptide analysis platforms, was confirmed by the targeted MS/MS result (see Figure 5-5).

Figure 5-5. Targeted MS/MS of m/z 1112.41 (2+), which is corresponding to glycan G2+2SA-F.
5.4.2.3 Comparison of results from 4 analytical approaches

The identified glycopeptides and glycans by these four platforms are listed in Tables 5-3, 5-4, 5-5 and 5-6. As is shown in Tables 5-3 and 5-4, both glycopeptide analysis platforms are able to differentiate glycan species on the glycopeptides isolated from different IgG subclasses and provided similar IgG subclass-specific profiles with similar amount of variation. However, tryptic peptides followed by C18 nanoHPLC- FTICR MS identified 13 IgG1 glycopeptides, 9 IgG2 glycopeptides and 6 IgG3/IgG4 glycopeptides. As we discussed previously, due to a similar elution time of a complex set of glycopeptides caused by higher flow rate, together with deamidated forms induced by the Lys-C digestion step, Lys-C peptides analyzed by C18-chip-Q-TOF identified fewer IgG glycopeptides and resulted in a somewhat higher RSD% for the glycan measurement (see Tables 5-3 and 5-4). The variations in peak area of both platforms for the high abundance glycopeptides (such as IgG1 with G0, G1 and G2, IgG2 with G0, IgG3/IgG4 with G0) are mostly less than 10% and 5%.

Both glycan analysis platforms (3 and 4) identified more glycoforms than the glycopeptide analysis platforms (1 and 2). While the glycan analysis platform (3 and 4) identified more glycan structures than the glycopeptide platform (1 and 2), it is important to note that potentially different sets of glycans were analyzed. IgG is N-glycosylated not only in the Fc regions but also in the variable regions of heavy and light chains. Furthermore, the types of N-glycans from variable regions are different from Fc N-glycans. Glycopeptide analysis specifically characterizes the Fc region glycans since Fc region peptide sequence information is also included for identification. In contrast, glycan analysis profiles total glycans from both variable and Fc regions. PNGase F off-line digests analyzed by PGC-chip-Q-TOF MS profiled the glycans from IgG Fc region since heavy chain of IgG was separated from light chain by
SDS-PAGE and only heavy chain was cut out and digested by PNGase F. On the other hand, mAb-Glyco-chip-Q-TOF MS directly analyzed samples and no separation of heavy and light chains before digestion. Thus, in addition to the glycans from IgG Fc region, glycans from light chain and variable region of the heavy chain were also identified and included in the total glycan profile results. Consequently, the most glycans (20 glycans) identified by PGC-chip-Q-TOF MS (platform 3) and a similar number of glycans (18) characterized by mAb-Glyco-chip-Q-TOF MS (platform 4) is a necessary result. Because the Fc region glycans account for 80-85% of total glycans on IgGs and the glycans from variable regions are structurally different from Fc glycans, however, the distribution of high abundance glycans from glycan analysis should be comparable with glycopeptide analysis in which only Fc glycans were analyzed.

The reproducibility of analysis is better for glycan than glycopeptide analysis with a RSD% for platform 3 and 4 of approximately 4% for high abundance glycans (G0, G1 and G2) and 10% for low abundance glycans vs. platform 1 and 2 of 5% and 10% for high abundance and up to 30% and 20% for low abundance glycans.

The advantage of glycopeptide analysis by LC-MS is that one method can allow site-specific analysis of glycosylation, simultaneous analysis of other peptide sequences, and posttranslational modifications. In the case of IgGs the site of glycosylation, Asn297 is known and the peptide sequence can be used to determine the mass of different glycans attached to this site (see Figure 5-3 for an example). The LC/MS analysis, however, is long; usually 1-1.5 hour is necessary for each sample. Besides, the digestion process is time-consuming and difficult to standardize between laboratories. In comparison, glycan analysis is a fast approach to structural analysis of glycans, where the released glycan can be separated by HPLC and then analyzed in the mass spectrometer. In this manner, glycan analysis provides a “glycan profile” of antibody
and no amino acid sequence information is required for data analysis. Since only glycans are
analyzed, the overall analysis time can be very short. However, glycosylation sites are not
determined by this approach, and this analysis would require the separation of IgG subclasses by
affinity chromatography, for example protein A binds IgG1, 2, and 4 while protein G binds only
IgG3.

5.4.3 Comparison of the IgG glycosylation profiles obtained for four different patients with
HIV-1 infection by four LC-MS platforms

Although all four patients were HIV-1 infected, patients B and D are still healthy and show high ADCC activity with higher binding capability to Fc-receptor and higher percentage of
NK cells that are activated (Table 5-7). On the contrary, patients A and C are progressively weak.
The IgG binding capability to Fc receptor is medium in patient A but low in patient C.
Accordingly, patient A shows mild inhibitory effect on HIV but on inhibition is shown in patient
C. It has been reported that the Fc glycosylation has a significant influence of IgG effector
function through changing the binding level of IgG to Fc receptor.\textsuperscript{3, 7-9, 23} Therefore, the IgGs Fc
glycosylation from patients A, B, C and D were detailed characterized.

As shown in Tables 5-3, 5-4, 5-5 and 5-6 and Figures 5-6, 5-7, 5-8 and 5-9, the lower
degrees of galactosylation (higher G0, lower G2) and higher level of defucosylated IgG1 species
(Gx-F) are obviously seen in patients B. But similar levels of IgG1 with G0, G2, G0-F and G1-F
were observed between patient A and D in both glycopeptide analysis platforms. In addition,
significantly lower levels of galactosylation on IgG1, IgG2 and IgG4 in HIV-1 infected
individuals have been reported.\textsuperscript{24} The research also showed that CDC but not ADCC activity is
affected by galactosylation level on IgGs.\textsuperscript{6} Therefore, IgG galactosylation and defucosylation
levels might not be involved in modification of IgG binding affinity to Fc receptor.

On the contrary, the elevated levels of bisecting N-acetylglucosamine (Gx+GlcNAc) observed in patient B and D but not in patient A and C in all four analysis platforms suggest that higher level of bisecting N-acetylglucosamine might cause more IgGs binding to Fc receptors. Accordingly, higher ADCC activity was triggered and HIV-1 infection was inhibited. And the connection between higher level of bisecting N-acetylglucosamine and increased ADCC has been reported in vitro. The increased levels of bisecting N-acetylglucosamine on IgG1 and IgG2 have been found in lambert-eaton myasthenic syndrome patients below 50 years.

5.5 Conclusions

Four LC-MS platforms are successfully applied for integrated analysis of glycans or glycopeptides on IgGs from HIV-1 infected patients and generated similar IgG glycosylation distribution. Among these four LC-MS platforms, PNGase F released glycans followed by PGC-chip QTOF identified the most glycans. The reproducibilities of four platforms are good with CV% of high abundant glycans less than or around 5% but two glycan analysis platforms generally produce better CV%. However, glycopeptide analysis platforms are more site-specific and the glycopeptides from three IgG subclasses can be differentiated by these analysis platforms. Furthermore, in addition to glycosylation information, peptide sequences and other posttranslational modifications are simultaneously obtained during glycopeptide analysis. Compared to glycopeptide analysis platforms, both glycan analysis platforms are high-throughput with only 4 minutes separation but similar total glycosylation profiles were produced. The novel mAb-Glyco-chip QTOF platform generates similar results with off-line released glycans analyzed by PGC-chip QTOF but minimum sample preparation is required.
The different glycosylation profiles were observed between two healthy and two progressively weak HIV-infected patients by all analysis platforms. The results from these four methods consistently suggest that in HIV-1 infected patients, bisecting N-acetylgalactosamine (Gx+GlcNAc) might influence the binding of IgGs to Fc receptor. The elevated level of bisecting N-acetylgalactosamine might induce higher binding of IgGs to Fc receptor and increased ADCC activity. Therefore, HIV-1 infection is inhibited.
Figure 5-6. Glycosylation profile of IgG subclasses obtained by FTICR MS analysis of tryptic glycopeptides from four HIV-1 infected patients. 
(A) IgG1 glycosylation profile (B) IgG2 glycosylation profile (C) IgG3/4 glycosylation profile.
Figure 5-7. Glycosylation profile of IgG subclasses obtained by C18 chip-Q-TOF MS analysis of tryptic glycopeptides from four HIV-1 infected patients. (A) IgG1 glycosylation profile (B) IgG2 glycosylation profile (C) IgG3/4 glycosylation profile.
Figure 5-8. Total glycan profile obtained by PGC-chip QTOF MS analysis of PNGase F released glycans.

Figure 5-9. Total glycan profile obtained by mAb-Glyco-chip QTOF MS analysis of intact IgGs.
Table 5-1. Possible tryptic glycopeptides of IgG subclasses

<table>
<thead>
<tr>
<th>Glycan Name</th>
<th>Glycan Structurea</th>
<th>Monoisotopic mass of glycan (in hydroxyl form)</th>
<th>Monoisotopic mass of tryptic glycopeptide</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>IgG1</td>
<td>IgG2</td>
</tr>
<tr>
<td>Swiss-Prot entry number</td>
<td></td>
<td>P01857</td>
<td>P01859</td>
</tr>
<tr>
<td>Glycopeptide sequence</td>
<td></td>
<td>E293EQYNSTYR301</td>
<td>E293EQFNSTFR301</td>
</tr>
<tr>
<td>Monoisotopic mass of peptide (no glycan)</td>
<td></td>
<td>1188.5048</td>
<td>1156.5149</td>
</tr>
<tr>
<td>G0</td>
<td></td>
<td>1462.5444</td>
<td>2633.0624</td>
</tr>
<tr>
<td>G1</td>
<td></td>
<td>1624.5973</td>
<td>2795.1153</td>
</tr>
<tr>
<td>G2</td>
<td></td>
<td>1786.6501</td>
<td>2957.1681</td>
</tr>
<tr>
<td>G0 – F</td>
<td></td>
<td>1316.4865</td>
<td>2487.0045</td>
</tr>
<tr>
<td>G1-F</td>
<td></td>
<td>1478.5393</td>
<td>2649.0573</td>
</tr>
<tr>
<td>G2 - F</td>
<td></td>
<td>1640.5922</td>
<td>2811.1102</td>
</tr>
<tr>
<td>G0 + GlcNAc</td>
<td></td>
<td>1665.6238</td>
<td>2836.1418</td>
</tr>
<tr>
<td>G2 + GlcNAc</td>
<td></td>
<td>1989.7295</td>
<td>3160.2475</td>
</tr>
<tr>
<td>Structure</td>
<td>Assignments</td>
<td>1st Peak</td>
<td>2nd Peak</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>G1 + SA(^b)</td>
<td><img src="image1" alt="Structure" /></td>
<td>1915.6927</td>
<td>3086.2107</td>
</tr>
<tr>
<td>G2 + SA</td>
<td><img src="image2" alt="Structure" /></td>
<td>2077.7455</td>
<td>3248.2635</td>
</tr>
<tr>
<td>G2 + 2SA</td>
<td><img src="image3" alt="Structure" /></td>
<td>2368.8409</td>
<td>3539.3589</td>
</tr>
<tr>
<td>G0 + GlcNAc – F</td>
<td><img src="image4" alt="Structure" /></td>
<td>1519.5659</td>
<td>2690.0839</td>
</tr>
<tr>
<td>G1 + GlcNAc – F</td>
<td><img src="image5" alt="Structure" /></td>
<td>1681.6187</td>
<td>2852.1367</td>
</tr>
<tr>
<td>G2 + GlcNAc – F</td>
<td><img src="image6" alt="Structure" /></td>
<td>1843.6715</td>
<td>3014.1895</td>
</tr>
<tr>
<td>G1 + GlcNAc + SA</td>
<td><img src="image7" alt="Structure" /></td>
<td>2118.7720</td>
<td>3289.29</td>
</tr>
<tr>
<td>G2 + GlcNAc + SA</td>
<td><img src="image8" alt="Structure" /></td>
<td>2280.8249</td>
<td>3451.3429</td>
</tr>
<tr>
<td>G2 + GlcNAc + 2SA</td>
<td><img src="image9" alt="Structure" /></td>
<td>2571.9203</td>
<td>3742.4383</td>
</tr>
<tr>
<td>G1 + SA – F</td>
<td><img src="image10" alt="Structure" /></td>
<td>1769.6347</td>
<td>2940.1395</td>
</tr>
<tr>
<td>G2+ SA – F</td>
<td><img src="image11" alt="Structure" /></td>
<td>1947.6825</td>
<td>3118.2005</td>
</tr>
<tr>
<td>G2 + 2SA – F</td>
<td><img src="image12" alt="Structure" /></td>
<td>2222.7830</td>
<td>3393.3010</td>
</tr>
</tbody>
</table>

\(^a\) N-acetylglucosamine (GlcNAc); \(^b\) if not mentioned specifically, sialic acid residue(s) is as N-acetylneuraminic acid
<table>
<thead>
<tr>
<th>Glycan Name</th>
<th>Glycan Structure</th>
<th>Monoisotopic mass of glycan (in hydroxyl form)</th>
<th>Monoisotopic mass of Lys-C glycopeptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swiss-Protein entry number</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IgG1</td>
<td>IgG2</td>
<td>IgG3</td>
<td>IgG4</td>
</tr>
<tr>
<td>P01857</td>
<td>P01859</td>
<td>P01860</td>
<td>P01861</td>
</tr>
<tr>
<td>Glycopeptide sequence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T289KPREEQYNSTYRVVSVLTVDWLNK317</td>
<td>T289KPREEQFNS</td>
<td>T289KPREEQFNS</td>
<td>T289KPREEQFNS</td>
</tr>
<tr>
<td>YRVVSVLTVDWLNK317</td>
<td>HQDWLNGK317</td>
<td>HQDWLNGK317</td>
<td>HQDWLNGK317</td>
</tr>
<tr>
<td>Monoisotopic mass of peptide (no glycan)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G0</td>
<td>G1</td>
<td>G2</td>
<td>G0 - F</td>
</tr>
<tr>
<td>1462.5444</td>
<td>1624.5973</td>
<td>1786.6501</td>
<td>1316.4865</td>
</tr>
<tr>
<td>4904.3475</td>
<td>5066.4004</td>
<td>5228.4532</td>
<td>4758.2896</td>
</tr>
<tr>
<td>4858.3421</td>
<td>5020.3950</td>
<td>5182.4478</td>
<td>4712.2842</td>
</tr>
<tr>
<td>4888.3526</td>
<td>5050.4055</td>
<td>5212.4583</td>
<td>4742.2947</td>
</tr>
<tr>
<td>4888.3526</td>
<td>5050.4055</td>
<td>5212.4583</td>
<td>4742.2947</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>G1 + SA&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td>1915.6927</td>
<td>5357.4958</td>
</tr>
<tr>
<td>G2 + SA</td>
<td></td>
<td>2077.7455</td>
<td>5519.5486</td>
</tr>
<tr>
<td>G2 + 2SA</td>
<td></td>
<td>2368.8409</td>
<td>5810.6440</td>
</tr>
<tr>
<td>G0 + GlcNAc – F</td>
<td></td>
<td>1519.5659</td>
<td>4961.3690</td>
</tr>
<tr>
<td>G1 + GlcNAc – F</td>
<td></td>
<td>1681.6187</td>
<td>5123.4218</td>
</tr>
<tr>
<td>G2 + GlcNAc – F</td>
<td></td>
<td>1843.6715</td>
<td>5285.4746</td>
</tr>
<tr>
<td>G1 + GlcNAc + SA</td>
<td></td>
<td>2118.7720</td>
<td>5560.5751</td>
</tr>
<tr>
<td>G2 + GlcNAc + SA</td>
<td></td>
<td>2280.8249</td>
<td>5722.6280</td>
</tr>
<tr>
<td>G2 + GlcNAc + 2SA</td>
<td></td>
<td>2571.9203</td>
<td>6013.7234</td>
</tr>
<tr>
<td>G1 + SA – F</td>
<td></td>
<td>1769.6347</td>
<td>5211.4246</td>
</tr>
<tr>
<td>G2 + SA – F</td>
<td></td>
<td>1947.6825</td>
<td>5389.4856</td>
</tr>
<tr>
<td>G2 + 2SA – F</td>
<td></td>
<td>2222.7830</td>
<td>5664.5861</td>
</tr>
</tbody>
</table>

<sup>a</sup> N-acetylglicosamine (GlcNAc); • Mannose; ▲ Fucose (F); ◆ Sialic acid (SA); ● Galactose

<sup>b</sup> if not mentioned specifically, sialic acid residue(s) is as N-acetyleneuraminic acid
Table 5-3. IgG isotype-specific Fc glycosylation profiles obtained by FTICR MS analysis of tryptic glycopeptides from four HIV-1 infected patients.

<table>
<thead>
<tr>
<th>IgG subclass</th>
<th>Glycan type</th>
<th>Patient A</th>
<th>Patient B</th>
<th>Patient C</th>
<th>Patient D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G0</td>
<td>1.41(^b) ± 5.06(^c)</td>
<td>2.98 ± 6.83</td>
<td>0.84 ± 4.17</td>
<td>1.23 ± 4.40</td>
</tr>
<tr>
<td></td>
<td>G1(^a)</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>G2</td>
<td>0.36 ± 3.83</td>
<td>0.13 ± 3.95</td>
<td>0.37 ± 2.63</td>
<td>0.35 ± 6.09</td>
</tr>
<tr>
<td></td>
<td>G0-F</td>
<td>0.02 ± 5.96</td>
<td>0.49 ± 6.79</td>
<td>0.02 ± 18.09</td>
<td>0.08 ± 15.87</td>
</tr>
<tr>
<td></td>
<td>G1-F</td>
<td>0.03 ± 2.23</td>
<td>0.29 ± 21.95</td>
<td>0.05 ± 8.08</td>
<td>0.06 ± 12.51</td>
</tr>
<tr>
<td></td>
<td>G2-F</td>
<td>0.09 ± 17.64</td>
<td>0.14 ± 7.27</td>
<td>0.19 ± 20.15</td>
<td>0.03 ± 5.54</td>
</tr>
<tr>
<td></td>
<td>G0+GlcNAc</td>
<td>0.31 ± 3.34</td>
<td>1.50 ± 9.05</td>
<td>0.22 ± 9.17</td>
<td>0.82 ± 10.18</td>
</tr>
<tr>
<td></td>
<td>G1+GlcNAc</td>
<td>0.25 ± 4.26</td>
<td>0.45 ± 7.58</td>
<td>0.23 ± 16.19</td>
<td>0.55 ± 7.23</td>
</tr>
<tr>
<td></td>
<td>G2+GlcNAc</td>
<td>0.01 ± 19.51</td>
<td>ND(^d)</td>
<td>0.01 ± 29.85</td>
<td>0.03 ± 9.76</td>
</tr>
<tr>
<td></td>
<td>G0+GlcNAc-F</td>
<td>ND</td>
<td>0.17 ± 12.88</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>G1+GlcNAc-F</td>
<td>0.01 ± 12.53</td>
<td>0.17 ± 14.25</td>
<td>0.03 ± 12.18</td>
<td>0.03 ± 11.47</td>
</tr>
<tr>
<td></td>
<td>G1+SA</td>
<td>0.05 ± 10.66</td>
<td>0.06 ± 16.51</td>
<td>0.05 ± 22.51</td>
<td>0.06 ± 5.89</td>
</tr>
<tr>
<td></td>
<td>G2+SA</td>
<td>0.20 ± 9.00</td>
<td>0.09 ± 12.39</td>
<td>0.24 ± 27.39</td>
<td>0.20 ± 3.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G0</td>
<td>1.18 ± 3.28</td>
<td>2.64 ± 3.75</td>
<td>0.87 ± 3.01</td>
<td>1.69 ± 7.35</td>
</tr>
<tr>
<td></td>
<td>G1</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>G2</td>
<td>0.27 ± 11.51</td>
<td>ND</td>
<td>0.28 ± 10.25</td>
<td>0.32 ± 9.02</td>
</tr>
<tr>
<td></td>
<td>G1-F</td>
<td>ND</td>
<td>0.12 ± 10.90</td>
<td>0.04 ± 20.54</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>G2-F</td>
<td>ND</td>
<td>0.00 ±</td>
<td>0.11 ± 28.11</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>G0 + GlcNAc</td>
<td>0.26 ± 5.17</td>
<td>0.57 ± 17.00</td>
<td>0.23 ± 19.94</td>
<td>0.42 ± 11.03</td>
</tr>
<tr>
<td></td>
<td>G1 + GlcNAc</td>
<td>0.04 ± 6.87</td>
<td>0.11 ± 6.22</td>
<td>0.09 ± 9.10</td>
<td>0.36 ± 6.00</td>
</tr>
<tr>
<td></td>
<td>G1+SA</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>0.06 ± 11.58</td>
</tr>
<tr>
<td></td>
<td>G2+SA</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>0.08 ± 23.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G0</td>
<td>0.55 ± 3.81</td>
<td>2.82 ± 2.42</td>
<td>0.70 ± 4.94</td>
<td>1.93 ± 11.46</td>
</tr>
<tr>
<td></td>
<td>G1</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>G2</td>
<td>0.51 ± 13.28</td>
<td>0.08 ± 12.42</td>
<td>0.52 ± 12.31</td>
<td>0.20 ± 9.94</td>
</tr>
<tr>
<td></td>
<td>G0+GlcNAc</td>
<td>0.17 ± 6.56</td>
<td>1.31 ± 5.56</td>
<td>0.12 ± 18.10</td>
<td>1.20 ± 8.63</td>
</tr>
<tr>
<td></td>
<td>G1+GlcNAc</td>
<td>0.14 ± 18.25</td>
<td>0.33 ± 5.30</td>
<td>0.16 ± 12.19</td>
<td>0.50 ± 8.00</td>
</tr>
<tr>
<td></td>
<td>G2+SA</td>
<td>0.15 ± 17.81</td>
<td>ND</td>
<td>0.11 ± 21.22</td>
<td>ND</td>
</tr>
</tbody>
</table>

\(^a\) Normalized to IgGx with G1; \(^b\) Average ratio to IgGx with G1; \(^c\) RSD\%; \(^d\) ND means not detected
Table 5-4. IgG isotype-specific Fc glycosylation profiles obtained by QTOF analysis of Lys-C glycopeptides from four HIV-1 infected patients.

<table>
<thead>
<tr>
<th>IgG subclass</th>
<th>Glycan type</th>
<th>Patient A</th>
<th>Patient B</th>
<th>Patient C</th>
<th>Patient D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G0</td>
<td>1.57 ± 12.43</td>
<td>4.8 ± 8.42</td>
<td>1.3 ± 2.50</td>
<td>1.7 ± 3.23</td>
</tr>
<tr>
<td></td>
<td>G1</td>
<td>1.00 ± 0.00</td>
<td>1.0 ± 0.00</td>
<td>1.0 ± 0.00</td>
<td>1.0 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>G2</td>
<td>0.35 ± 12.62</td>
<td>0.2 ± 1.90</td>
<td>0.4 ± 2.08</td>
<td>0.3 ± 5.50</td>
</tr>
<tr>
<td></td>
<td>G0-F</td>
<td>ND</td>
<td>1.2 ± 10.2</td>
<td>ND</td>
<td>0.0 ± 4.31</td>
</tr>
<tr>
<td></td>
<td>G0+GlcNAc</td>
<td>0.29 ± 12.71</td>
<td>1.2 ± 10.5</td>
<td>0.2 ± 3.31</td>
<td>0.5 ± 4.59</td>
</tr>
<tr>
<td></td>
<td>G1+GlcNAc</td>
<td>0.16 ± 13.28</td>
<td>0.4 ± 11.5</td>
<td>0.1 ± 6.00</td>
<td>0.3 ± 3.41</td>
</tr>
<tr>
<td></td>
<td>G2+GlcNAc</td>
<td>ND</td>
<td>ND</td>
<td>0.0 ± 12.7</td>
<td>0.0 ± 5.25</td>
</tr>
<tr>
<td></td>
<td>G1+SA</td>
<td>0.05 ± 10.12</td>
<td>0.0 ± 16.5</td>
<td>0.0 ± 2.97</td>
<td>0.0 ± 3.57</td>
</tr>
<tr>
<td></td>
<td>G2+SA</td>
<td>0.17 ± 11.58</td>
<td>0.1 ± 16.8</td>
<td>0.2 ± 3.70</td>
<td>0.1 ± 5.09</td>
</tr>
<tr>
<td></td>
<td>G0</td>
<td>1.59 ± 3.00</td>
<td>2.6 ± 2.42</td>
<td>1.1 ± 2.01</td>
<td>2.3 ± 2.66</td>
</tr>
<tr>
<td></td>
<td>G1</td>
<td>1.00 ± 0.00</td>
<td>1.0 ± 0.00</td>
<td>1.0 ± 0.00</td>
<td>1.0 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>G2</td>
<td>0.30 ± 4.08</td>
<td>ND</td>
<td>0.3 ± 4.75</td>
<td>0.3 ± 9.36</td>
</tr>
<tr>
<td></td>
<td>G2+SA</td>
<td>0.11 ± 0.97</td>
<td>ND</td>
<td>0.1 ± 5.18</td>
<td>0.1 ± 3.69</td>
</tr>
<tr>
<td></td>
<td>G0</td>
<td>0.83 ± 0.53</td>
<td>3.1 ± 21.1</td>
<td>1.1 ± 2.59</td>
<td>1.6 ± 2.02</td>
</tr>
<tr>
<td></td>
<td>G1</td>
<td>1.00 ± 0.00</td>
<td>1.0 ± 0.00</td>
<td>1.0 ± 0.00</td>
<td>1.0 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>G2</td>
<td>0.43 ± 7.18</td>
<td>ND</td>
<td>0.4 ± 8.07</td>
<td>0.2 ± 2.56</td>
</tr>
<tr>
<td></td>
<td>G0+GlcNAc</td>
<td>ND</td>
<td>0.8 ± 18.2</td>
<td>0.3 ± 7.33</td>
<td>0.5 ± 12.6</td>
</tr>
<tr>
<td></td>
<td>G2+SA</td>
<td>0.48 ± 8.15</td>
<td>0.2 ± 14.8</td>
<td>0.2 ± 10.8</td>
<td>1.0 ± 7.65</td>
</tr>
</tbody>
</table>

a Normalized to IgGx with G1; b Average ratio to IgGx with G1; c RSD%; d ND means not detected
Table 5-5. Total glycan profile obtained by PGC-chip QTOF MS analysis of PNGase F released glycans

<table>
<thead>
<tr>
<th>Glycan type</th>
<th>Patient A</th>
<th>Patient B</th>
<th>Patient C</th>
<th>Patient D</th>
</tr>
</thead>
<tbody>
<tr>
<td>G0</td>
<td>1.14 ± 1.14</td>
<td>2.44 ± 1.52</td>
<td>0.80 ± 1.85</td>
<td>1.47 ± 3.20</td>
</tr>
<tr>
<td>G1</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
</tr>
<tr>
<td>G2</td>
<td>0.38 ± 3.21</td>
<td>0.23 ± 3.42</td>
<td>0.45 ± 2.06</td>
<td>0.37 ± 3.69</td>
</tr>
<tr>
<td>G0-F</td>
<td>0.01 ± 13.83</td>
<td>0.11 ± 4.03</td>
<td>0.01 ± 8.32</td>
<td>0.03 ± 11.27</td>
</tr>
<tr>
<td>G1-F</td>
<td>0.03 ± 8.63</td>
<td>0.67 ± 4.35</td>
<td>0.16 ± 0.27</td>
<td>0.05 ± 3.83</td>
</tr>
<tr>
<td>G2-F</td>
<td>0.02 ± 8.89</td>
<td>0.27 ± 2.49</td>
<td>0.19 ± 2.68</td>
<td>0.03 ± 3.65</td>
</tr>
<tr>
<td>G0+GlcNAc</td>
<td>0.10 ± 1.40</td>
<td>0.43 ± 4.18</td>
<td>0.07 ± 4.04</td>
<td>0.27 ± 9.17</td>
</tr>
<tr>
<td>G1+GlcNAc</td>
<td>0.09 ± 1.33</td>
<td>0.22 ± 1.29</td>
<td>0.09 ± 0.73</td>
<td>0.20 ± 7.26</td>
</tr>
<tr>
<td>G0+GlcNAc-F</td>
<td>0.01 ± 12.93</td>
<td>0.06 ± 6.65</td>
<td>0.01 ± 6.46</td>
<td>ND</td>
</tr>
<tr>
<td>G1+GlcNAc-F</td>
<td>0.01 ± 7.99</td>
<td>0.18 ± 4.25</td>
<td>0.03 ± 8.37</td>
<td>ND</td>
</tr>
<tr>
<td>G2+GlcNAc-F</td>
<td>ND</td>
<td>0.08 ± 0.46</td>
<td>0.02 ± 7.72</td>
<td>ND</td>
</tr>
<tr>
<td>G1+SA</td>
<td>0.03 ± 7.03</td>
<td>0.03 ± 13.44</td>
<td>0.04 ± 0.88</td>
<td>0.03 ± 4.43</td>
</tr>
<tr>
<td>G2+SA</td>
<td>0.15 ± 3.57</td>
<td>0.09 ± 5.76</td>
<td>0.14 ± 4.58</td>
<td>0.10 ± 5.80</td>
</tr>
<tr>
<td>G2-F+SA</td>
<td>0.57 ± 3.08</td>
<td>1.53 ± 3.91</td>
<td>0.59 ± 4.54</td>
<td>0.40 ± 8.44</td>
</tr>
<tr>
<td>G2+2SA</td>
<td>0.13 ± 8.15</td>
<td>0.09 ± 4.54</td>
<td>0.06 ± 5.94</td>
<td>0.07 ± 4.81</td>
</tr>
<tr>
<td>G2+GlcNAc+SA</td>
<td>0.03 ± 3.93</td>
<td>0.07 ± 4.64</td>
<td>0.03 ± 5.95</td>
<td>0.03 ± 17.90</td>
</tr>
<tr>
<td>G2+GlcNAc+2SA</td>
<td>0.11 ± 4.93</td>
<td>0.15 ± 2.09</td>
<td>0.11 ± 1.57</td>
<td>0.09 ± 10.87</td>
</tr>
<tr>
<td>G1-F+SA</td>
<td>0.02 ± 6.56</td>
<td>0.04 ± 9.07</td>
<td>0.02 ± 1.55</td>
<td>0.02 ± 13.39</td>
</tr>
<tr>
<td>G2-F+2SA</td>
<td>0.37 ± 4.72</td>
<td>0.87 ± 5.76</td>
<td>0.43 ± 3.48</td>
<td>0.43 ± 17.41</td>
</tr>
</tbody>
</table>

*a Normalized to IgGx with G1;  b Average ratio to IgGx with G1;  c RSD%;  d ND means not detected*
Table 5-6. Total glycan profile obtained by mAb-Glyco-chip QTOF MS analysis of intact IgGs.

<table>
<thead>
<tr>
<th>Glycan</th>
<th>Patient A</th>
<th>Patient B</th>
<th>Patient C</th>
<th>Patient D</th>
</tr>
</thead>
<tbody>
<tr>
<td>G0</td>
<td>1.28 ± 6.43°</td>
<td>2.66 ± 3.88</td>
<td>0.88 ± 1.18</td>
<td>1.56 ± 3.18</td>
</tr>
<tr>
<td>G1°</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
<td>1.00 ± 0.00</td>
</tr>
<tr>
<td>G2</td>
<td>0.30 ± 5.20</td>
<td>0.19 ± 5.78</td>
<td>0.24 ± 2.04</td>
<td>0.19 ± 1.94</td>
</tr>
<tr>
<td>G0-F</td>
<td>0.02 ± 4.54</td>
<td>0.27 ± 1.62</td>
<td>0.02 ± 3.64</td>
<td>0.05 ± 2.62</td>
</tr>
<tr>
<td>G1-F</td>
<td>0.04 ± 6.53</td>
<td>0.12 ± 5.19</td>
<td>0.02 ± 5.27</td>
<td>0.05 ± 6.22</td>
</tr>
<tr>
<td>G2-F</td>
<td>0.04 ± 1.69</td>
<td>0.05 ± 12.74</td>
<td>0.03 ± 5.41</td>
<td>0.03 ± 7.66</td>
</tr>
<tr>
<td>G0+GlcNAc</td>
<td>0.21 ± 7.55</td>
<td>0.77 ± 3.09</td>
<td>0.10 ± 4.55</td>
<td>0.39 ± 4.10</td>
</tr>
<tr>
<td>G1+GlcNAc</td>
<td>0.13 ± 4.82</td>
<td>0.25 ± 8.18</td>
<td>0.09 ± 4.06</td>
<td>0.19 ± 3.07</td>
</tr>
<tr>
<td>G2+GlcNAc</td>
<td>0.02 ± 2.06</td>
<td>0.02 ± 8.95</td>
<td>0.01 ± 8.12</td>
<td>0.02 ± 6.83</td>
</tr>
<tr>
<td>G0+GlcNAc-F</td>
<td>0.00 ± 10.94</td>
<td>0.12 ± 5.06</td>
<td>0.00 ± 7.35</td>
<td>0.02 ± 6.01</td>
</tr>
<tr>
<td>G1+GlcNAc-F</td>
<td>0.01 ± 10.21</td>
<td>0.05 ± 2.18</td>
<td>0.00 ± 4.32</td>
<td>0.02 ± 6.40</td>
</tr>
<tr>
<td>G1+SA</td>
<td>0.04 ± 4.45</td>
<td>0.05 ± 6.91</td>
<td>0.03 ± 4.05</td>
<td>0.04 ± 5.25</td>
</tr>
<tr>
<td>G2+SA</td>
<td>0.16 ± 5.86</td>
<td>0.09 ± 2.37</td>
<td>0.11 ± 2.93</td>
<td>0.11 ± 2.99</td>
</tr>
<tr>
<td>G2-F+SA</td>
<td>0.02 ± 2.50</td>
<td>0.08 ± 13.46</td>
<td>0.03 ± 14.79</td>
<td>0.08 ± 11.52</td>
</tr>
<tr>
<td>G2+2SA</td>
<td>0.13 ± 10.82</td>
<td>0.08 ± 5.49</td>
<td>0.05 ± 2.53</td>
<td>0.04 ± 4.73</td>
</tr>
<tr>
<td>G2+GlcNAc+2SA</td>
<td>0.06 ± 10.73</td>
<td>0.06 ± 4.37</td>
<td>0.04 ± 4.14</td>
<td>0.01 ± 14.85</td>
</tr>
<tr>
<td>G2-F+2SA</td>
<td>0.22 ± 5.36</td>
<td>0.65 ± 5.21</td>
<td>0.17 ± 4.43</td>
<td>0.36 ± 7.02</td>
</tr>
</tbody>
</table>

° Normalized to IgGx with G1; ° Average ratio to IgGx with G1; ° RSD%

Table 5-7. Patients Information

<table>
<thead>
<tr>
<th>Patient</th>
<th>µg/µL (protein assay)</th>
<th>% NK cells that are activated by the antibodies</th>
<th>Binds Fc-Receptor (RIIIα)</th>
<th>Inhibitory effect on HIV</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>7.94</td>
<td>4.10%</td>
<td>Medium (80 Rus)</td>
<td>Mild</td>
</tr>
<tr>
<td>B</td>
<td>8.31</td>
<td>38%</td>
<td>High (140 Rus)</td>
<td>Total inhibition</td>
</tr>
<tr>
<td>C</td>
<td>9.2</td>
<td>6%</td>
<td>Low (40 Rus)</td>
<td>None</td>
</tr>
<tr>
<td>D</td>
<td>7.84</td>
<td>41.4%</td>
<td>High (160 Rus)</td>
<td>Total inhibition</td>
</tr>
</tbody>
</table>
5.6 References


17. Chaudhury, C.; Mehnaz, S.; Robinson, J. M.; Hayton, W. L.; Pearl, D. K.; Roopenian,


25. Umana, P.; Jean-Mairet, J.; Moudry, R.; Amstutz, H.; Bailey, J. E. Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular
Future Work

This thesis has demonstrated the power of LC-MS for characterization of antibody structure and isoforms. Further directions for this area research could include:

1. Apply our LC-MS with ETD method, which is described in Chapter 2, to characterize more complicated disulfide bond, for example one peptide with multiple disulfide bonds and disulfide knot.

2. N-linked glycosylation analysis of protein drug (not mAb) by mAb-Glyco-chip MS method can be investigated and could be another application of this method.

3. Since only four patient samples were analyzed in Chapter 5, the results should be further confirmed by LC-MS analysis of more samples or by other analytical approach, e.g. lectin affinity assay.

4. Glycosylation plays an important role in a lot of biological activities, such as protein folding, subcellular localization, cell-cell interaction. The changes of glycosylation also correlate with cancer and other diseases. The four LC-MS analysis platforms in Chapter 5 can be used to resolve other biological questions and discover any quantitative change of glycosylation between normal and disease sample. This would be very useful strategy for the detection of new potential diagnostic markers and therapeutic targets.